

ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ
ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ ΔΙΟΙΚΗΣΗ»
ΤΜΗΜΑΤΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ



AdsRec:

Σύστημα Συστάσεων Διαφημίσεων για τα Μέσα Κοινωνικής Δικτύωσης

**Διπλωματική Εργασία της
Έλενας Μισιργή (ΑΕΜ: 452)**

Εξεταστική Επιτροπή

Επιβλέπων: κ. Συμεωνίδης Παναγιώτης

Μέλη: κ. Βασιλειάδης Νικόλαος
κ. Πετρίδου Ευγενία

**ΘΕΣΣΑΛΟΝΙΚΗ
ΣΕΠΤΕΜΒΡΙΟΣ 2015**

Πρόλογος

Το θέμα που πραγματεύεται η συγκεκριμένη εργασία είναι τα συστήματα συστάσεων (recommender systems) που έχουν αναπτυχθεί ιδιαίτερα το τελευταίο διάστημα στο διαδίκτυο. Αποφασίσαμε να μελετήσουμε σε βάθος τα συστήματα αυτά, αλλά και να αναπτύξουμε το δικό μας σύστημα συστάσεων. Έτσι δημιουργήθηκε το AdsRec (<http://ads.rec.xelixis.net/>), ένας δικτυακός τόπος που παρέχει στους χρήστες του συστάσεις διαφημιστικών banners καθώς αυτοί ενημερώνονται για διάφορα θέματα τα οποία τους αφορούν.

Στο σημείο αυτό θα ήθελα να ευχαριστήσω ιδιαίτερα τον κ. Παναγιώτη Συμεωνίδη για το άκρως ενδιαφέρον θέμα που διαλέξαμε για την εκπόνηση της διπλωματικής μου εργασίας, αλλά και για την πολύτιμη βοήθεια του καθ' όλη τη διάρκεια υλοποίησης της. Ήταν πάντα δίπλα μου όταν τον χρειαζόμουν και μου έδινε κίνητρα για να προχωρήσουμε και να φτάσουμε πλέον στο τελικό αυτό αποτέλεσμα.

Τέλος, θα ήθελα να ευχαριστήσω όλους τους καθηγητές στο Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών «Πληροφορική & Διοίκηση» για τις γνώσεις που μου μετέφεραν κατά τη διάρκεια των σπουδών μου, καθώς και όλους τους δικούς μου ανθρώπους που ήταν δίπλα μου και χωρίς τη βοήθεια τους όλα θα φάνταζαν μεγαλύτερης δυσκολίας.

Μισιργή Έλενα

10/9/2015

Περιεχόμενα

ΠΡΟΛΟΓΟΣ	I
ΠΕΡΙΕΧΟΜΕΝΑ.....	III
ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ	VI
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ.....	VIII
ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ.....	ΣΦΑΛΜΑ! ΔΕΝ ΕΧΕΙ ΟΡΙΣΤΕΙ ΣΕΛΙΔΟΔΕΙΚΤΗΣ.
1 ΕΙΣΑΓΩΓΗ	1
1.1 ΣΚΟΠΟΣ ΕΡΓΑΣΙΑΣ	1
1.2 ΔΟΜΗ ΤΗΣ ΕΡΓΑΣΙΑΣ	2
2 ΔΙΑΦΗΜΙΣΗ ΣΤΟ ΔΙΑΔΙΚΤΥΟ.....	5
2.1 ΔΙΑΦΗΜΙΣΗ	5
2.1.1 Ορισμός διαφήμισης	5
2.1.2 Διαφημιστικά μέσα.....	6
2.2 Η ΔΙΑΦΗΜΙΣΗ ΣΤΟ ΔΙΑΔΙΚΤΥΟ	6
2.3 ΙΣΤΟΡΙΚΗ ΕΞΕΛΙΞΗ ΗΛΕΚΤΡΟΝΙΚΗΣ ΔΙΑΦΗΜΙΣΗΣ	7
2.4 ΜΕΘΟΔΟΙ ΔΙΑΦΗΜΙΣΗΣ	8
2.4.1 Διαφημίσεις σε δικτυακούς τόπους.....	8
2.4.2 Καταχώρηση σε μηχανές αναζήτησης	12
2.4.3 Viral περιεχόμενο	14
2.4.4 Email – Ηλεκτρονικό Ταχυδρομείο	16
2.5 ΔΙΑΔΙΚΤΥΑΚΗ ΔΙΑΦΗΜΙΣΗ ΕΝΑΝΤΙ ΤΩΝ ΑΛΛΩΝ ΜΕΣΩΝ	16
2.6 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΚΑΙ ΜΕΙΟΝΕΚΤΗΜΑΤΑ	18
3 RECOMMENDER SYSTEMS	21
3.1 ΟΡΙΣΜΟΣ	21
3.2 ΔΙΑΔΙΚΑΣΙΑ ΣΥΣΤΑΣΗΣ	23
3.3 ΠΡΟΣΕΓΓΙΣΕΙΣ ΣΥΣΤΑΣΕΩΝ	24

3.3.1 Μη εξατομικευμένη προσέγγιση (non-personalized).....	24
3.3.2 Φιλτράρισμα βάσει περιεχομένου (content-based filtering).....	24
3.3.3 Συνεργατικό φιλτράρισμα (collaborative filtering)	26
3.3.3.1Συνεργατικό φιλτράρισμα με βάση τη μνήμη (memory-based collaborative filtering).....	27
3.3.3.2 Συνεργατικό φιλτράρισμα με βάση το μοντέλο (model-based collaborative filtering).....	30
3.3.4 Φιλτράρισμα βασισμένο στη γνώση (knowledge-based filtering)	30
3.3.5 Φιλτράρισμα βασισμένο σε δημογραφικά στοιχεία (demography-based filtering).....	31
3.3.6 Υβριδικά συστήματα συστάσεων (hybrid recommender systems)	31
4 ΣΥΣΤΗΜΑ ADSREC	33
4.1 ΔΙΚΤΥΑΚΟΣ ΤΟΠΟΣ.....	34
4.1.1 Αρχική σελίδα.....	34
4.1.2 Σελίδα κατηγορίας άρθρων	37
4.1.3 Σελίδα με τα πλήρη άρθρα	38
4.1.4 Βαθμολογία συστάσεων.....	39
4.1.5 Προφίλ χρήστη.....	40
4.2 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ	42
4.3 ΜΗΧΑΝΗ ΣΥΣΤΑΣΕΩΝ (RECOMMENDER ENGINE).....	45
4.3.1 Αλγόριθμος φιλτραρίσματος βάσει περιεχομένου.....	45
4.3.2 Αλγόριθμος συνεργατικού φιλτραρίσματος.....	47
5 ΜΕΛΕΤΗ ΧΡΗΣΤΩΝ	51
5.1 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ	51
5.2 ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ.....	53
5.2.1 Η διαδικασία.....	54
5.2.2 Σχετικότητα ως προς το περιεχόμενο	55
5.2.3 Σχετικότητα ως προς τα ενδιαφέροντα του χρήστη	57
5.3 ΠΑΡΟΥΣΙΑΣΗ ΣΤΟΙΧΕΙΩΝ	58
6 ΣΥΜΠΕΡΑΣΜΑΤΑ	63

6.1 ΓΕΝΙΚΑ ΣΥΜΠΕΡΑΣΜΑΤΑ.....	63
6.2 ΕΜΠΟΡΙΚΗ ΑΞΙΟΠΟΙΗΣΗ ΣΥΣΤΗΜΑΤΟΣ	65
ΒΙΒΛΙΟΓΡΑΦΙΑ	67
ΠΑΡΑΡΤΗΜΑ Ι.....	69
ΠΑΡΑΡΤΗΜΑ ΙΙ.....	70

Κατάλογος σχημάτων

Εικόνα 1. Το πρώτο banner στην ιστορία του διαδικτύου (Πηγή: [19]).....	8
Εικόνα 2. Αποτελέσματα σε αναζητήσεις στο Google με text links	10
Εικόνα 3. Banners και scrolling ads (Πηγή: http://intention.al/)	10
Εικόνα 4. Pop-up διαφήμιση (Πηγή: http://inkstory.eu/)	11
Εικόνα 5. Η splash screen διαφήμιση στο Inkscape (Πηγή: www.inkscape.org)	12
Εικόνα 6. Αποτελέσματα αναζήτησης και διαφημίσεις Adwords.....	14
Εικόνα 7. Εξάπλωση ενός viral περιεχομένου (Πηγή: www.eyegroup.gr)	15
Εικόνα 8. Διαγωνισμός "Like & Share"	15
Εικόνα 9. Ποσοστά των μέσων διαφήμισης (Πηγή: ZenithOptimedia).....	17
Εικόνα 10. Σελίδα προϊόντος στο eBay και οι συστάσεις.....	22
Εικόνα 11. Συστάσεις βάσει περιεχομένου - YouTube.....	26
Εικόνα 12. Συστάσεις με βάση το προϊόν - Amazon.....	28
Εικόνα 13. Συστάσεις με βάση το χρήστη - Amazon	29
Εικόνα 14. Η αρχιτεκτονική του συστήματος AdsRec	33
Εικόνα 15. Η αρχική σελίδα.....	35
Εικόνα 16. Διαφημιστικό banner διάστασης 300x250	35
Εικόνα 17. Διαφημιστικό banner διάστασης 336x280	36
Εικόνα 18. Διαφημιστικό banner διάστασης 468x60.....	36
Εικόνα 19. Διαφημιστικό banner διάστασης 728x90.....	36
Εικόνα 20. Φόρμα σύνδεσης εγγεγραμμένων χρηστών.....	37
Εικόνα 21. Φόρμα εγγραφής νέων χρηστών.....	37
Εικόνα 22. Η σελίδα κατηγορίας	38
Εικόνα 23. Η σελίδα του άρθρου.....	39
Εικόνα 24. Pop-up παράθυρο αξιολόγησης banner	40

Εικόνα 25. Το ιστορικό των άρθρων στο προφίλ του χρήστη.....	41
Εικόνα 26. Τα banners στο προφίλ του χρήστη	41
Εικόνα 27. Ο αλγόριθμος σε ψευδοκώδικα.....	46
Εικόνα 28. Ο αλγόριθμος συνεργατικού φιλτραρίσματος	48
Εικόνα 29. Οι βαθμολογίες που καταχωρεί ο χρήστης για το banner	52
Εικόνα 30. Ο πίνακας με τις βαθμολογίες των χρηστών	53

Κατάλογος πινάκων

Πίνακας 1. Περιγραφικά μέτρα των δύο δειγμάτων	56
Πίνακας 2. Αποτελέσματα t-test	56
Πίνακας 3. Περιγραφικά μέτρα	58
Πίνακας 4. Αποτελέσματα t-test	58
Πίνακας 5. Συχνότητες εμφάνισης βαθμολογιών σε σχέση με περιεχόμενο.....	59
Πίνακας 6. Αποτελέσματα βαθμολογιών σε σχέση με περιεχόμενο	59
Πίνακας 7. Συχνότητες εμφάνισης βαθμολογιών σε σχέση με ενδιαφέροντα....	61
Πίνακας 8. Αποτελέσματα βαθμολογιών σε σχέση με ενδιαφέροντα.....	61

Κατάλογος διαγραμμάτων

Διάγραμμα 1. Συχνοτήτων των δύο κατηγοριών αλγορίθμων.....	55
Διάγραμμα 2. Μέσης τιμής και τυπικής απόκλισης με βάση το περιεχόμενο ...	60
Διάγραμμα 3. Μέσης τιμής και τυπικής απόκλισης σε σχέση με ενδιαφέροντα	62

1 Εισαγωγή

Η δραματική εξάπλωση της χρήσης του διαδικτύου τα τελευταία χρόνια έχει επιφέρει ραγδαία αύξηση του όγκου πληροφορίας που διακινείται στον παγκόσμιο ιστό. Έτσι οι χρήστες του διαδικτύου έρχονται αντιμέτωποι καθημερινά με έναν τεράστιο όγκο δεδομένων, γνωστό και ως φαινόμενο υπερπληροφόρησης (information overload). Αυτή η τάση έχει ως συνέπεια να δυσκολεύονται οι χρήστες να ξεχωρίσουν αυτό που τους ενδιαφέρει πραγματικά και να χάνονται σε αυτό το «χάος» της πληροφορίας.

Για το λόγο αυτό, δημιουργήθηκαν τα Συστήματα Συστάσεων (Recommender Systems) και χρησιμοποιήθηκαν με μεγάλη επιτυχία στο ηλεκτρονικό εμπόριο καθώς και τη διαφήμιση στο διαδίκτυο. Σκοπός τους είναι να φιλτράρουν αυτόν τον τεράστιο όγκο δεδομένων και να βοηθήσουν τους χρήστες να εντοπίσουν προϊόντα και υπηρεσίες που τους ταιριάζουν. Μπορούν να στρέψουν την προσοχή των χρηστών σε συγκεκριμένα προϊόντα μέσα από ένα μεγάλο πλήθος διαθέσιμων προϊόντων και να επηρεάσουν στη λήψη αποφάσεων για αγορά.

1.1 Σκοπός εργασίας

Οι συστάσεις των συστημάτων αυτών δεν είναι καθόλου τυχαίες αλλά κρύβουν έναν ολόκληρο μηχανισμό από πίσω τους. Στηρίζονται σε τεχνικές ανάλυσης δεδομένων που προέρχονται από τη μέχρι τώρα συμπεριφορά του χρήστη και δηλώνουν τις προτιμήσεις του. Αυτή η ιδέα μας παρακίνησε να κατασκευάσουμε ένα σύστημα το οποίο εκμεταλλεύεται πληροφορίες από το προφίλ των χρηστών και τους προτείνει διαφημιστικά banners σύμφωνα με τις επιθυμίες τους.

Στα πλαίσια της εργασίας αναπτύχθηκε ένα σύστημα συστάσεων διαφημιστικών banners, το οποίο έχει αναλάβει την ενημέρωση των χρηστών για διάφορα θέματα. Σκοπός του είναι να προτείνει κατάλληλα διαφημιστικά banners στους χρήστες του καθώς αυτοί διαβάζουν τα άρθρα τα οποία τους ενδιαφέρουν. Ο ένας αλγόριθμος εκμεταλλεύεται πληροφορίες από το προφίλ των χρηστών και τους προτείνει διαφημιστικά banners σύμφωνα με τις επιθυμίες τους και ο άλλος σύμφωνα με το περιεχόμενο των άρθρων που διαβάζουν.

Το σύστημα που διαλέξαμε να κατασκευάσουμε αναμένεται να προκαλέσει ενδιαφέρον διότι η διαφήμιση στο διαδίκτυο είναι ένα άκρως επίκαιρο ζήτημα. Δεν υπάρχει επιχείρηση που θέλει να διαφημιστεί και δε στρέφεται προς τη λύση του διαδικτύου. Ιδιαίτερο ενδιαφέρον επίσης παρουσιάζει το πώς θα μπορέσει ένα σύστημα να παρουσιάσει τα κατάλληλα banners στους κατάλληλους χρήστες ώστε να μείνουν ικανοποιημένοι και οι δύο.

Αυτό που επιζητούμε για το σύστημα μας είναι να διαδραματίσει διπλό ρόλο κατά τη χρήση του. Αφενός να τραβήξει το ερευνητικό ενδιαφέρον για τον τρόπο με τον οποίο διαμορφώνεται το προφίλ του χρήστη και δημιουργούνται εξατομικευμένες συστάσεις και αφετέρου να αποτελέσει έναν έμπιστο δικτυακό τόπο στον οποίο μπορούν οι χρήστες να ενημερώνονται για οτιδήποτε τους ενδιαφέρει.

1.2 Δομή της εργασίας

Η εργασία απαρτίζεται από πέντε κεφάλαια. Στο πρώτο κεφάλαιο δίνουμε έναν ορισμό της διαφήμισης γενικά και αναφέρουμε τα μέσα τα οποία χρησιμοποιούνται για να διαφημισθούν προϊόντα και υπηρεσίες. Στη συνέχεια γινόμαστε πιο συγκεκριμένοι με το θέμα της εργασίας και ορίζουμε τη διαφήμιση στο Διαδίκτυο. Αφού αναφέρουμε κάποιες γενικές πληροφορίες για να κατανοήσουμε καλύτερα τη λειτουργία της, κάνουμε μία μικρή ιστορική αναδρομή ξεκινώντας από το πρώτο διαφημιστικό banner – που και πότε εμφανίστηκε. Αναφέρουμε όλες τις μορφές της διαδικτυακής διαφήμισης, από τις οποίες εμείς έχουμε επιλέξει την καταχώρηση διαφήμισης σε δικτυακό τόπο μέσω banners. Εξηγούμε γιατί επιλέγεται η διαφήμιση στο διαδίκτυο έναντι των υπολοίπων μέσων και αναλύουμε τα πλεονεκτήματα και μειονεκτήματα της.

Στο δεύτερο κεφάλαιο παρουσιάζουμε τα Συστήματα Συστάσεων (Recommender Systems). Αρχικά δίνουμε τον ορισμό των συστημάτων αυτών και στη συνέχεια αναπτύσσουμε τη διαδικασία με την οποία γίνονται οι συστάσεις. Προκειμένου να κατανοηθούν πλήρως τα συστήματα και ο τρόπος με τον οποίο παράγονται οι συστάσεις τους, αναλύουμε τις προσεγγίσεις τους. Δύο από τις προσεγγίσεις χρησιμοποιούνται και στο δικό μας σύστημα: το φιλτράρισμα βάσει περιεχομένου (content-based filtering) και το συνεργατικό φιλτράρισμα (collaborative filtering).

Το τρίτο κεφάλαιο είναι αφιερωμένο στο σύστημα συστάσεων το οποίο σχεδιάσαμε, το AdsRec. Ορίζουμε την αρχιτεκτονική του συστήματος και περιγράφουμε όλα τα

απαραίτητα υποσυστήματα που βοηθούν στην πραγματοποίηση των συστάσεων στους χρήστες του συστήματος. Ο δικτυακός τόπος (website) είναι η διεπαφή με την οποία οι χρήστες αλληλεπιδρούν με το σύστημα AdsRec, το οποίο περιλαμβάνει άρθρα και τις συστάσεις των διαφημιστικών banners. Η μηχανή συστάσεων (Recommender Engine) και οι αλγόριθμοι σύστασης περιγράφονται λεπτομερώς ώστε να κατανοήσουμε πλήρως τον τρόπο λειτουργίας του συστήματος συστάσεων AdsRec. Η βάση δεδομένων (Database) αναλύεται με όλες τις οντότητες και συσχετίσεις της, δείχνοντας σχηματικά τη δομή της ώστε να κατανοηθεί ευκολότερα.

Στο τέταρτο κεφάλαιο εξετάζουμε τη μελέτη περιπτώσεων χρηστών που πραγματοποιήσαμε για το σύστημα AdsRec. Στη μελέτη αυτή παίρνουν μέρος είκοσι εγγεγραμμένοι χρήστες, οι οποίοι έχουν επιλεγεί ως τυχαίο δείγμα. Οι χρήστες αυτοί χρησιμοποιούν το σύστημα, αλληλεπιδρούν μαζί του, αξιολογούν τις συστάσεις που τους γίνονται και εμείς βγάζουμε τα συμπεράσματα μας.

Τέλος, στο πέμπτο κεφάλαιο παρουσιάζουμε τα τελικά συμπεράσματα από τη χρήση του συστήματος και τι έδειξε η μελέτη χρηστών. Ποιο banner πατήθηκε περισσότερο και γιατί; Τέλος, προτείνουμε ποια θα μπορούσε να είναι η εμπορική αξιοποίηση του συστήματος που κατασκευάσαμε στα πλαίσια του μεταπτυχιακού διπλώματος, καθώς πιστεύουμε ότι είναι ένα σύστημα το οποίο μπορεί να αξιοποιηθεί για εμπορικούς σκοπούς και να εξελιχθεί.

2 Διαφήμιση στο Διαδίκτυο

Αναμφισβήτητα τα τελευταία χρόνια η διαφήμιση έχει περάσει από διάφορες μορφές και μέσα. Η τεχνολογία αλλάζει με ραγδαίους ρυθμούς και από τις εξελίξεις αυτές δε θα μπορούσε να μην επηρεαστεί και η διαφήμιση. Το καταναλωτικό κοινό και ο τρόπος με τον οποίο πραγματοποιούνται οι συναλλαγές έχουν επηρεαστεί και αυτά. Εδώ και χρόνια έχουμε περάσει στην εποχή του ηλεκτρονικού εμπορίου, όπου μία ηλεκτρονική συναλλαγή μπορεί να αντικαταστήσει πλήρως μία συναλλαγή στον πραγματικό κόσμο. Έτσι, οι επιχειρήσεις που θέλουν να διαφημίσουν προϊόντα και υπηρεσίες οφείλουν να ακολουθήσουν την τάση αυτή και να απευθυνθούν στους χρήστες του διαδικτύου.

2.1 Διαφήμιση

Διαφήμιση [1] είναι η διαδικασία γνωστοποίησης και επηρεασμού του καταναλωτικού κοινού για ένα προϊόν ή μία υπηρεσία επί πληρωμή. Στην πιο απλή σημασία της λέξης «διαφήμιση» σημαίνει να τραβάς την προσοχή σε κάτι, να γνωστοποιείς κάτι σε κάποιον ή να τον πληροφορείς γι' αυτό.

2.1.1 Ορισμός διαφήμισης

Ως διαφήμιση [1] νοείται η μονομερής (δηλαδή από τον πομπό, χωρίς τη συμμετοχή του δέκτη) επικοινωνία και μετάδοση πληροφοριών με πειστικό τρόπο για προϊόντα και υπηρεσίες ή ιδέες όπου ο διαφημιζόμενος (αυτός που έχει πληρώσει για τη μετάδοση της πληροφορίας) είναι εμφανής (δηλαδή δεν είναι κρυφή ή έμμεση ή μαύρη διαφήμιση και είναι φανερό ποιος πληρώνει για την προβολή).

Άλλος ορισμός της διαφήμισης [1] είναι "η επί πληρωμή μονομερής επικοινωνία ενός εμφανούς χορηγού που χρησιμοποιεί μέσα επικοινωνίας για να πείσει ή να επηρεάσει ένα κοινό".

2.1.2 Διαφημιστικά μέσα

Η διαφήμιση μπορεί να χρησιμοποιήσει διάφορα μέσα [1, 2] προκειμένου να πετύχει την αποστολή της, τη γνωστοποίηση δηλαδή του προϊόντος ή της υπηρεσίας. Κριτήρια επιλογής αποτελεί η αγορά που έχει σαν στόχο, η αποτελεσματικότητα του μέσου και το κόστος που είναι διατεθειμένη να πληρώσει η επιχείρηση. Ορισμένα από τα μέσα αυτά είναι:

- Η τηλεόραση
- Το ραδιόφωνο
- Ο κινηματογράφος
- Διάσημα πρόσωπα
- Τύπος (εφημερίδες, περιοδικά κλπ.)
- Αφίσες
- Διαδίκτυο
- Φυλλάδια κλπ.

Το διαδίκτυο προσφέρει ορισμένες διευκολύνσεις σε σχέση με τα υπόλοιπα μέσα. Με τη χρήση του διαδικτύου επιτυγχάνεται άμεση επικοινωνία με τον πελάτη αλλά και με την επιχείρηση (αμφίδρομη επικοινωνία), μείωση του κόστους με αρκετούς τρόπους δωρεάν διαφήμισης (μέσα κοινωνικής δικτύωσης, «word of mouth» κλπ.) και η αποκλειστική δυνατότητα να παρέχει εξατομικευμένες διαφημίσεις ανάλογα με τα ενδιαφέροντα κάθε χρήστη.

2.2 Η διαφήμιση στο διαδίκτυο

Αρκετές διαδικασίες της καθημερινής μας ζωής έχουν αλλάξει σε σημαντικό βαθμό από την εμφάνιση του διαδικτύου. Ένα από αυτά είναι ο τρόπος επικοινωνίας, ο οποίος έχει εξελιχθεί με τη χρήση του διαδικτύου καθώς μπορούμε να στείλουμε ένα ηλεκτρονικό μήνυμα οπουδήποτε και οποιαδήποτε στιγμή το επιθυμούμε. Ακόμη, έχει αλλάξει ο τρόπος με τον οποίο ψωνίζουμε, καθώς όλο και περισσότερο αυξάνεται το ποσοστό των χρηστών που χρησιμοποιούν τα ηλεκτρονικά καταστήματα (e-shops) τα οποία επιτρέπουν στους καταναλωτές να αγοράζουν προϊόντα, και όχι μόνο, μέσω διαδικτύου.

Έτσι γίνεται κατανοητό ότι δε θα μπορούσε να μην επηρεαστεί και ο τρόπος με τον οποίο γίνεται η διαφήμιση των προϊόντων και υπηρεσιών. Η διαφήμιση στο διαδίκτυο κερδίζει συνεχώς έδαφος τα τελευταία χρόνια λόγω των σημαντικών πλεονεκτημάτων

που παρουσιάζει έναντι των υπολοίπων μορφών διαφήμισης. Βασικά πλεονεκτήματα με τη χρήση του διαδικτύου στη διαφήμιση είναι ότι γίνεται πιο στοχευμένη, μειώνεται το κόστος και η αποτελεσματικότητα της είναι μετρήσιμη.

Βέβαια παρά τα σημαντικά πλεονεκτήματα που έχει η διαφήμιση στο διαδίκτυο, δεν πρέπει να θεωρηθεί ότι είναι εύκολη υπόθεση. Χρειάζεται αρκετή μελέτη και προσπάθεια για να στηθεί σωστά μία διαφημιστική καμπάνια στο διαδίκτυο και να καταστεί επιτυχημένη. Υπάρχει πολύ μεγάλος ανταγωνισμός που σημαίνει ότι το κόστος μπορεί να αυξηθεί σημαντικά. Το μεγαλύτερο όμως πρόβλημα της διαφήμισης στο διαδίκτυο έγκειται στο γεγονός ότι οι περισσότεροι χρήστες έχουν αναπτύξει το φαινόμενο της διαφημιστικής τυφλότητας (ad blindness) [3].

Πολλοί είναι οι χρήστες που αντιμετωπίζουν τη διαφήμιση ως αναγκαίο κακό και ορισμένες φορές τη βρίσκουν ενοχλητική. Θεωρούν τη διαφήμιση ως έναν τρόπο επιβολής του διαφημιστικού προϊόντος και ένας τρόπος εξαναγκασμού από αυτόν που έχει πληρώσει για να εμφανιστεί η διαφήμιση αυτή. Σπάνια κάποιος χρήστης σκέφτεται ότι η διαφήμιση εμφανίζεται σε αυτόν γιατί μπορεί να έχει κάποια αξία ως περιεχόμενο για τον ίδιο και μπορεί όντως να τον ενδιαφέρει. Παρόμοια είναι και η αντιμετώπιση των τηλεθεατών στην τηλεόραση, οι οποίοι κάνουν zapping αλλάζοντας το κανάλι ώστε να μη δουν τις διαφημίσεις που προβάλλονται.

Τα τελευταία χρόνια ενισχύεται η τάση της διαφήμισης σε συγκεκριμένους δικτυακούς τόπους, όπως είναι τα αποτελέσματα αναζήτησης του Google και το Facebook. Ο ανταγωνισμός σε αυτές τις σελίδες είναι ιδιαίτερα μεγάλος οπότε χρειάζεται ιδιαίτερη προσοχή και μελέτη για να ξεχωρίσει μία διαφημιστική καμπάνια. Σημαντικό ρόλο παίζει η σχεδίαση και το διαφημιστικό μήνυμα που περιλαμβάνεται προκειμένου να τραβήξει το ενδιαφέρον του χρήστη, ώστε να κάνει κλικ στη συγκεκριμένη διαφήμιση και όχι στη διπλανή.

2.3 Ιστορική εξέλιξη ηλεκτρονικής διαφήμισης

Είναι γεγονός ότι το διαδίκτυο προκάλεσε επανάσταση στη διαφήμιση. Το ντεμπούτο της σημειώθηκε τον Οκτώβριο του 1994, όταν ο Joe McCambley, ιδιοκτήτης μιας μικρής διαφημιστικής, έφτιαξε την πρώτη ψηφιακή διαφήμιση. Στην Εικόνα 1 βλέπουμε το πρώτο banner, το οποίο εμφανίστηκε στο πρώτο ηλεκτρονικό περιοδικό του κόσμου, το Hotwired.com. Κάνοντας κλικ κάποιος χρήστης μεταφερόταν σε έναν άλλον δικτυακό τόπο όπου μπορούσε να δει επτά από τα μεγαλύτερα μουσεία του

κόσμου. Το κοινό πραγματικά το λάτρεψε και το 44% των επισκεπτών του Hotwired.com έκαναν κλικ πάνω στη διαφήμιση.



Εικόνα 1. Το πρώτο banner στην ιστορία του διαδικτύου (Πηγή: [19])

Από αυτή τη στιγμή και μετά η διαφήμιση αλλάζει ριζικά και καθιερώνεται η διαδικτυακή διαφήμιση. Για πρώτη φορά οι διαφημιστές έχουν στην διάθεσή τους απολύτως μετρήσιμα δεδομένα, όχι μόνο για το πόσα άτομα είδαν τη διαφήμιση αλλά και πόσα άτομα έχουν αλληλεπιδράσει με αυτήν. Αν και η συνέχεια για τη διαδικτυακή διαφήμιση φαντάζει εύκολη, διαπιστώνουμε ότι δεν ήταν και τόσο τελικά. Το διαδίκτυο γέμισε γρήγορα με μεγάλο πλήθος υπηρεσιών και ιδεών, με συνέπεια να αυξάνονται όλο και περισσότερο και οι προσδοκίες. Αρκετές ήταν οι εταιρείες που πίστεψαν ότι θα γίνουν κερδοφόρες αποκλειστικά μέσω του διαδικτύου.

Όμως το 2001 με τη φούσκα των dot.com προκλήθηκε μεγάλη ζημιά σε πολλούς επενδυτές και πολλές ήταν οι επιχειρήσεις που χρεωκόπησαν. Ύστερα από την κατάρρευση, δοκίμασαν άλλες μεθόδους διαδικτυακής διαφήμισης και από τότε παρουσιάστηκε σημαντική βελτίωση της διαφήμισης στο διαδίκτυο. Διαπιστώθηκε ότι τα πρώτα δέκα χρόνια το εισόδημα που επέφερε η διαδικτυακή διαφήμιση αυξήθηκε ραγδαία και από λίγες εκατοντάδες χιλιάδες δολάρια το 1994 έφτασε σχεδόν τα 10.000.000.000 δολάρια το 2004 [5].

2.4 Μέθοδοι διαφήμισης

Υπάρχουν διάφοροι τρόποι και τεχνικές [6] που μπορούν να χρησιμοποιηθούν για να διαφημιστούν προϊόντα και υπηρεσίες στο διαδίκτυο. Οι τρόποι αυτοί εξελίσσονται με την πάροδο του χρόνου και παρουσιάζονται στην ενότητα αυτή.

2.4.1 Διαφημίσεις σε δικτυακούς τόπους

Πρόκειται για διαφημίσεις οι οποίες τοποθετούνται σε κάποιο σημείο ενός δικτυακού τόπου παρουσιάζοντας κάποιο μήνυμα ή κάποιο προϊόν/υπηρεσία. Ακολουθούν τα δομικά στοιχεία του είδους των διαφημίσεων αυτών.

Banners

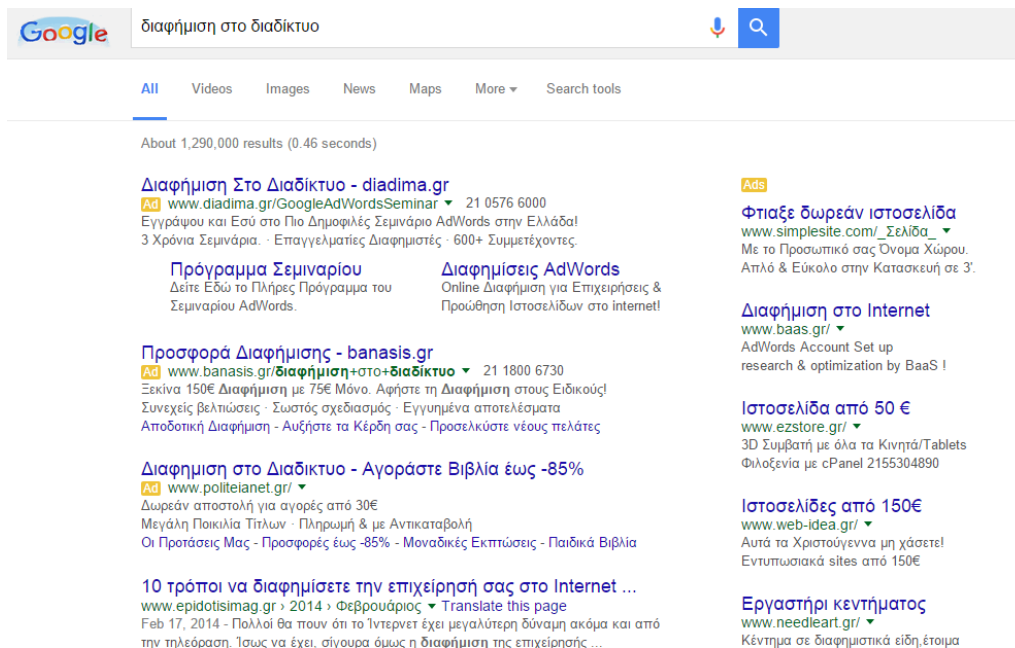
Τα banners είναι γραφικά εικονίδια, τα οποία συνδέονται με το site του διαφημιζόμενου. Κάθε φορά που ο χρήστης πατάει πάνω στο εικονίδιο μεταφέρεται σε μία νέα ιστοσελίδα, όπου μπορεί να πληροφορηθεί για το προϊόν ή την υπηρεσία που τον ενδιαφέρει. Υπάρχουν διάφορα είδη banners ανάλογα με την προσοχή που θέλει να τραβήξει ο διαφημιζόμενος.

Τα banners αποτελούν τον πιο διαδεδομένο τρόπο διαφήμισης, όμως η απόδοσή τους δεν είναι ξεκάθαρη. Μπορεί σίγουρα να δυναμώσει το brand name μιας εταιρίας, όμως για να κάνει κλικ ο χρήστης πρέπει να είναι αρκετά έξυπνο και εντυπωσιακό ή να τον ενδιαφέρει πραγματικά. Λόγω της εμφάνισης τεράστιου πλήθους banners, πολλοί είναι οι χρήστες που τα απορρίπτουν πριν καν τα δουν. Έτσι η πλειοψηφία των χρηστών δύσκολα τα παρατηρεί και ακόμα πιο δύσκολα κάνει κλικ πάνω τους.

Υπάρχουν διάφορα χαρακτηριστικά που επηρεάζουν την απόδοση ενός banner. Ένα από αυτά είναι σίγουρα το μέρος στο οποίο εμφανίζεται, καθώς όσο πιο ψηλά είναι τοποθετημένο τόσο περισσότερες πιθανότητες έχει να το προσέξουν. Πολύ σημαντικό στοιχείο είναι η σχεδίαση του, η οποία πρέπει να είναι έξυπνη και εντυπωσιακή ώστε να τραβήξει την προσοχή του χρήστη. Το περιεχόμενο του banner πρέπει να είναι σχετικό με το περιεχόμενο της σελίδας που εμφανίζεται, αλλά και με τα ενδιαφέροντα του χρήστη, για να έχει περισσότερες πιθανότητες να κάνει κλικ πάνω του.

Text link – Σύνδεσμος κειμένου

Η λειτουργία τους είναι παρόμοια με αυτή των banners, με τη διαφορά ότι αντί για εικόνα ή γραφικό έχουμε κείμενο, το οποίο είναι σύνδεσμος (text link) στην ιστοσελίδα της διαφημιζόμενης εταιρίας. Είναι από τους πιο διαδεδομένους και οικονομικούς τρόπους διαφήμισης σε καταλόγους, θεματικές ιστοσελίδες και σε προγράμματα διαφήμισης των μηχανών αναζήτησης. Για παράδειγμα, η μηχανή αναζήτησης Google με τα αποτελέσματα της και το πρόγραμμα Adwords, όπως βλέπουμε στην Εικόνα 2, έφερε επανάσταση σε αυτόν τον τρόπο διαφήμισης και έδειξε το δρόμο σε πολλούς διαφημιζόμενους και διαφημιστές.



Εικόνα 2. Αποτελέσματα σε αναζητήσεις στο Google με text links

Scrolling ads

Τα banners και τα text links που μελετήσαμε πριν είναι τοποθετημένα σε ένα συγκεκριμένο σημείο κάποιας ιστοσελίδας. Τα scrolling ads είναι παρόμοια με τα banners, με τη διαφορά ότι η θέση τους δεν είναι σταθερή αλλά ακολουθεί τον χρήστη σε οποιοδήποτε σημείο της ιστοσελίδας βρίσκεται ώστε να μη χαθεί ποτέ από το οπτικό του πεδίο. Η διαφορά αυτή ακριβώς αποτυπώνεται στην Εικόνα 3.

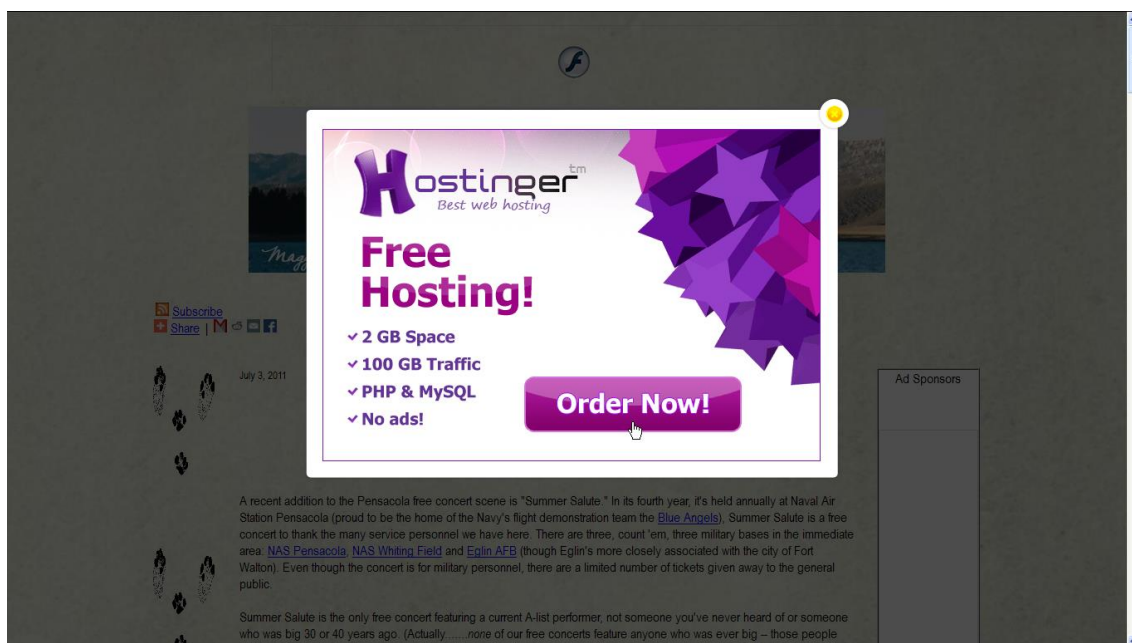


Εικόνα 3. Banners και scrolling ads (Πηγή: <http://intention.al/>)

Pop-up & pop-under windows

Αυτού του είδους οι διαφημίσεις περιλαμβάνουν ξεχωριστά παράθυρα πολύ μικρότερου μεγέθους από αυτό της κανονικής σελίδας και εμφανίζονται αυτόματα σε κάποιον χρήστη χωρίς να το έχει επιλέξει ο ίδιος. Τα pop-up παράθυρα ανοίγουν μπροστά από το ενεργό παράθυρο του χρήστη όπως βλέπουμε στην Εικόνα 4, ενώ τα pop-under ανοίγουν πίσω από το παράθυρο αυτό.

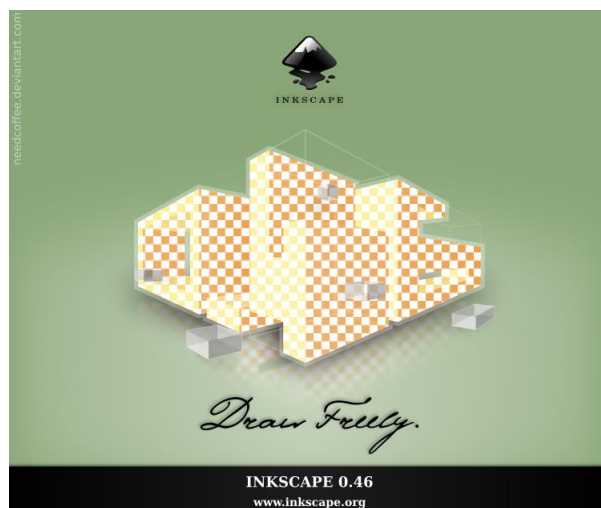
Τα παράθυρα αυτά μπορούν να θεωρηθούν ενοχλητικά καθώς ανοίγουν παρά τη θέληση των χρηστών με αποτέλεσμα οι περισσότεροι χρήστες να τα κλείνουν πριν καν προλάβουν να δουν το περιεχόμενό τους. Το πλεονέκτημα που έχουν είναι ότι δεν αναγκάζουν τους χρήστες να αφήσουν την ιστοσελίδα που έχουν επισκεφτεί, κάτι το οποίο τα καθιστά λίγο πιο φιλικά προς το χρήστη.



Εικόνα 4. Pop-up διαφήμιση (Πηγή: <http://inkstory.eu/>)

Splash screen - Εμβόλιμη οθόνη

Η εμβόλιμη οθόνη είναι μία εισαγωγική ιστοσελίδα, όπως φαίνεται στην Εικόνα 5, η οποία εμφανίζεται πριν την αρχική σελίδα με σκοπό να τραβήξει την προσοχή του χρήστη για ένα μικρό χρονικό διάστημα. Αυτός ο τρόπος διαφήμισης είναι αρκετά προκλητικός και μπορεί να θεωρηθεί σε μεγάλο βαθμό ενοχλητικός. Συνίσταται μόνο σε περιπτώσεις όπου η διαφήμιση έχει άμεση σχέση με την ιστοσελίδα που τη φιλοξενεί και σε περιπτώσεις όπου κρίνεται πολύ ενδιαφέρουσα για τον ίδιο τον χρήστη.



Εικόνα 5. Η splash screen διαφήμιση στο Inkscape (Πηγή: www.inkscape.org)

2.4.2 Καταχώρηση σε μηχανές αναζήτησης

Από τη στιγμή που δημιουργείται μία νέα ιστοσελίδα πρέπει να δηλωθεί το περιεχόμενο της και ο τρόπος που μπορούν να την επισκεφτούν οι χρήστες του διαδικτύου. Έτσι η διαφήμιση μέσω των μηχανών αναζήτησης κρίνεται από τις πιο αποτελεσματικές και κερδοφόρες μεθόδους, καθώς με ένα αρκετά χαμηλό κόστος μπορεί να επιφέρει θεαματικά αποτελέσματα.

Από τη στιγμή που δημιουργείται ένας καινούργιος δικτυακός τόπος, οι μηχανές αναζήτησης τον εντοπίζουν σε διάστημα λίγων ωρών. Υπάρχει η δυνατότητα να καταχωρήσει κάποιος μόνος του τον δικτυακό τόπο που επιθυμεί στις μηχανές αναζήτησης με μία πολύ απλή διαδικασία κατά την οποία δηλώνει κάποιες λέξεις-κλειδιά (keywords). Έτσι όταν κάποιος χρήστης αναζητήσει κάποια από τις λέξεις-κλειδιά, η μηχανή αναζήτησης θα εμφανίσει τη διεύθυνση του δικτυακού τόπου.

Ο τρόπος καταχώρησης που αναφέρθηκε είναι δωρεάν και αφορά στα οργανικά αποτελέσματα. Υπάρχει ακόμη ένας τρόπος καταχώρησης κατά τον οποίο ενοικιάζεται κάποιο τμήμα των μηχανών αναζήτησης για διαφημίσεις. Αυτή η μέθοδος είναι επί πληρωμή και με αυτό τον τρόπο μπορεί μία επιχείρηση να ξεχωρίσει από τον ανταγωνισμό της. Για παράδειγμα, στη μηχανή αναζήτησης Google έχουμε τα οργανικά αποτελέσματα και τις διαφημίσεις Adwords.

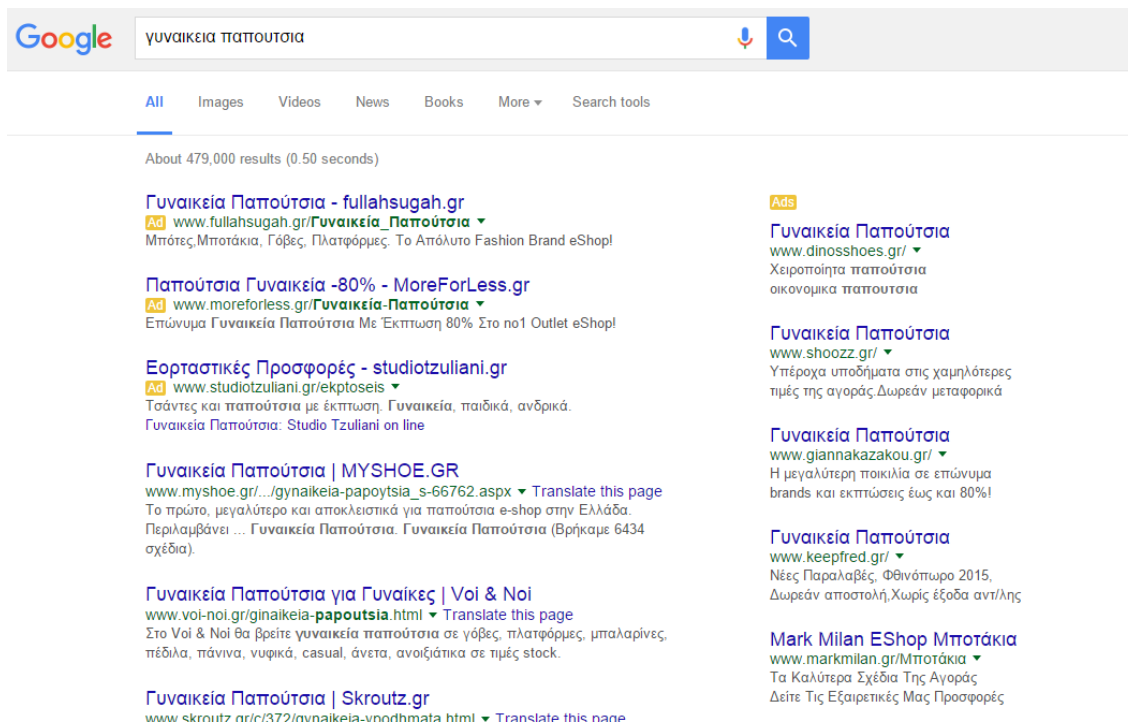
Μία ιστοσελίδα οφείλει να είναι φιλική προς τις μηχανές αναζήτησης ώστε να δημιουργούνται οι ιδανικότερες συνθήκες για την όσο το δυνατό πιο ψηλά κατάταξη

της στα αποτελέσματα αναζήτησης. Αυτό πραγματοποιείται με μία διαδικασία βελτιστοποίησης, που ονομάζεται SEO (Search Engine Optimization). Πρόκειται για ένα αρκετά χρήσιμο εργαλείο για αποτελεσματική διαφήμιση χωρίς πολλά έξοδα.

Για τα πληρωμένα αποτελέσματα με ένα καλό SEO μία διαφήμιση μπορεί να εμφανιστεί ψηλότερα από τις υπόλοιπες ανταγωνιστικές διαφημίσεις, το κόστος ενός κλικ να είναι μικρότερο και η διαφήμιση να εμφανίζεται πιο τακτικά. Διαπιστώνουμε λοιπόν το πόσο σημαντική είναι αυτή η διαδικασία για μία σελίδα που θέλει να διαφημιστεί. Ακόμη, σε συνδυασμό με επιπλέον στόχευση (στόχευση περιεχομένου με κάποια keywords, στόχευση κοινού κ.α.) τα αποτελέσματα μπορεί να γίνουν θαυματουργά.

Τα αποτελέσματα αναζητήσεων στις δημοφιλείς μηχανές αναζήτησης είναι ένα πολύ σημαντικό μέσο προώθησης. Αυτό συμβαίνει διότι όλο και μεγαλύτερο ποσοστό των χρηστών αναζητούν στο διαδίκτυο οτιδήποτε τους ενδιαφέρει. Πλέον στις μέρες μας οι καταναλωτές πριν αγοράσουν κάποιο προϊόν κάνουν μία μικρή έρευνα στο διαδίκτυο αναζητώντας πληροφορίες για το προϊόν που ενδιαφέρονται, εμπειρίες και γνώμες άλλων χρηστών.

Στην Εικόνα 6 βλέπουμε τα αποτελέσματα αναζήτησης του όρου «γυναικεία παπούτσια» στη μηχανή αναζήτησης Google. Πάνω από τα κανονικά αποτελέσματα και στη στήλη δεξιά έχουμε τις πληρωμένες διαφημίσεις Adwords. Μελετώντας τα αποτελέσματα αυτά, διαπιστώνουμε τη χρησιμότητα του SEO τόσο στα οργανικά όσο και στα πληρωμένα αποτελέσματα. Όσον αφορά τα οργανικά αποτελέσματα, που είναι τα κανονικά αποτελέσματα, η Google αντιμετωπίζει με διαφορετικό τρόπο τις ιστοσελίδες που είναι καλογραμμένες, σωστά δομημένες και με σωστό περιεχόμενο.



Εικόνα 6. Αποτελέσματα αναζήτησης και διαφημίσεις Adwords

2.4.3 Viral περιεχόμενο

Η διαφήμιση μέσω ενός viral περιεχομένου με διαδραστικό τρόπο είναι η τελευταία τάση της εποχής. Πρόκειται για μία νέα μορφή διαφήμισης και η φιλοσοφία της στηρίζεται στη δημιουργία σταθερών σχέσεων της διαφημιζόμενης επιχείρησης με τους πελάτες της. Το περιεχόμενο αυτό, το οποίο βρίσκεται στην ιστοσελίδα του διαφημιζόμενου, στη σελίδα του στα social media ή στο YouTube, θα πρέπει να εξυπηρετεί την πολιτική του αλλά ταυτόχρονα να είναι αρεστό και χρήσιμο στο κοινό. Σκοπός είναι να προκαλέσει τους χρήστες να μοιραστούν το περιεχόμενο αυτό σε κοινωνικά δίκτυα και blogs, να γίνει δηλαδή viral.

Ο αγγλικός όρος viral που έχει επικρατήσει για το περιεχόμενο το οποίο διαδίδεται με ταχύτατους ρυθμούς μέσω διαδικτυακής διαμοίρασης (sharing), σημαίνει «υικός» και δηλώνει αυτόν που έχει σχέση με το ιό. Το όνομα αυτό δηλώνει την ταχύτητα με την οποία εξαπλώνεται στο διαδίκτυο ένα viral περιεχόμενο. Στην Εικόνα 7 φαίνεται αυτή η ραγδαία εξάπλωση του περιεχομένου που γίνεται viral.



Εικόνα 7. Εξάπλωση ενός viral περιεχομένου (Πηγή: www.eyegroup.gr)

Η λογική του viral περιεχομένου αγγίζει σε αρκετά μεγάλο βαθμό αυτή του “word of mouth” και μπορούμε να πούμε ότι είναι η ηλεκτρονική εκδοχή του. Σύμφωνα με αυτή τη λογική οι άνθρωποι στηρίζονται κατά πολύ στη γνώμη των φίλων τους και σε άτομα από τον κοντινό περίγυρο τους. Με τη χρήση του διαδικτύου η εξάπλωση της γνώμης αυτής μπορεί να γίνει με ταχύτατους ρυθμούς. Πολλές είναι οι επιχειρήσεις που επιδιώκουν με διάφορους τρόπους να γίνει viral κάποιο περιεχόμενο τους. Οι στόχοι τους δεν είναι μόνο οικονομικοί, αλλά επιδιώκουν αναγνωρισιμότητα, δημοτικότητα, προσέλκυση νέων πελατών και ανάπτυξη των σχέσεων με τους ήδη υπάρχοντες πελάτες.

Για παράδειγμα στην Εικόνα 8 βλέπουμε έναν διαγωνισμό από ένα εστιατόριο, το οποίο παροτρύνει να κάνουν “like & share” το περιεχόμενο για να κερδίσουν κάποιες προσφορές. Με τον τρόπο αυτό διαφημίζει το προϊόν του και παρακινεί τους πελάτες του να κάνουν viral το περιεχόμενο. Συνήθως σε αυτού του είδους τους διαγωνισμούς κερδίζει αυτός που συγκεντρώνει τα περισσότερα likes, που σημαίνει ότι το είδαν οι περισσότεροι χρήστες και αυτός είναι ουσιαστικά και ο πρωταρχικός στόχος.



Εικόνα 8. Διαγωνισμός “Like & Share”

2.4.4 Email – Ηλεκτρονικό Ταχυδρομείο

Ένας πολύ διαδεδομένος τρόπος διαφήμισης είναι αυτός μέσω email. Αποτελεί βασικό εργαλείο για τους διαφημιστές καθώς απευθύνονται σε συγκεκριμένο κοινό το οποίο ενδιαφέρεται για τις διαφημίσεις αυτές και έτσι αυξάνεται η αποτελεσματικότητά τους. Όταν αυτά αποστέλλονται μαζί σε ένα συγκεκριμένο κοινό με σκοπό να προωθήσουν κάποιο προϊόν ή υπηρεσία, ονομάζεται newsletter και αποτελεί ένα πολύ βασικό εργαλείο του internet marketing.

Μια καμπάνια newsletter δεν θεωρείται εύκολη υπόθεση καθώς χρειάζεται ιδιαίτερη προσοχή στο στήσιμο της. Ένα σωστό newsletter πρέπει να είναι απόλυτα προσεγμένο ως προς το περιεχόμενο και τη σχεδίαση, να εμφανίζεται σωστά σε όλα τα προγράμματα ανάγνωσης email και να ελαχιστοποιεί όσο το δυνατό περισσότερο την πιθανότητα να θεωρηθεί ανεπιθύμητη αλληλογραφία (spam). Ακόμη, πρέπει να δίνει τη δυνατότητα σε κάποιον που δεν επιθυμεί να λαμβάνει τα ενημερωτικά email να διαγραφεί από τη λίστα.

Υπάρχουν δύο χρήσεις όσον αφορά το σκοπό της αποστολής ενός newsletter. Ο ένας έχει να κάνει με την προσέλκυση νέων πελατών (newsletter campaign) και πρέπει να συλλεχθούν όσο το δυνατό περισσότερα emails πιθανών υποψηφίων. Ο άλλος απευθύνεται στο ήδη υπάρχον πελατολόγιο με σκοπό τη δημιουργία στενής επαφής και σχέσης με τους πελάτες μέσω της ενημέρωσης τους για προσφορές ή νέα (ενημερωτικό newsletter). Η αποστολή ενός newsletter πρέπει να είναι συχνή αλλά να μην ξεπερνάει τη μία φορά την εβδομάδα, και το περιεχόμενο του πρέπει να είναι χρήσιμο και ουσιώδες.

2.5 Διαδικτυακή διαφήμιση έναντι των άλλων μέσων

Το διαδίκτυο είναι ένα μέσο που έχει επηρεάσει τις ζωές όλων των ανθρώπων καθώς τις έχει διευκολύνει κατά πολύ. Οι χρήστες του διαδικτύου έρχονται σε επαφή με έναν τεράστιο όγκο δεδομένων από όπου μπορούν να ενημερωθούν για τα πάντα και να ασχοληθούν με οτιδήποτε τους ενδιαφέρει. Ακόμη, μπορούν να έρθουν σε επαφή με ανθρώπους που βρίσκονται οπουδήποτε πάνω στη γη.

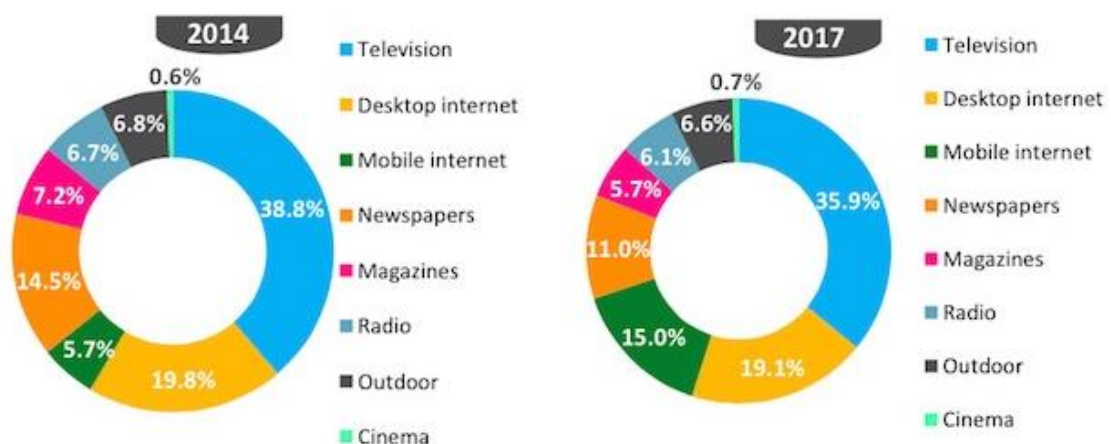
Παράλληλα το ηλεκτρονικό εμπόριο αλλάζει τα δεδομένα και τον τρόπο με τον οποίο γίνονται οι αγορές καθώς και όλες τις ενέργειες που σχετίζονται με τις αγορές. Κάθε χρήστης όπου και να βρίσκεται έχει τη δυνατότητα να προβεί σε αγορά του προϊόντος ή της υπηρεσίας που επιθυμεί. Ακόμη μπορεί να κάνει πολύ εύκολα έρευνα αγοράς για το

αντικείμενο που ενδιαφέρεται, να συγκρίνει παρόμοια προϊόντα και τελικά να επιλέξει αυτό που θέλει.

Επίσης, οι χρήστες μπορούν να συμμετέχουν σε συζητήσεις με άλλους καταναλωτές που έχουν δοκιμάσει το συγκεκριμένο προϊόν ή υπηρεσία και να ενημερωθούν για τις γνώμες τους. Αρκετοί δικτυακοί τόποι παρέχουν στους καταναλωτές δυνατότητες βαθμολόγησης ή κριτικής στα προϊόντα τους και τη διαδικασία αγοράς από το ηλεκτρονικό τους κατάστημα. Αυτή η δυνατότητα δεν υπάρχει στα πραγματικά καταστήματα και ακόμα και να γίνει μία κριτική είναι δύσκολο να τη μάθει κάποιος μελλοντικός αγοραστής.

Τα τελευταία χρόνια με την ανάπτυξη των κοινωνικών δικτύων (όπως το Facebook κ.ά.) κάθε χρήστης έχει το δικαίωμα να γράψει τη γνώμη του για κάτι και είναι δύσκολο να λογοκριθεί. Κανένα άλλο μέσο δε διαθέτει αυτά τα χαρακτηριστικά και ποτέ ξανά όλες αυτές οι πληροφορίες δεν ήταν τόσο εύκολα προσβάσιμες από τους καταναλωτές. Ας δούμε όμως και έμπρακτα με ποσοστά τη χρήση των διαφημίσεων στα διάφορα μέσα.

Στην Εικόνα 9 βλέπουμε τα ποσοστά που συγκεντρώνουν τα μέσα διαφήμισης παγκοσμίως για το 2014 και μία πρόβλεψη για το 2017. Σύμφωνα με την πρόβλεψη αυτή αναμένεται πολύ μεγάλη αύξηση του ποσοστού των διαφημίσεων στο mobile internet, το οποίο σχεδόν τριπλασιάζεται. Ακόμη, προβλέπεται μείωση του ποσοστού διαφημίσεων της τηλεόρασης μέσα στα επόμενα χρόνια και τα ποσοστά αθροιστικά σε internet και κινητά αναμένονται να το φτάσουν.



Εικόνα 9. Ποσοστά των μέσων διαφήμισης (Πηγή: ZenithOptimedia)

Τα χαρακτηριστικά [4,5] που κάνουν τη διαδικτυακή διαφήμιση να διαφέρει από τα υπόλοιπα μέσα συνοψίζονται στα εξής:

- Αγορά on-line

Οι καταναλωτές έχουν τη δυνατότητα να βλέπουν τα προϊόντα και να τα αγοράζουν αμέσως. Αυτό το στοιχείο διακρίνει το διαδίκτυο από τα υπόλοιπα μέσα διαφήμισης. Μπορούν να επηρεάσουν τη λήψη απόφασης για αγορά, ακόμα και να προτείνουν προϊόντα σύμφωνα με τις επιθυμίες τους. Πολλοί είναι οι καταναλωτές που προτιμούν το διαδίκτυο για τις αγορές τους.

- Αλληλεπίδραση

Οι καταναλωτές προτιμούν την αλληλεπίδραση του διαδικτύου από άλλες μορφές διαφήμισης, καθώς νιώθουν ότι έτσι διαθέτουν τον έλεγχο οι ίδιοι και όχι οι επιχειρήσεις.

- Οικειότητα πελατών

Μέσω του διαδικτύου οι καταναλωτές μπορούν να επιζητήσουν πληροφορίες για τα προϊόντα και τις υπηρεσίες που τους ενδιαφέρουν με μεγάλη ευκολία. Στις πραγματικές αγορές δεν υπάρχει αυτή η δυνατότητα. Αυτό συμβαίνει μόνο στο διαδίκτυο το οποίο παρέχει πολλαπλά μέσα επικοινωνίας. Τα οφέλη αυτής της οικειότητας είναι αμφίδρομα αφού οι καταναλωτές εξυπηρετούνται καλύτερα και οι επιχειρήσεις μαθαίνουν τις τάσεις των καταναλωτών.

2.6 Πλεονεκτήματα και μειονεκτήματα

Αρκετά είναι τα πλεονεκτήματα της διαφήμισης στο διαδίκτυο έναντι των υπόλοιπων μορφών, γι' αυτό και είναι τόσο διαδεδομένη. Τα πλεονεκτήματα αφορούν και τις δύο πλευρές, και τις επιχειρήσεις που θέλουν να διαφημιστούν και τους καταναλωτές. Το κυριότερο πλεονέκτημα [4] είναι ότι φτάνουν οι κατάλληλες διαφημίσεις στους κατάλληλους χρήστες την κατάλληλη στιγμή.

Υπάρχουν πολλές δυνατότητες στις διαδικτυακές διαφημίσεις και σπάνια μία συγκεκριμένη διαφήμιση απευθύνεται σε όλο το εύρος του καταναλωτικού κοινού. Συνήθως στοχεύει σε συγκεκριμένο τμήμα του στηριζόμενη σε κάποια χαρακτηριστικά, τα οποία είναι είτε τα δημογραφικά στοιχεία του καθενός (φύλο, ηλικία κλπ.) είτε τα ενδιαφέροντα τους.

Σημαντικό κομμάτι είναι η επικοινωνία που επιτυγχάνεται μέσω του διαδικτύου μεταξύ πελάτη και επιχείρησης. Μπορούν να επικοινωνούν είτε άμεσα είτε έμμεσα και έτσι αναπτύσσεται μία σχέση εμπιστοσύνης. Ο πελάτης μπορεί να ενημερώνεται για νέα προϊόντα και για υπηρεσίες που τον ενδιαφέρουν και η επιχείρηση μαθαίνει τι αρέσει στους πελάτες της και τι όχι, παίρνοντας feedback για τα προϊόντα τους, τυχόν βελτιώσεις ή αλλαγές κ.α.

Η επιχείρηση μπορεί να περνάει τα μηνύματα που θέλει στο καταναλωτικό κοινό με πολύ ευκολότερο τρόπο από ότι σε άλλα μέσα και χωρίς μεγάλο κόστος. Επίσης μειώνεται το κόστος προβολής πολλών μηνυμάτων σε διαφορετικές ομάδες. Σημαντική είναι η δυνατότητα αλλαγής του μηνύματος στις διαφημίσεις, κάτι το οποίο σε άλλα μέσα (έντυπα, τηλεόραση κ.α.) στοιχίζει σημαντικά σε κόστος και σε χρόνο.

Από τα μεγαλύτερα πλεονεκτήματα [5] είναι ότι το διαδίκτυο είναι το μοναδικό μέσο που μπορεί να μετρήσει την αποτελεσματικότητα μιας διαφήμισης με νούμερα και να δούμε την απήχηση που έχει στο καταναλωτικό κοινό. Καταγράφει ποιος είδε τη διαφήμιση, ποιος έκανε κλικ, πόσο έμεινε κ.α. και έτσι δίνει τη δυνατότητα να αλλαχθεί κάτι στη διαφημιστική καμπάνια ώστε να αυξηθεί η αποτελεσματικότητά της.

Ιδιαίτερα μειονεκτήματα στη διαφήμιση μέσω διαδικτύου δεν υπάρχουν. Από την μεριά της επιχείρησης το μόνο που μπορεί να θεωρηθεί ως μειονέκτημα είναι η αύξηση του κόστους λόγω του ανταγωνισμού, αλλά πάλι συνολικά σίγουρα είναι μικρότερο από ότι σε άλλα μέσα. Από την πλευρά του χρήστη όμως εγκυμονούν διάφοροι κίνδυνοι και χρειάζεται ιδιαίτερη προσοχή στις συναλλαγές.

Υπάρχει ο κίνδυνος προσβολής από ιούς [7] καθώς υπάρχουν διαφημίσεις με κακόβουλο περιεχόμενο, οι οποίες κλονίζουν την εμπιστοσύνη των χρηστών. Υπάρχει ανασφάλεια για την καταχώρηση των προσωπικών τους στοιχείων και κυρίως στη χρήση πιστωτικών καρτών, αμφιβολίες για την εγκυρότητα των πληροφοριών και ενδοιασμοί για την «παρακολούθηση» της συμπεριφοράς τους. Τέλος, η πιο ακραία περίπτωση είναι το ηλεκτρονικό έγκλημα και χρειάζεται μεγάλη προσοχή.

3 Recommender Systems

Τα τελευταία χρόνια το διαδίκτυο έχει εξαπλωθεί ραγδαία και συνεχίζει συνεχώς να εξελίσσεται με ταχύτατους ρυθμούς. Η εξέλιξη αυτή έχει σαν αποτέλεσμα την εκθετική αύξηση του όγκου πληροφοριών. Με αυτό το «χάος» των πληροφοριών που βρίσκεται διασκορπισμένο μέσα στο διαδίκτυο έρχεται αντιμέτωπος ο απλός χρήστης, ο οποίος συναντά ένα σωρό άχρηστες πληροφορίες. Απαιτείται αρκετός χρόνος και μεγάλη προσπάθεια από τον ίδιο για να ψάξει μέσα αυτόν τον τεράστιο όγκο των πληροφοριών, να συλλέξει, να αξιολογήσει, να επεξεργαστεί και τέλος να διαλέξει αυτό που τον ενδιαφέρει.

Έτσι προέκυψε η ανάγκη δημιουργίας ενός συστήματος, το οποίο θα κάνει όλη αυτή την διαδικασία για τον χρήστη. Ένα σύστημα το οποίο θα είναι σε θέση, ανάλογα με τις επιθυμίες και τις ανάγκες του κάθε χρήστη, να τον καθοδηγεί και να τον κατευθύνει ώστε να μπορέσει να εντοπίσει και να συλλέξει τα δεδομένα που πραγματικά τον ενδιαφέρουν.

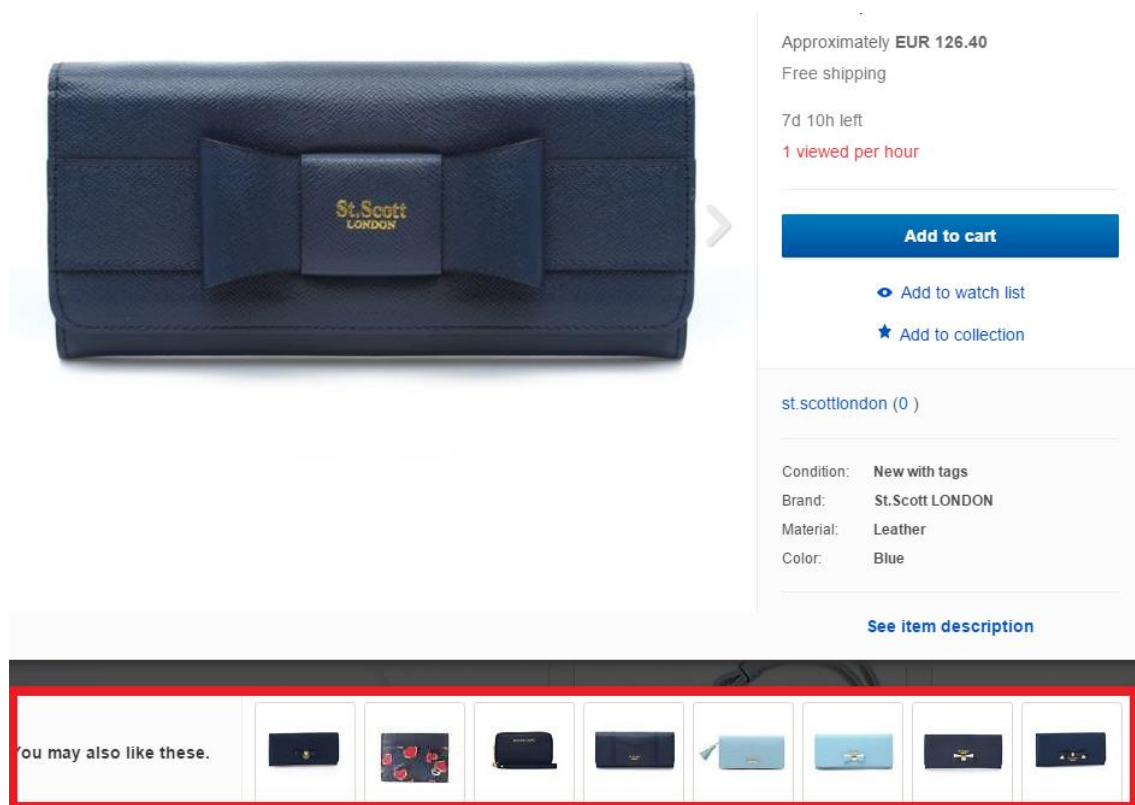
3.1 Ορισμός

Τα *Συστήματα Συστάσεων (Recommender Systems)* [9] περιγράφουν τις ευφυείς τεχνικές που παρέχουν εξατομικευμένες υπηρεσίες, γνώσεις ή συμβουλές, οι οποίες αναμένεται να ικανοποιήσουν τις απαιτήσεις του χρήστη. Τα εν λόγω συστήματα παρέχουν τεχνικές φιλτραρίσματος πληροφορίας με σκοπό την επιλεκτική προώθηση πληροφορίας σε κάθε χρήστη ξεχωριστά με κύριο γνώμονα την προσωπικότητα και τα ενδιαφέροντα του, η οποία είναι πιθανό να τον ενδιαφέρει. Η διαδικασία επιλογής και παρουσίασης των πληροφοριών αυτών στον χρήστη, αποτελεί μία πρόταση/σύσταση προς τον ίδιο.

Τα συστήματα συστάσεων άρχισαν να εμφανίζονται και να εξαπλώνονται συνεχώς τα τελευταία χρόνια. Μεγάλη είναι η δραστηριοποίηση των συστημάτων αυτών στο χώρο του ηλεκτρονικού εμπορίου καθώς οι επιτυχημένες συστάσεις μπορούν να οδηγήσουν σε αυξημένες πωλήσεις και κέρδη. Αρκετά είναι τα ηλεκτρονικά καταστήματα που χρησιμοποιούν τεχνικές συστάσεων προκειμένου να προτείνουν ενδιαφέροντα και

χρήσιμα προϊόντα αλλά και να παρέχουν πληροφορίες στους καταναλωτές προκειμένου να αποφασίσουν ποια προϊόντα θα αγοράσουν. Ένα σύστημα συστάσεων [10] μπορεί να χρησιμοποιηθεί για να καθοδηγήσει έναν πελάτη στη σωστή επιλογή στοχεύοντας στο φιλτράρισμα των άχρηστων πληροφοριών ή προβλέποντας τα ενδιαφέροντα προϊόντα αυτόματα εκ μέρους των χρηστών σύμφωνα με τις προσωπικές τους προτιμήσεις.

Τα συστήματα συστάσεων χρησιμοποιούνται σε πολλές εφαρμογές προτείνοντας προϊόντα, υπηρεσίες και πληροφορίες στοιχείων στους δυνητικούς καταναλωτές. Πολλά τα παραδείγματα των εταιρειών, όπως το eBay, η Amazon, η Procter & Gamble, οι οποίες έχουν επεκτείνει τα εμπορικά συστήματα συστάσεων αυξάνοντας με αυτό τον τρόπο τις πωλήσεις τους αλλά και την εμπιστοσύνη των πελατών τους. Βλέπουμε στην Εικόνα 10 τη σελίδα ενός προϊόντος από το eBay και από κάτω ακριβώς τις συστάσεις άλλων προϊόντων που γίνονται.



Εικόνα 10. Σελίδα προϊόντος στο eBay και οι συστάσεις

3.2 Διαδικασία σύστασης

Κάθε σύστημα συστάσεων ακολουθεί μία συγκεκριμένη διαδικασία κατά την οποία λαμβάνει κάποια δεδομένα εισόδου και τα επεξεργάζεται προκειμένου να παραχθούν αποτελέσματα. Οι πιθανές πηγές [11] από όπου μπορούν να προέλθουν τα δεδομένα εισόδου είναι:

- Τα δεδομένα χρήστη,
- Τα δεδομένα αντικειμένου και
- Οι πιθανές αλληλεπιδράσεις του χρήστη με το αντικείμενο (αξιολογήσεις, αγορές, κλικ κλπ.)

Τα δεδομένα του χρήστη συνήθως είναι δημογραφικά στοιχεία (ηλικία, φύλο κ.α.) αλλά και στοιχεία σχετικά με τις προτιμήσεις του. Όλα αυτά τα δεδομένα συνθέτουν το «προφίλ» του χρήστη, όπως ονομάζεται. Η καταγραφή των δεδομένων του χρήστη μπορεί να γίνει είτε άμεσα (explicitly) είτε έμμεσα (implicitly). Άμεσα σημαίνει ότι μόνος του ο χρήστης καταχωρεί τα στοιχεία του εφόσον του ζητηθούν. Παράδειγμα άμεσης καταχώρησης είναι το registration σε ένα eshop. Έμμεσα γίνεται όταν το σύστημα καταγράφει από μόνο του κάποια στοιχεία καθώς ο χρήστης αλληλεπιδρά μαζί του για άλλο λόγο, όπως για παράδειγμα τη στιγμή που γίνεται μία συναλλαγή.

Τα αντικείμενα περιγράφονται από ένα σύνολο χαρακτηριστικών γνωρισμάτων, το οποίο διαμορφώνει το περιεχόμενο του κάθε αντικειμένου. Και αυτή η καταχώρηση μπορεί να είναι άμεση ή έμμεση. Με άμεσο τρόπο είναι πολύ εύκολο να ληφθούν καθώς σχεδόν πάντα τα προϊόντα καταχωρούνται μαζί με τα χαρακτηριστικά τους. Έμμεσα τα στοιχεία μπορούν να εξαχθούν από κατάλληλα έγγραφα χρησιμοποιώντας τα απαραίτητα εργαλεία.

Επιπλέον, και η αλληλεπίδραση μεταξύ χρήστη και αντικειμένου μπορεί να πραγματοποιηθεί με άμεσο ή έμμεσο τρόπο [12]. Άμεσα πραγματοποιείται με τη διαδικασία της αξιολόγησης του κάθε αντικειμένου από τον χρήστη από την προσωπική του εμπειρία και συνήθως υποστηρίζεται από τα περισσότερα συστήματα. Η αξιολόγηση μπορεί να έχει διάφορες μορφές, όπως δυαδική, αριθμητική, συμβολική, κείμενο, ηχητική, οπτική κ.α. Επειδή όμως δεν είναι πάντα όλοι οι χρήστες πρόθυμοι να αξιολογήσουν τα αντικείμενα, υπάρχουν και έμμεσοι τρόποι αξιολόγησης μέσω της παρατήρησης της συμπεριφοράς τους.

3.3 Προσεγγίσεις συστάσεων

Τα πρώτα συστήματα συστάσεων που είχαν δημιουργηθεί στηρίζονταν στη χρησιμοποίηση της πληροφορίας βάσει στατιστικής ανάλυσης (πόσο έχει χρησιμοποιηθεί η κάθε πληροφορία), βάσει ιστορικής ανάλυσης (ποια πληροφορία έχει χρησιμοποιηθεί πιο πρόσφατα) ή βάσει περιεχομένου (πόσο ταιριάζει μία πληροφορία με κάποια άλλη ή με τον χρήστη). Πολλές οι προσεγγίσεις που αναπτύχθηκαν για συστάσεις, η απλούστερη από τις οποίες είναι η μη εξατομικευμένη (non-personalized).

Όμως στα περισσότερα συστήματα οι συστάσεις απευθύνονται σε έναν συγκεκριμένο χρήστη, όπως επίσης και εξατομικευμένα προϊόντα προτείνονται με σκοπό να μετατραπούν οι απλοί χρήστες σε καταναλωτές. Οι σημαντικότερες προσεγγίσεις [12] των εξατομικευμένων συστημάτων είναι η προσέγγιση βάσει περιεχομένου (content-based filtering) και η προσέγγιση συνεργατικού φιλτραρίσματος (collaborative filtering). Υπάρχουν όμως και άλλες συμπληρωματικές προσεγγίσεις, όπως η προσέγγιση βασισμένη στη γνώση (knowledge-based filtering) και η βασισμένη σε δημογραφικά στοιχεία (demography-based filtering).

3.3.1 Μη εξατομικευμένη προσέγγιση (non-personalized)

Στα συστήματα αυτά οι συστάσεις είναι ίδιες για όλους τους χρήστες. Καμία πληροφορία για τους ίδιους, αφού όλοι φαίνονται να είναι το ίδιο για το σύστημα. Οι συστάσεις αυτές είτε επιλέγονται βάσει της δημοτικότητας τους (περισσότερες εμφανίσεις, μεγαλύτερη βαθμολογία κ.α.) ή αποτελούν τα νέα προϊόντα που θέλουν να προωθηθούν. Παρόλο που η κατασκευή ενός τέτοιου συστήματος συστάσεων είναι πολύ εύκολη, δε λαμβάνει όμως καθόλου υπόψη του τα ενδιαφέροντα του κάθε χρήστη. Καθώς δεν περιλαμβάνει καθόλου εξατομίκευση αναμένεται οι συστάσεις του να μην είναι και τόσο ικανοποιητικές για τους χρήστες του.

3.3.2 Φιλτράρισμα βάσει περιεχομένου (content-based filtering)

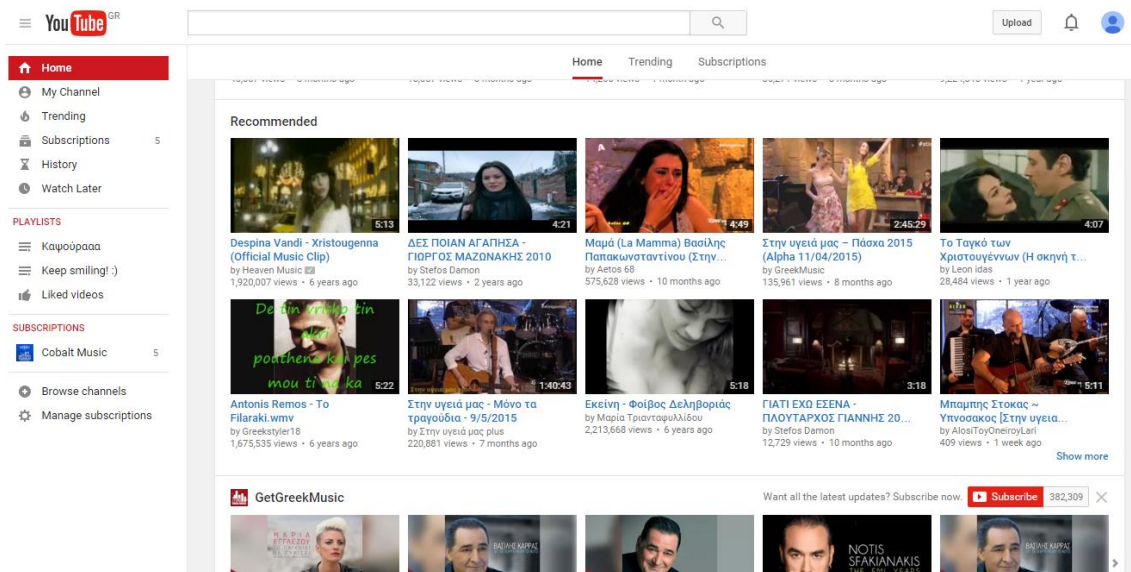
Στις προσεγγίσεις φιλτραρίσματος βάσει περιεχομένου (content-based filtering, CB), το σύστημα συστήνει αντικείμενα στον χρήστη λαμβάνοντας υπόψη τη συσχέτιση μεταξύ του περιεχομένου του αντικειμένου και των προτιμήσεων του χρήστη. Αυτή η προσέγγιση έχει πρόσβαση σε ένα σύνολο χαρακτηριστικών που παρουσιάζουν τα προϊόντα και παρόμοιες τεχνικές συναντάμε κυρίως σε συστήματα ανάκτησης πληροφοριών και σε συστήματα μηχανικής μάθησης.

Ο χρήστης εκφράζει τις ανάγκες, τα οφέλη και τους περιορισμούς και το σύστημα συγκρίνοντας το περιεχόμενο των αντικειμένων (κάποια keywords συνήθως) και το προφίλ του χρήστη εξάγει κάποια συμπεράσματα για τις προτιμήσεις του. Στη συνέχεια χρησιμοποιεί αυτή την πληροφορία για τις προτιμήσεις του χρήστη που αφορούν τα αντικείμενα και έτσι του προτείνει αντικείμενα με παρόμοια γνωρίσματα. Η λογική αυτών των συστάσεων είναι ότι όταν ένας χρήστης δείχνει κάποια ιδιαίτερη προτίμηση σε ένα σύνολο αντικειμένων με παρόμοιο περιεχόμενο, τότε είναι πολύ πιθανό ο συγκεκριμένος χρήστης να ενδιαφέρεται και για κάποιο άλλο αντικείμενο που ανήκει στην ίδια κατηγορία αντικειμένων.

Αυτές οι προσεγγίσεις συνήθως βασίζονται σε καταχωρήσεις του ίδιου του χρήστη ή στο ιστορικό προηγούμενων αγορών για να χτίσουν το προφίλ του και δεν έχουν καμία σχέση με τη γνώμη άλλων χρηστών. Οι αλγόριθμοι που υπολογίζουν την ομοιότητα ανάμεσα σε αντικείμενο και χρήστη μπορούν να λειτουργήσουν και χωρίς σύνδεση (offline). Επιπλέον, χρησιμοποιούν συχνά επεξηγήσεις σχετικά με τον τρόπο που προκύπτουν οι προτάσεις που γίνονται σε κάθε χρήστη, οπότε με αυτόν τον τρόπο αυξάνεται η εμπιστοσύνη του στο σύστημα και ταυτόχρονα και οι πιθανότητες να δεχτεί τις προτάσεις που του γίνονται. Πλεονέκτημα των συστημάτων αυτών θεωρείται η πρόταση νέων προϊόντων χωρίς να έχουν αξιολογηθεί ακόμη, αφού οι πληροφορίες που συλλέγουν προκύπτουν από την περιγραφή του ίδιου του προϊόντος.

Από την άλλη πλευρά όμως, η μέθοδος φιλτραρίσματος βάσει περιεχομένου δε μπορεί να χρησιμοποιηθεί σε όλες τις μορφές περιεχομένου. Είναι εύκολο να χρησιμοποιηθεί σε κείμενα και εικόνες, αλλά όχι τόσο σε πληροφορίες πολυμέσων όπου το περιεχόμενο τους είναι δύσκολο να οριστεί. Ένα τέτοιο σύστημα δεν μπορεί να κάνει σωστές συστάσεις αν το περιεχόμενο δεν περιέχει αρκετή πληροφορία ώστε να βγάλει συμπεράσματα για το αν αρέσει στο χρήστη ή όχι και έτσι χρειάζεται περαιτέρω ανάλυση του περιεχομένου. Άλλα προβλήματα των συστημάτων αυτών είναι οι μη αξιόπιστες συστάσεις σε νέους χρήστες αφού ακόμα δεν έχουν το προφίλ των προτιμήσεων του (γνωστό και ως cold start new user problem) και η έλλειψη δυνατότητας προσαρμογής στις αλλαγές των προτιμήσεων, όπως συμβαίνει στην πραγματικότητα (stability vs plasticity problem [13]). Τέλος, δε δίνεται η δυνατότητα στους χρήστες να εξερευνήσουν αντικείμενα, τα οποία δεν είναι παρόμοια με αυτά που ανήκουν στο προφίλ τους.

Στην Εικόνα 11 βλέπουμε τις συστάσεις που γίνονται σε κάποιον χρήστη στο YouTube. Οι συστάσεις αυτές έχουν να κάνουν με video που έχει παρακολουθήσει ο χρήστης στο παρελθόν.



Εικόνα 11. Συστάσεις βάσει περιεχομένου – YouTube

3.3.3 Συνεργατικό φιλτράρισμα (collaborative filtering)

Οι αλγόριθμοι συνεργατικού ή κοινωνικού φιλτραρίσματος (collaborative filtering or social filtering, CF) παρέχουν συστάσεις που βασίζονται στις απόψεις άλλων χρηστών με τους οποίους παρουσιάζουν παρόμοια συμπεριφορά. Η βασική διαφορά με το φιλτράρισμα βάσει περιεχομένου είναι ότι δεν απαιτούνται πληροφορίες ούτε για τους χρήστες, ούτε για τα προϊόντα. Οι συστάσεις βασίζονται καθαρά στις προτιμήσεις παρόμοιων χρηστών και όχι του ίδιου χρήστη, στον οποίο γίνονται οι προτάσεις. Δε χρειάζονται καμία πληροφορία για τα αντικείμενα και δε γνωρίζουν τίποτα γι' αυτά, παρά μόνο τι πιστεύουν οι χρήστες τους. Η παραδοχή στην οποία στηρίζονται τα συστήματα αυτά είναι ότι οι χρήστες που συμφώνησαν στο παρελθόν σε κάτι τείνουν να συμφωνήσουν ξανά στο μέλλον.

Το συνεργατικό φιλτράρισμα συχνά αναφέρεται ως η πιο ευρέως υιοθετημένη και επιτυχημένη προσέγγιση, καθώς διαθέτει κάποια χαρακτηριστικά τα οποία αντιμετωπίζουν ορισμένες αδυναμίες των συστάσεων βάσει περιεχομένου. Όπως αναφέρθηκε και πιο πριν, τα συστήματα αυτά δεν χρησιμοποιούν το περιεχόμενο των αντικειμένων και δεν απαιτούν καμία γνώση σχετικά με τα χαρακτηριστικά των

προϊόντων και καμία ανάλυση των στοιχείων αυτών. Πολύ σημαντικό είναι το γεγονός πως τα συστήματα αυτά βλέπουν πέρα από τις προτιμήσεις του κάθε χρήστη και έτσι μπορούν να προτείνουν νέα αντικείμενα στον χρήστη δημιουργώντας έτσι ανάγκες και ενδιαφέροντα που ίσως δεν είχε στο παρελθόν.

Ένα μειονέκτημα των συστημάτων αυτών είναι ότι τα προϊόντα των συστάσεων πρέπει να είναι τυποποιημένα, δηλαδή να έχουν συγκεκριμένη μορφή. Προϊόντα τα οποία δεν είναι τυποποιημένα δε μπορούν να προταθούν. Ακόμη, οι χρήστες που έχουν ασυνήθιστες προτιμήσεις ενδέχεται να αντιμετωπίσουν πρόβλημα καθώς εάν δεν υπάρχουν άλλοι χρήστες κοντά σε αυτές τις προτιμήσεις, οι συστάσεις δε θα είναι καθόλου ικανοποιητικές (grey sheep problem [12]).

Κοινό πρόβλημα που αντιμετωπίζουν και οι δύο προσεγγίσεις είναι το stability vs plasticity problem που αναφέρθηκε παραπάνω. Το πρόβλημα “αργής εκκίνησης” (cold start new item problem) εμφανίζεται όταν ένα καινούργιο αντικείμενο προστεθεί στο σύστημα και δε μπορεί να προταθεί μέχρι κάποιος να το αξιολογήσει. Η αξιολόγηση αυτή μπορεί να καθυστερήσει να πραγματοποιηθεί διότι μπορεί ο χρήστης να δυσκολευτεί να το εντοπίσει και έτσι θα καθυστερήσει ακόμα περισσότερο να ξεκινήσει να προτείνεται. Ταυτόχρονα, υπάρχει πιθανότητα ορισμένα αντικείμενα να μην αξιολογηθούν ποτέ εφόσον υπάρχει μεγάλο πλήθος αντικειμένων και λίγοι χρήστες. Οι τεχνικές συνεργατικού φιλτραρίσματος απαιτούν ένα μεγάλο ποσό ανάδρασης χρηστών (user feedback) για την ικανοποιητική παραγωγή συστάσεων.

Ορισμένες ακόμα δυσκολίες που αντιμετωπίζουν τα συστήματα συνεργατικού φιλτραρίσματος είναι η συνωνυμία (ίδια στοιχεία καταχωρούνται με διαφορετικά ονόματα ή καταχωρήσεις και έτσι εκλαμβάνονται σαν ξεχωριστά στοιχεία), οι επιθέσεις ανταγωνισμού (μπορούν να αξιολογήσουν θετικά τα δικά τους στοιχεία και αρνητικά τα ανταγωνιστικά ώστε να κερδίσουν έδαφος έναντι του ανταγωνισμού) και η διατήρηση του απορρήτου (όταν θέλουν να κρατήσουν την ιδιωτικότητα τους).

Τα συστήματα συνεργατικού φιλτραρίσματος χωρίζονται σε δύο κατηγορίες, τα συστήματα με βάση τη μνήμη (memory-based) και με βάση το μοντέλο (model-based).

3.3.3.1 Συνεργατικό φιλτράρισμα με βάση τη μνήμη (memory-based collaborative filtering)

Οι αλγόριθμοι αυτοί διατηρούν στη μνήμη τους τη συμπεριφορά των χρηστών, η οποία μπορεί να περιλαμβάνει αξιολογήσεις ή προτιμήσεις σε αντικείμενα. Έτσι προσπαθεί να

δημιουργήσει συσχετίσεις ανάμεσα σε χρήστες και αντικείμενα. Η πληροφορία αυτή αποθηκεύεται σε έναν δισδιάστατο πίνακα χρηστών-αντικειμένων. Για τις αξιολογήσεις στις memory-based προσεγγίσεις μπορεί να χρησιμοποιηθεί ολόκληρος ο πίνακας ή τμήμα του. Στην κατηγορία των αλγορίθμων αυτών υπάρχουν οι προσεγγίσεις με βάση τον χρήστη (user-based collaborative filtering) και οι προσεγγίσεις με βάση το αντικείμενο (item-based collaborative filtering). Στις Εικόνες 12 και 13 βλέπουμε το παράδειγμα της Amazon και για τις δύο περιπτώσεις. Στην Εικόνα 12 βλέπουμε την προσέγγιση με βάση το προϊόν, ενώ στην Εικόνα 13 την προσέγγιση με βάση τον χρήστη.

70 New from \$5.25
15 Collectible from \$10.95

40 New from \$7.00
3 Collectible from \$14.95

Audible trial

With over 110 million copies of his novels in print, David Baldacci is one of the most widely read storytellers in the world. Now he introduces a startling, original new character: a man with perfect memory who must solve his own family's murder.

MEMORY MAN

Amos Decker's life changed forever—twice.

Read more

Amazon Signed Editions
Books signed by your favorite authors
Shop now

Gift-wrap available.

☐ Yes, I want FREE Two-Day Shipping with Amazon Prime

Add to Cart

Turn on 1-Click ordering for this browser

Want it Wednesday, Dec. 30? Order within 70 hrs 22 mins and choose Two-Day Shipping at checkout. Details

Ship to:
Select a shipping address: ▼

☐ Buy Used \$4.99

Add to List

Have one to sell? Sell on Amazon

Frequently Bought Together

DAVID BALDACCIO
MEMORY MAN

DAVID BALDACCIO
THE GUILTY

JOHN GRISHAM
ROGUE LAWYER

Total price: \$44.08

Add all three to Cart

Add all three to List

☒ This item: Memory Man by David Baldacci Paperback \$9.91

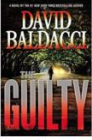
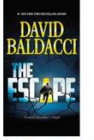
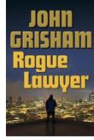
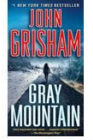
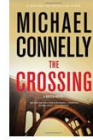

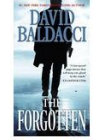
☒ The Guilty (Will Robie series) by David Baldacci Hardcover \$16.80

☒ Rogue Lawyer by John Grisham Hardcover \$17.37





Εικόνα 12. Συστάσεις με βάση το προϊόν - Amazon

☒ This Item: Memory Man by David Baldacci Paperback \$9.91
☒ The Guilty (Will Robie series) by David Baldacci Hardcover \$16.80
☒ Rogue Lawyer by John Grisham Hardcover \$17.37

Customers Who Bought This Item Also Bought Page 1 of 15

 The Guilty (Will Robie series) David Baldacci 786 reviews Hardcover \$16.80 Prime	 The Escape (John Puller Series) David Baldacci 6,658 reviews Mass Market Paperback \$6.92 Prime	 Rogue Lawyer John Grisham 3,598 reviews Hardcover \$17.37 Prime	 Gray Mountain: A Novel John Grisham 16,103 reviews Paperback \$7.54 Prime	 The Crossing (Bosch) Michael Connelly 898 reviews Hardcover \$15.10 Prime	 The Survivor (A Mitch Rapp Novel) Vince Flynn 1,877 reviews Hardcover \$16.80 Prime	 The Forgotten (John Puller) David Baldacci 4,562 reviews Mass Market Paperback \$6.22 Prime
--	--	--	--	---	--	--

Sponsored Products Related To This Item (What's this?)

				
---	---	---	---	---

Εικόνα 13. Συστάσεις με βάση το χρήστη – Amazon

Στην πρώτη περίπτωση οι συστάσεις στηρίζονται στη γειτνίαση, όπου επιλέγεται ένα σύνολο χρηστών με βάση την ομοιότητα τους με τον ενεργό χρήστη. Πλεονεκτήματα των αλγορίθμων αυτών είναι υπάρχει η δυνατότητα να προταθούν διαφορετικά αντικείμενα από αυτά που έχει επισκεφτεί και αξιολογήσει. Και αυτό συμβαίνει διότι οι συστάσεις στηρίζονται στη συμπεριφορά παρόμοιων χρηστών και όχι στη δική του.

Όμως, οι user-based αλγόριθμοι δε μπορούν να κλιμακωθούν εύκολα όταν έχουμε στο σύστημα έναν πολύ μεγάλο αριθμό χρηστών και αντικειμένων λόγω της μεγάλης υπολογιστικής πολυπλοκότητας. Ακόμη, όταν έχω μεγάλο αριθμό αντικειμένων είναι εξαιρετικά δύσκολο κάθε χρήστης να αξιολογήσει όλα τα αντικείμενα, οπότε ο πίνακας θα είναι αραιός (sparsity problem) με συνέπεια οι συστάσεις για κάποιους χρήστες να μην είναι ακριβείς.

Στη δεύτερη περίπτωση έχουμε συσχετίσεις μεταξύ των αντικειμένων αντί να έχουμε συσχετίσεις μεταξύ των χρηστών. Ο αριθμός των αντικειμένων είναι περιορισμένος οπότε οι αλγόριθμοι αυτοί μπορούν να κλιμακωθούν καλύτερα σε σχέση με τους προηγούμενους. Ακόμη υπάρχει μικρότερο πρόβλημα με αραιά δεδομένα στον πίνακα και τα αντικείμενα μπορούν να κατηγοριοποιηθούν ευκολότερα.

3.3.3.2 Συνεργατικό φιλτράρισμα με βάση το μοντέλο (model-based collaborative filtering)

Στα συστήματα αυτά δημιουργείται ένα μοντέλο προβλέψεων συστάσεων με βάση το σύνολο των δεδομένων των αξιολογήσεων των χρηστών. Έτσι για να γίνουν οι συστάσεις δε χρειάζεται όλο το σύνολο των δεδομένων αλλά παράγεται αυτόματα από το μοντέλο αυτό. Για την κατασκευή του μοντέλου χρησιμοποιούνται τεχνικές μάθησης, όπως τα νευρωνικά δίκτυα, μοντέλα συσταδοποίησης (clustering) κ.α. Προκειμένου να είναι το μοντέλο ακριβές πρέπει να είναι διαθέσιμος μεγάλος αριθμός δεδομένων.

Τα συστήματα αυτά αναιρούν τα προβλήματα του προηγούμενου μοντέλου. Είναι γρήγορα, περιορίζει το πρόβλημα με τα αραιά δεδομένα, μπορεί να διαθέτει μεγάλα σύνολα δεδομένων, έχει καλύτερη κλιμάκωση και υψηλότερη απόδοση καθώς το μοντέλο δημιουργείται πριν την online διαδικασία των συστάσεων.

Το μειονέκτημα των συγκεκριμένων αλγορίθμων είναι η δυσκολία προσαρμογής σε αλλαγές δεδομένων. Αν έχουμε πολλές αλλαγές θα πρέπει να ξαναδημιουργηθεί το μοντέλο ώστε οι συστάσεις να γίνονται σωστά με τα ανανεωμένα δεδομένα. Αυτό σημαίνει ότι θα μας κοστίζει αρκετά. Επίσης, υπάρχει η πιθανότητα να χαθεί χρήσιμη πληροφορία καθώς τα συστήματα αυτά χρησιμοποιούν συχνά μεθόδους μείωσης διαστάσεων.

3.3.4 Φιλτράρισμα βασισμένο στη γνώση (knowledge-based filtering)

Τα συστήματα φιλτραρίσματος βασισμένου στη γνώση (knowledge-based filtering, KB) βασίζονται στη γνώση που σχετίζεται με τους χρήστες και τα αντικείμενα και εξάγουν πληροφορίες που αφορούν το ποια προϊόντα ικανοποιούν τις απαιτήσεις του χρήστη. Έτσι εξαρτάται από συμπεράσματα για τις προτιμήσεις και τις ανάγκες του χρήστη και όχι από τα προϊόντα. Δε καταγράφεται ιστορικό με όλες τις προτιμήσεις των χρηστών αποφεύγοντας πολλά μειονεκτήματα που έχουν οι υπόλοιπες συστάσεις, όπως το πρόβλημα αργής εκκίνησης για ένα νέο χρήστη, για ένα νέο προϊόν ή ένα προϊόν που δε χρησιμοποιείται συχνά από τη φύση του.

Το μεγάλο πλεονέκτημα των αλγορίθμων αυτών είναι ότι δε χρειάζεται να αποθηκεύσει οποιαδήποτε πληροφορία σχετικά με τη συμπεριφορά του χρήστη. Ένας χρήστης δηλώνει μόνος του τις προτιμήσεις του, οπότε μπορεί να λάβει σωστές συστάσεις ακόμη και ως νέος χρήστης. Επίσης, αν οι προτιμήσεις του αλλάξουν τότε μπορεί

εύκολα να τις επαναπροσδιορίσει, κάτι το οποίο στα συστήματα που καταγράφουν τη συμπεριφορά του είναι αδύνατο. Το μειονέκτημα είναι το γεγονός ότι ο χρήστης μόνος του πρέπει να καταχωρήσει τα ενδιαφέροντα του προκειμένου να γίνουν συστάσεις σύμφωνα με αυτά.

3.3.5 Φιλτράρισμα βασισμένο σε δημογραφικά στοιχεία (demography-based filtering)

Η προσέγγιση του φιλτραρίσματος βασισμένου σε δημογραφικά στοιχεία (demography-based filtering, DF) στηρίζεται στα δημογραφικά στοιχεία των χρηστών (ηλικία, φύλο κ.α.) για να τους κατηγοριοποιήσει. Δημιουργεί κατηγορίες χρηστών με παρόμοια χαρακτηριστικά και παρακολουθεί τη συνολική συμπεριφορά τους. Έτσι για να γίνουν οι συστάσεις σε κάποιο χρήστη, εντοπίζεται αρχικά σε ποια κατηγορία ανήκει και κατόπιν λαμβάνονται υπόψη οι προτιμήσεις των υπόλοιπων χρηστών της ίδιας κατηγορίας.

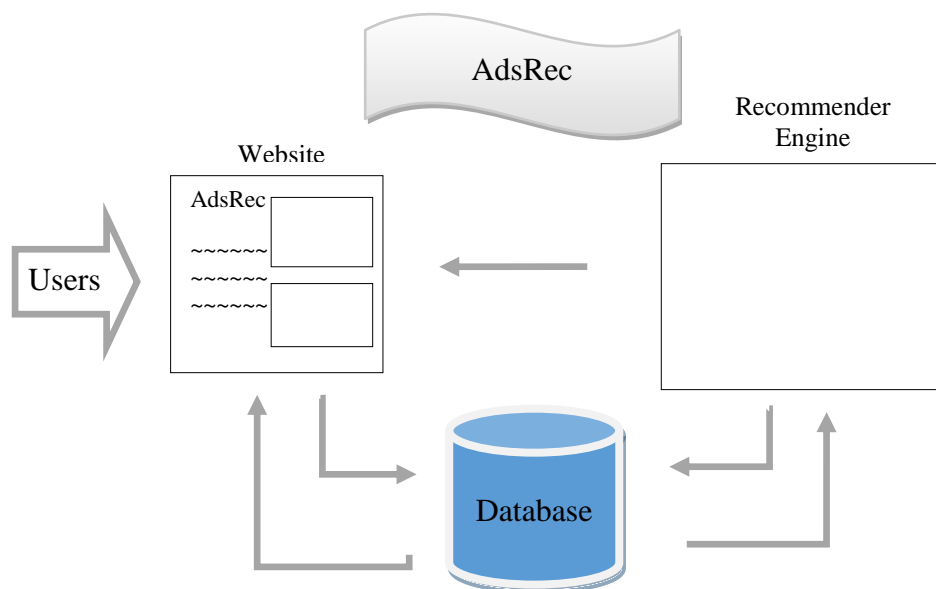
Βασικό πλεονέκτημα είναι ότι δε χρησιμοποιούνται αξιολογήσεις και έτσι νέοι χρήστες μπορούν να πάρουν συστάσεις και δεν είναι απαραίτητη καμία γνώση για τα αντικείμενα. Το μειονέκτημα είναι η δυσκολία συγκέντρωσης των δημογραφικών στοιχείων καθώς δεν είναι όλοι οι χρήστες διατεθειμένοι να τις μοιραστούν έτσι εύκολα. Υπάρχει το grey sheep πρόβλημα για τους χρήστες με ασυνήθιστη συμπεριφορά και το stability vs plasticity πρόβλημα για τους χρήστες που αλλάζουν προτιμήσεις. Η δημογραφική κατηγοριοποίηση χρειάζεται αρκετό χρόνο και οι γενικεύσεις που δημιουργούνται από αυτή είναι συχνά ψευδείς.

3.3.6 Υβριδικά συστήματα συστάσεων (hybrid recommender systems)

Τα υβριδικά συστήματα συστάσεων συνδυάζουν δύο ή περισσότερες προσεγγίσεις από αυτές που παρουσιάσαμε, με σκοπό την καλύτερη απόδοση. Συνδυάζοντας διαφορετικές προσεγγίσεις περιορίζονται οι αδυναμίες, καθώς το ένα καλύπτει τις αδυναμίες του άλλου. Ο πιο διαδεδομένος συνδυασμός είναι αυτός που συνδυάζει προσεγγίσεις με βάση το περιεχόμενο και προσεγγίσεις συνεργατικού φιλτραρίσματος. Με αυτό τον τρόπο αντιμετωπίζεται το πρόβλημα της “αργής εκκίνησης” (cold start new item problem) για ένα νέο αντικείμενο.

4 Σύστημα AdsRec

Το σύστημα AdsRec είναι ένας δικτυακός τόπος που φιλοξενεί άρθρα της επικαιρότητας και εμφανίζει ορισμένες διαφημίσεις – συστάσεις προς τους χρήστες της. Στην Εικόνα 14 εμφανίζεται η αρχιτεκτονική του συστήματος AdsRec. Αυτό απαρτίζεται από τον δικτυακό τόπο (Website), τη μηχανή συστάσεων (Recommender Engine) και τη βάση δεδομένων (Database). Ο δικτυακός τόπος AdsRec αποτελείται από ένα σύνολο άρθρων και ένα σύνολο διαφημιστικών banners. Οι χρήστες εγγράφονται/συνδέονται στο σύστημα προκειμένου να γίνουν οι κατάλληλες γι' αυτούς συστάσεις διαφημιστικών banners. Όλα αυτά τα δεδομένα καταχωρούνται σε μία βάση δεδομένων. Το σύστημα μας καταγράφει τη συμπεριφορά του κάθε χρήστη και στη συνέχεια εκτελούνται οι αλγόριθμοι σύστασης προκειμένου να εμφανίσει τα πλέον κατάλληλα διαφημιστικά banners για κάθε χρήση. Εν τέλει, οι χρήστες αξιολογούν τις συστάσεις που τους έχουν γίνει με σκοπό να μετρηθεί η αποτελεσματικότητα του συστήματος AdsRec.



Εικόνα 14. Η αρχιτεκτονική του συστήματος AdsRec

4.1 Δικτυακός τόπος

Το υποσύστημα αυτό αποτελεί τον δικτυακό τόπο (website) όπου οι χρήστες μπορούν να ενημερώνονται για διάφορα θέματα. Πρόκειται για ένα ιστολόγιο (blog) που περιλαμβάνει άρθρα από διάφορες κατηγορίες. Προσφέρει μία διεπαφή προς τους χρήστες ώστε να αλληλεπιδράσουν με το σύστημα μας και να εμφανιστούν τα κατάλληλα για τον κάθε χρήστη διαφημιστικά banners. Η μόνη προϋπόθεση είναι ότι πρέπει ο χρήστης να εγγραφεί και να συνδεθεί με τα στοιχεία του, χωρίς να είναι απαραίτητη όμως. Μπορεί κάποιος χρήστης να ενημερωθεί διαβάζοντας τα άρθρα χωρίς να κάνει κάποια εγγραφή στο σύστημα αν δεν τον ενδιαφέρουν οι εξατομικευμένες προτάσεις. Για τη δημιουργία του δικτυακού τόπου χρησιμοποιείται κώδικας html και php.

4.1.1 Αρχική σελίδα

Η αρχική σελίδα του δικτυακού τόπου μας αποτελείται από διάφορα τμήματα, πολλά από τα οποία είναι κοινά με τις υπόλοιπες σελίδες. Όπως φαίνεται στην Εικόνα 15, πάνω πάνω έχουμε τη φόρμα σύνδεσης των χρηστών. Δίπλα από τη φόρμα υπάρχει το κουμπί εγγραφής των νέων χρηστών και ακριβώς από κάτω εμφανίζεται το μενού των κατηγοριών. Τα άρθρα χωρίζονται σε κατηγορίες ανάλογα με το θέμα τους. Από κάτω έχουμε ένα slider, όπου εμφανίζονται διάφορα άρθρα και εναλλάσσονται αυτόματα. Αμέσως μετά εμφανίζονται τα τελευταία και πιο πρόσφατα άρθρα με τον τίτλο τους, την εικόνα και μια μικρή περιγραφή. Πατώντας πάνω ο χρήστης μεταφέρεται στο αντίστοιχο άρθρο, όπου μπορεί να το διαβάσει ολόκληρο το περιεχόμενό του.



Εικόνα 15. Η αρχική σελίδα

Ακόμη, στην αρχική σελίδα εμφανίζονται 4 διαφημίσεις (banners). Συνολικά έχω 4 κατηγορίες στα διαφημιστικά banners, ανάλογα με τις διαστάσεις τους. Έτσι έχω τις εξής κατηγορίες:

- 300 x 250

Είναι η πρώτη θέση πάνω πάνω. Εδώ εμφανίζονται οι προβεβλημένες διαφημίσεις, αυτοί δηλαδή που έχουν πληρώσει για να εμφανίζονται οι διαφημίσεις τους.



Εικόνα 16. Διαφημιστικό banner διάστασης 300x250

- 336 x 280

Σε αυτή τη θέση εμφανίζεται το banner με τη μεγαλύτερη απήχηση, δηλαδή με τα περισσότερα hits (κλικ)



Εικόνα 17. Διαφημιστικό banner διάστασης 336x280

- 468 x 60

Αυτή είναι η θέση του footer (τελευταίο κάτω κάτω) και εμφανίζονται τυχαία οι διαφημίσεις



Εικόνα 18. Διαφημιστικό banner διάστασης 468x60

- 728 x 90

Σε αυτή τη θέση εμφανίζονται τυχαία οι διαφημίσεις



Εικόνα 19. Διαφημιστικό banner διάστασης 728x90

Στην Εικόνα 20 βλέπουμε τη φόρμα όπου μπορεί ένας χρήστης να συνδεθεί στο σύστημα. Παρέχεται έλεγχος λογαριασμού χρήστη σε περίπτωση που δοθούν λανθασμένα στοιχεία εμφανίζοντας ένα μήνυμα λάθους.

Είσοδος: Username : Password : [Δεν είστε μέλος της ιστοσελίδας μας; Κάντε εγγραφή τώρα!](#)

Παρακαλώ συμπληρώστε τα στοιχεία σας

Εικόνα 20. Φόρμα σύνδεσης εγγεγραμμένων χρηστών

Στην περίπτωση που κάποιος χρήστης δεν είναι ήδη μέλος της ιστοσελίδας, πατώντας στο «Κάντε εγγραφή τώρα!» εμφανίζεται η φόρμα που φαίνεται στην Εικόνα 21, όπου μπορεί να καταχωρήσει τα στοιχεία του και να εγγραφεί στο σύστημα AdsRec.

Εγγραφή στην ιστοσελίδα μας

Κάνοντας εγγραφή στην ιστοσελίδα μας συμμετέχετε στο σύστημα παρακολούθησης συμπεριφορών των χρηστών και της αλληλεπίδρασης χρηστών. Πληκτρολογήστε το όνομα χρήστη που επιθυμείτε

UserName :

Πληκτρολογήστε τον κωδικό εισόδου που επιθυμείτε

Password :

Εικόνα 21. Φόρμα εγγραφής νέων χρηστών

4.1.2 Σελίδα κατηγορίας άρθρων

Υπάρχουν συνολικά έντεκα κατηγορίες άρθρων στον δικτυακό τόπο, οι οποίες επιλέχθηκαν με τυχαίο τρόπο για να αντιπροσωπεύσουν ενδεικτικά τα θέματα των άρθρων. Η κάθε κατηγορία περιέχει άρθρα που ανήκουν στο συγκεκριμένο θέμα. Για παράδειγμα, η κατηγορία «Μουσική» περιλαμβάνει άρθρα με θέμα τη μουσική. Οι κατηγορίες των άρθρων που βρίσκονται στο σύστημα μας είναι οι εξής:

- Ειδήσεις
- Αθλητικά
- Μουσική
- Lifestyle
- Τεχνολογία

- Fitness
- City Guide
- Ομορφιά
- Μόδα
- Σχέσεις
- Τουρισμός

Πατώντας σε κάποια κατηγορία, όπως φαίνεται στην Εικόνα 22, εμφανίζεται η λίστα των άρθρων με παρόμοιο τρόπο όπως στην αρχική σελίδα (τίτλος, εικόνα και μία μικρή περιγραφή). Κάνοντας κλικ σε κάθε ένα από αυτά, ο χρήστης μεταφέρεται στο άρθρο όπου μπορεί να το διαβάσει ολόκληρο. Η μόνη διαφήμιση που εμφανίζεται σε αυτές τις σελίδες είναι στο footer με τυχαίο τρόπο.

[Θα θυμίζει... Λιμάνι η «Allianz Arena»](#)



Για ακόμη μια φορά ο Ολυμπιακός θα έχει την έντονη παρουσία του κόσμου του στο πλευρό του. Πάνω από 5.000 φίλοι της ομάδας αναμένονται στη Γερμανία. Στις 21 του μήνα, το «Allianz Arena» αναμένεται να θυμίζει... «Γεώργιος Καραϊσκάκης». Πάνω από 5.000 Ολυμπιακοί αναμένεται να δώσουν έντονο ερυθρόλευκο χρώμα, κάνοντας με τις ιαχές τους το βαυαρικό γήπεδο να φέρνει σε... Φάληρο. Τα εισιτήρια που έχει λάβει επίσημα ο Ολυμπιακός είναι 4.200, τα οποία και έχουν κάνει... φτερά! Ωστόσο, σύμφωνα με τον «ΓΑΥΡΟ» αρκετά...

[Διαβάστε περισσότερα](#)

[Σοκ στον Ολυμπιακό με Γιανγκ!](#)



Άσχημα νέα στο... στρατόπεδο του Ολυμπιακού, καθώς ο Πάτρικ Γιανγκ υπέστη ρήξη πρόσθιου χιαστού αριστερού γονάτου, όπως έδειξαν οι εξετάσεις! Ο Αμερικανός σέντερ τραυματίστηκε στη διάρκεια της εκτός έδρας αναμέτρησης με την Εφές κι οι εξετάσεις στις οποίες υποβλήθηκε, έδειξαν πως υπέστη ρήξη πρόσθιου χιαστού αριστερού γονάτου. Αυτό συνεπάγεται ότι ο Γιανγκ θα μείνει εκτός δράσης από 6 έως 9 μήνες και πλέον στον Ολυμπιακό... τρέχουν! Αναλυτικά: Ο Πάτρικ Γιανγκ υποβλήθηκε το πρωί της Παρα...

[Διαβάστε περισσότερα](#)



Εικόνα 22. Η σελίδα κατηγορίας

4.1.3 Σελίδα με τα πλήρη άρθρα

Αυτή είναι η σελίδα όπου εμφανίζεται κάθε άρθρο με ολοκληρωμένο περιεχόμενο. Κάθε χρήστης, συνδεδεμένος και μη, μπορεί να διαβάσει ολόκληρο το περιεχόμενο του άρθρου και να ενημερωθεί. Αυτός είναι ο ένας από τους δύο σκοπούς του δικτυακού αυτού τόπου, να ενημερώσει τους χρήστες του. Ο τρόπος εμφάνισης του περιεχομένου του άρθρου φαίνεται στην Εικόνα 23 και περιλαμβάνει τον τίτλο, μία εικόνα και το κείμενο του άρθρου.

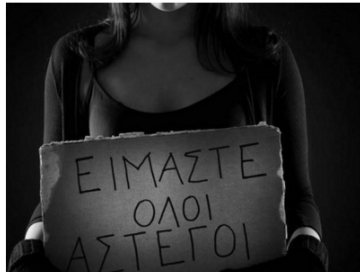
Ο δεύτερος σκοπός του δικτυακού τύπου AdsRec είναι να προτείνει διαφημιστικά banners, όσο το δυνατό πιο σχετικά σε κάθε χρήστη. Σε αυτή τη σελίδα, όπως φαίνεται και στην Εικόνα 23, έχουμε δύο κατηγορίες διαφημιστικών banners, το κάθε ένα από τα οποία προκύπτει με διαφορετικό τρόπο και έχει διαφορετικές διαστάσεις. Το πρώτο banner έχει διαστάσεις 240x200 και προκύπτει από τον αλγόριθμο σύστασης βάσει περιεχομένου και το δεύτερο με διαστάσεις 246x205 προκύπτει από τον αλγόριθμο συνεργατικού φιλτραρίσματος.



[Ειδήσεις](#) [Αθλητικά](#) [Μουσική](#) [Lifestyle](#) [Τεχνολογία](#) [Fitness](#) [City guide](#) [Ομορφιά](#) [Μόδα](#) [Σχέσεις](#) [Τουρισμός](#)

“Είμαστε όλοι άστεγοι”: Πίνοντας ένα ποτό, βοηθάμε όσους κοιμούνται στους δρόμους της Θεσσαλονίκης

Στην κατηγορία: [Ειδήσεις](#)



Μια παρέα... χιλίων ατόμων κινητοποιείται στη Θεσσαλονίκη, προκειμένου να βάλει τέλος σε μία από τις χειρότερες μορφές βίας και να στηρίξει το δικαίωμα όλων των ανθρώπων στην αξιοπρεπή διαβίωση. Η πρωτοβουλία «Είμαστε όλοι άστεγοι» έχει ως στόχο να συγκεντρωθούν χρήματα για να στηριχτούν οι άνθρωποι που έκαναν δεύτερο σπίτι τους τους δρόμους της Θεσσαλονίκης. Με την εθελοντική συμμετοχή εννέα μπαρ της πόλης και περισσότερων από 50 μουσικών, dj και καλλιτεχνών, το βράδυ της Δευτέρας 12 Ιανουαρίου θα πραγματοποιηθεί πλήθος εκδηλώσεων, τα έσοδα των οποίων, θα διατεθούν στους αστέγους. Εμπνευστής της ιδέας είναι ο Χρήστος Εξαρχόπουλος, ο οποίος τα τελευταία τρία χρόνια ζει μόνιμα στο Βερολίνο και έρχεται στη Θεσσαλονίκη κατά τη διάρκεια των εορτών. «Περπατώντας ένα βράδυ στην Τσιμισκή, είδα πάρα πολύ κόσμο να μένει στο δρόμο. Το ότι ζω στο εξωτερικό με έκανε να το συνειδητοποιήσω. Όταν ήμουν εδώ δεν το συνειδητοποιούσα γιατί ήταν κάτι που το έβλεπα κάθε μέρα και δεν το παρατηρούσα» εξηγεί στο thestival



Εικόνα 23. Η σελίδα του άρθρου

4.1.4 Βαθμολογία συστάσεων

Στο σύστημα AdsRec δίνεται στους χρήστες η δυνατότητα αξιολόγησης των προτάσεων που τους γίνονται με σκοπό να μετρηθεί η αποτελεσματικότητά τους. Έτσι κάθε εγγεγραμμένος χρήστης κάνοντας κλικ σε κάποιο banner εμφανίζεται το pop-up παράθυρο που βλέπουμε στην Εικόνα 24, όπου μπορεί να καταχωρήσει δύο βαθμολογίες. Η πρώτη (Ερώτηση 1) αφορά το πόσο σχετικό είναι το διαφημιστικό banner με το περιεχόμενο του άρθρου και δεύτερη (Ερώτηση 2) πόσο σχετικό είναι με τα ενδιαφέροντα του χρήστη. Οι βαθμολογίες κυμαίνονται σε κλίμακα από το 1 έως το 10, με το 1 να δηλώνει τη μικρότερη τιμή και το 10 τη μεγαλύτερη.

[Κλείσιμο](#)

Πόσο σχετικό είναι το διαφημιστικό banner με το περιεχόμενο του/των άρθρ-ου/-ων της σελίδας μας? (σε κλίμακα 1 μέχρι 10)

Πόσο σχετικό είναι το διαφημιστικό banner με τα δικά σας ενδιαφέροντα/προτιμήσεις σε προϊόντα/υπηρεσίες? (σε κλίμακα 1 μέχρι 10)

Αποστολή βαθμολογίας

Εικόνα 24. Pop-up παράθυρο αξιολόγησης banner

4.1.5 Προφίλ χρήστη

Οποιοσδήποτε εγγεγραμμένος χρήστης μόλις συνδεθεί έχει τη δυνατότητα να δει το ιστορικό του. Το ιστορικό περιλαμβάνει δύο πίνακες με στοιχεία από τη συμπεριφορά του όπως αυτά αποτυπώνονται στις Εικόνες 25 και 26. Ο πρώτος πίνακας (Εικόνα 25) περιλαμβάνει τη λίστα με τα άρθρα που έχει επισκεφτεί ο συνδεδεμένος χρήστης. Στην πρώτη στήλη εμφανίζεται το id του άρθρου, στη δεύτερη ο τίτλος του και στην τρίτη η ημερομηνία και ώρα προσπέλασης. Μέσα από την τέταρτη και τελευταία στήλη του πίνακα του δίνεται η δυνατότητα να διαγράψει το άρθρο που επιθυμεί από τη λίστα αυτή.

Ο δεύτερος πίνακας, που φαίνεται στην Εικόνα 26, εμφανίζει το σύνολο των banners στα οποία έχει κάνει κλικ ο χρήστης. Στην πρώτη στήλη εμφανίζεται το id του banner που πατήθηκε, στη δεύτερη βλέπουμε το banner και στην τρίτη το id του άρθρου που διάβαζε όταν εμφανίστηκε το συγκεκριμένο banner. Στην τέταρτη στήλη εμφανίζεται η βαθμολογία που καταχώρησε ο χρήστης στην Ερώτηση 1, στην πέμπτη η βαθμολογία στην Ερώτηση 2 και στην έκτη στήλη από ποιον αλγόριθμο σύστασης προήλθε το banner.

Προφίλ χρήστη elena




Μέσα από αυτή τη σελίδα μπορείτε να δείτε το προφίλ σας!

elena αυτά είναι τα άρθρα που έχετε επισκεφθεί

Id άρθρου	Όνομα άρθρου	Ημερομηνία και ώρα προσπέλασης	Ενέργειες
1	"Είμαστε όλοι άστεγοι". Πίνοντας ένα ποτό, βοηθάμε όσους κοιμούνται στους δρόμους της Θεσσαλονίκης	2015-11-8 - 3:39	Διαγραφή
2	Ποιες μέρες δε θα λειτουργήσουν τα Πανεπιστήμια λόγω εκλογών	2015-11-3 - 1:33	Διαγραφή
3	ΠΑΕ ΠΑΟΚ: Δώσατε το πρωτάθλημα εκεί που θέλετε, μήπως γλιτώσουμε αθώους ανθρώπους από τη φυλακή	2015-11-12 - 6:60	Διαγραφή
5	Η Καίτη Γαρμπή ξεκινάει εμφανίσεις στη Θεσσαλονίκη με τον Γιώργο Παπαδόπουλο στο Magique Music Hall	2015-11-3 - 10:55	Διαγραφή
6	Ελεωνόρα Ζουγανέλη: "Να με θυμάσαι και να μ'αγαπάς" βρεθήκαμε στην παρουσίαση του δίσκου	2015-11-16 - 8:38	Διαγραφή
7	Συνάντηση της Αντζελίνα Τζολί με τον Πάπα	2015-11-16 - 4:47	Διαγραφή
9	Ψυγείο που λειτουργεί χωρίς ηλεκτρισμό από Ουαλλική start-up	2015-11-30 - 11:43	Διαγραφή
10	4 είδη ψωμιού για να καταναλώσετε χωρίς ενοχές	2015-11-21 - 4:4	Διαγραφή
12	13η ExpoWedding 2015 στην Θεσσαλονίκη	2015-11-14 - 9:1	Διαγραφή
13	La Pastaria: Εγκαινιάζει το πρώτο της κατάστημα στη Θεσσαλονίκη	2015-11-18 - 8:58	Διαγραφή
14	Tips για να κάνετε σωστά την απολέπιση	2015-11-21 - 10:17	Διαγραφή
20	Νέες συντάξεις στο Font Design στο Χαλάνδρι	2015-11-12 - 2:44	Διαγραφή

Εικόνα 25. Το ιστορικό των άρθρων στο προφίλ του χρήστη

elena αυτά είναι τα banners που έχετε κάνει κλικ και η βαθμολογία τους

Id banner	banner	Id άρθρου που πατήθηκε	Πόσο σχετικό είναι το διαφημιστικό banner με το περιεχόμενο του/των άρθρ-ου/-ων της σελίδας μας? (σε κλίμακα 1 μέχρι 10)	Πόσο σχετικό είναι το διαφημιστικό banner με τα δικά σας ενδιαφέροντα/προτιμήσεις σε προϊόντα/υπηρεσίες? (σε κλίμακα 1 μέχρι 10)	Από ποιόν αλγόριθμο προέκυψε ως σύσταση
93		21	10	8	Συνεργατικό φιλτράρισμα
92		5	10	8	Βασει περιεχομένου
					

Εικόνα 26. Τα banners στο προφίλ του χρήστη

4.2 Βάση Δεδομένων

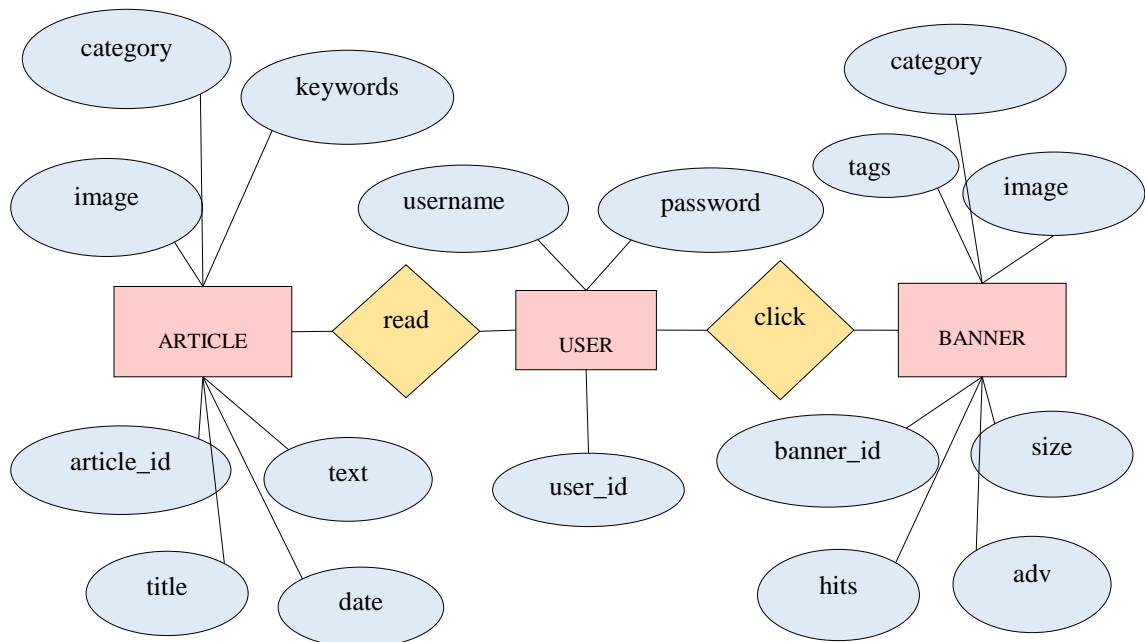
Η βάση δεδομένων περιλαμβάνει όλες τις απαραίτητες πληροφορίες για τις οντότητες του συστήματος AdsRec και τις μεταξύ τους συσχετίσεις. Η βάση του συστήματος είναι MySQL και περιλαμβάνει επτά (7) πίνακες. Σύμφωνα με το μοντέλο οντοτήτων – συσχετίσεων (μοντέλο Ο/Σ – ER model) διακρίνουμε τις εξής οντότητες στο σύστημα μας:

1. **Εγγεγραμμένοι χρήστες (Users):** Πρόκειται για ένα σύνολο χρηστών που χρησιμοποιούν το σύστημα AdsRec για την ενημέρωσή τους. Ο κάθε χρήστης έχει τα ενδιαφέροντά του και από τη στιγμή που θα συνδεθεί στο σύστημα, όλη του η συμπεριφορά καταγράφεται και έτσι διαμορφώνεται το προφίλ του. Σύμφωνα με το προφίλ αυτό γίνονται οι κατάλληλες προς αυτόν συστάσεις διαφημιστικών banners.
2. **Άρθρα (Articles):** Η οντότητα αυτή απαρτίζεται από τα άρθρα του συστήματος AdsRec και ουσιαστικά είναι το κυρίως περιεχόμενο του. Περιγράφονται από ένα σύνολο χαρακτηριστικών, τα σημαντικότερα από τα οποία είναι η κατηγορία στην οποία ανήκουν, ένα σύνολο λέξεων (keywords) και ο αριθμός των εμφανίσεών τους.
3. **Διαφημιστικά banners – διαφημίσεις (Banners):** Ένα σύνολο αντικειμένων τα οποία προτείνονται στους χρήστες ως αποτέλεσμα των αλγορίθμων σύστασης. Αναφερόμαστε σε banners τεσσάρων διαφορετικών διαστάσεων, όπου πάνω τους υπάρχουν διαφημιστικά μηνύματα. Κάποια από τα χαρακτηριστικά τους είναι η κατηγορία στην οποία ανήκουν, οι λέξεις που τα αντιπροσωπεύουν, αν είναι πληρωμένη διαφήμιση ή όχι και ο αριθμός των κλικ.

Το σύστημα AdsRec περιλαμβάνει δύο συσχετίσεις μεταξύ των οντοτήτων. Οι χρήστες αλληλεπιδρούν με το σύνολο των άρθρων και των banners. Η συμπεριφορά κάθε χρήστη καταγράφεται από τα συνολικά άρθρα που έχει διαβάσει και έτσι γίνονται οι κατάλληλες προτάσεις διαφημίσεων. Στο τέλος, βαθμολογούνται οι συστάσεις από τους χρήστες. Στο Σχήμα 1 βλέπουμε το συνολικό διάγραμμα οντοτήτων-συσχετίσεων που περιέχει αναλυτικά τις οντότητες με τα χαρακτηριστικά τους και τις συσχετίσεις τους.

Τα ορθογώνια απεικονίζουν τις τρεις οντότητες του συστήματος μας (Article, User, Banner). Οι ρόμβοι δείχνουν τις συσχετίσεις που υπάρχουν μεταξύ των οντοτήτων. Έχουμε τη συσχέτιση «read» μεταξύ ενός Article και ενός User. Να τονίσουμε ότι ένας

χρήστης μπορεί να διαβάσει πολλά άρθρα και ένα άρθρο μπορεί να διαβαστεί από πολλούς χρήστες. Η δεύτερη συσχέτιση είναι η «click» ανάμεσα σε User και Banner. Επίσης ένας χρήστης μπορεί να κάνει κλικ σε πολλά banners και ένα banner μπορεί να πατηθεί από πολλούς χρήστες.



Σχήμα 1. Το διάγραμμα οντοτήτων-συσχετίσεων

Η βάση δεδομένων περιλαμβάνει πίνακες για όλες τις οντότητες του συστήματος αλλά και κάποιους επιπρόσθετους πίνακες που μας βοηθούν να καταγράψουμε συμπεριφορές, να γίνουν οι συστάσεις και να μετρήσουμε την αποτελεσματικότητα των συστάσεων αυτών.

- mods_articles (id, title, text, category, cat_id, property, date, image_path)

Ο πίνακας αυτός περιλαμβάνει τα άρθρα του συστήματος και ένα σύνολο χαρακτηριστικών τους. Η εικόνα (image_path), ο τίτλος (title), το κείμενο (text), η ημερομηνία δημιουργίας (date) και οι λέξεις- κλειδιά (property) χαρακτηρίζουν το κάθε άρθρο. Επίσης κάθε άρθρο έχει ένα δικό του μοναδικό αριθμό (id), ώστε να μπορεί να ξεχωρίζει από οποιοδήποτε άλλο.

- mods_banners (id, hits, size, image_path, cat_id, tags, adv)

Στον πίνακα αυτό αποθηκεύονται τα διαφημιστικά banners, δηλαδή οι διαφημίσεις του συστήματος. Το σύνολο των banners χωρίζονται σε τέσσερις

κατηγορίες ανάλογα με το μέγεθος τους (size). Στη μεταβλητή hits καταχωρούμε τα κλικ που έχουν γίνει στο συγκεκριμένο banner, δηλαδή πόσες φορές πατήθηκε από το σύνολο των χρηστών. Η μεταβλητή adv μας δείχνει αν μία διαφήμιση είναι πληρωμένη ή όχι.

- mods_banners_collaboration (id, user_id, banner)

Στον πίνακα mods_banners_collaboration αποθηκεύονται οι συσχετίσεις ανάμεσα στους χρήστες και τα banners. Ουσιαστικά αποθηκεύονται όλα τα στιγμιότυπα των «κλικ» που γίνονται, τα οποία ορίζονται σε ζεύγος πληροφορίας χρήστης - banner. Αυτός ο πίνακας είναι ιδιαίτερα χρήσιμος για να καταγράψουμε τη συμπεριφορά των χρηστών και συγκεκριμένα, σε ποια banners έχουν κάνει κλικ.

- mods_categories (id, name)

Εδώ είναι αποθηκευμένες οι κατηγορίες του συστήματος. Οι κατηγορίες είναι έντεκα και ο αριθμός αυτός προέκυψε με τυχαίο τρόπο όπως αναφέρθηκε.

- mods_opinions (username, banurl, opinion1, opinion2, article, action)

Στον πίνακα mods_opinions αποθηκεύονται οι βαθμολογίες των χρηστών στα διαφημιστικά banners. Κάθε χρήστης κάνοντας κλικ σε κάποιο banner, καλείται να βαθμολογήσει (από το 1 μέχρι το 10) σύμφωνα με το πόσο χρήσιμη ήταν γι' αυτόν η διαφήμιση (opinion2) και πόσο σχετική τη θεωρεί με το περιεχόμενο της σελίδας όπου βρίσκεται (opinion1). Το χαρακτηριστικό action δείχνει από ποιον αλγόριθμο προήλθε η σύσταση. Ακόμα αποθηκεύεται το όνομα του χρήστη (username), το banner το οποίο βαθμολογεί (banurl) και το id του άρθρου όπου εμφανίστηκε το banner (article).

- mods_users (id, username, password)

Ο πίνακας mods_users αφορά τους χρήστες που είναι εγγεγραμμένοι στο σύστημα. Περιλαμβάνει το username και το password (που πρέπει να είναι μοναδικά), τα στοιχεία δηλαδή που πρέπει να εισάγει για να συνδεθεί στο σύστημα. Το id καταχωρείται αυτόματα από το σύστημα και είναι ένας μοναδικός αριθμός (και πρωτεύον κλειδί για τον πίνακα), ο οποίος ακολουθεί τον χρήστη σε όλη τη διάρκεια που είναι συνδεδεμένος και τον κάνει να ξεχωρίζει από τους υπόλοιπους χρήστες.

- mods_users_collaborate (id, user_id, activity)

Εδώ καταγράφεται συμπεριφορά κάθε χρήστη αναφορικά με τα άρθρα που έχει επισκεφτεί. Με ανάλογο τρόπο με τα banners, καταγράφουμε τον χρήστη και το άρθρο που διάβασε. Σε συνδυασμό με τον πίνακα `mods_banners_collaboration`, εξάγουμε χρήσιμα συμπεράσματα για τη συμπεριφορά κάθε χρήστη και μέσω των αλγορίθμων σύστασης, γίνονται οι συστάσεις.

4.3 Μηχανή συστάσεων (Recommender Engine)

Η μηχανή συστάσεων – Recommender Engine αποτελεί βασικό τμήμα του συστήματος, το οποίο συνδέεται με το δικτυακό τόπο και τη βάση δεδομένων. Είναι το κομμάτι το οποίο εκτελεί τους αλγορίθμους σύστασης λαμβάνοντας σαν είσοδο δεδομένα από τη βάση δεδομένων και τελικά εμφανίζει τα αποτελέσματα στους χρήστες μέσω του δικτυακού τόπου.

Στο σύστημα AdsRec υλοποιούνται δύο αλγόριθμοι συστάσεων, ο κάθε ένας από τους οποίους παράγει το δικό του αποτελέσματα που είναι μία πρόταση διαφημιστικού banner προς τον χρήστη. Έτσι εμφανίζονται δύο προτεινόμενα banners σε κάθε χρήστη. Αφού εμφανιστούν, υπάρχει η δυνατότητα αξιολόγησης των banners από τον χρήστη σύμφωνα με δύο κριτήρια. Οι αξιολογήσεις αφορούν αριθμητικές βαθμολογίες που κυμαίνονται από το 1 έως το 10, όπου το 1 αντιπροσωπεύει την ελάχιστη τιμή και το 10 τη μέγιστη.

Για την υλοποίηση της μηχανής συστάσεων κάθε αλγόριθμος αντλεί διαφορετικές πληροφορίες από τη βάση δεδομένων. Τα δεδομένα που χρησιμοποιούνται είναι πληροφορίες για τα άρθρα και τα banners, αλλά και το προφίλ των εγγεγραμμένων χρηστών. Το προφίλ κάθε χρήστη διαμορφώνεται από δεδομένα που προκύπτουν από τη συμπεριφορά του, όπως είναι τα άρθρα που έχει επισκεφτεί και τα διαφημιστικά banners στα οποία έχει κάνει κλικ. Παρακάτω αναλύεται ο τρόπος λειτουργίας των αλγορίθμων ώστε να κατανοηθεί πλήρως ο τρόπος που παράγονται οι συστάσεις προς τους χρήστες.

4.3.1 Αλγόριθμος φιλτραρίσματος βάσει περιεχομένου

Σκοπός αυτού του αλγορίθμου είναι να εμφανίσει το πιο σχετικό με το περιεχόμενο του άρθρου banner. Αυτό πραγματοποιείται με τη σύγκριση των λέξεων-κλειδίων (keywords) που έχουν καταχωρηθεί για την περιγραφή του άρθρου και των banners. Έτσι, συγκρίνονται τα δύο σύνολα μεταξύ τους και εμφανίζεται το banner με τα

περισσότερα κοινά keywords με αυτά του άρθρου. Οι διαφημίσεις που προκύπτουν από τον αλγόριθμο αυτό βρίσκονται στην πρώτη θέση δεξιά από το κείμενο του άρθρου και έχουν διάσταση 250x300.

Ο αλγόριθμος ελέγχει τη σχετικότητα ενός άρθρου με τα διαθέσιμα banners που έχει. Αφού συγκρίνει όλα τα keywords των banners με αυτά του άρθρου, μετράει πόσα κοινά υπάρχουν ανάμεσα στα δύο αυτά σύνολα και τελικά εμφανίζει το banner με τα περισσότερα. Η λειτουργία του αλγορίθμου συστάσεων βάσει περιεχομένου αναπτύσσεται στην Εικόνα 27 με ψευδοκώδικα σε λογικά βήματα και ακολουθεί αναλυτική περιγραφή. Στη συνέχεια στο Σχήμα 4 αποτυπώνεται μία γραφική απεικόνιση του αλγορίθμου.

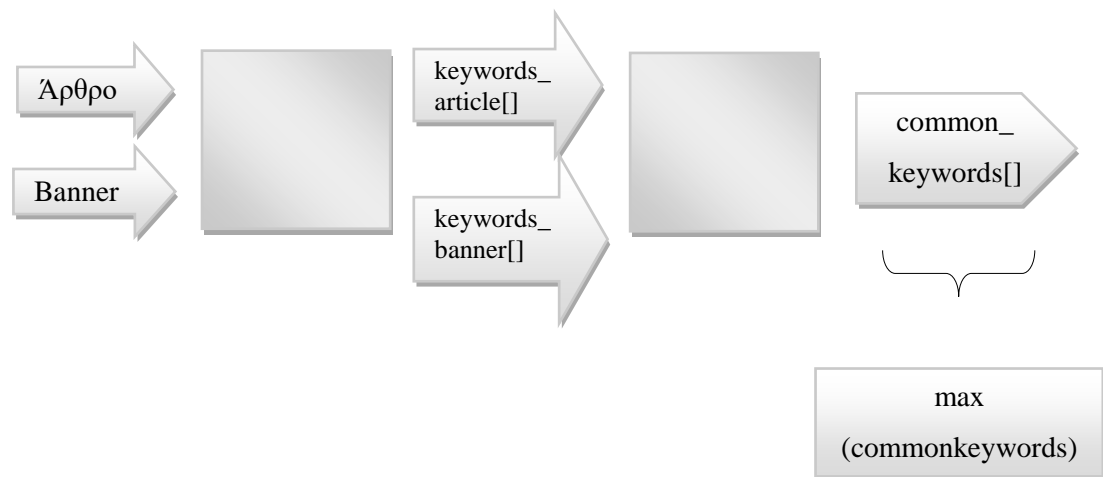
```
1. keywords_article [] = {keywords άρθρου} // μέγεθος a
2. banners [] = {banners συγκεκριμένης διάστασης, n σε αριθμό}
3. for i=0 to n // για όλα τα banners
4.     keywords_banner [] = {keywords banner} // μέγεθος b
5.     for x=0 to (a-1) // για όλα τα keywords του άρθρου
6.         for y=0 to (b-1) // για όλα τα keywords του banner
7.             if keywords_article[x]=keywords_banner[y]
8.                 ++ pointsystem // αυξάνω μετρητή
9.             Τέλος if
10.        Τέλος επαναλήψεων όσα είναι τα keywords του banner
11.    Τέλος επαναλήψεων όσα είναι τα keywords του άρθρου
12.    commonkeywords [] = {οι τιμές του pointsystem - κοινά στοιχεία}
13. Τέλος επαναλήψεων για όλα τα banners
14. Εμφανίζω banner με το max(commonkeywords)
```

Εικόνα 27. Ο αλγόριθμος σε ψευδοκώδικα

Στην Εικόνα 27 παρουσιάζονται αναλυτικά όλα τα βήματα που ακολουθεί ο αλγόριθμος σύστασης βάσει περιεχομένου. Στη γραμμή 1 τοποθετούμε σε μία δομή πίνακα (array) τα keywords του άρθρου – keywords_article[]. Στη γραμμή 2 βάζουμε όλα τα banners διάστασης 250x300 σε μία array – banners[]. Στις γραμμές 3 μέχρι 13 με τη χρήση μιας επαναληπτικής δομής δημιουργούμε (γραμμή 4) μία δομή πίνακα για κάθε banner με το

σύνολο των keywords του – keywords_banner[]. Στη συνέχεια χρησιμοποιούνται δύο επαναλήψεις από τη γραμμή 5 έως την 11 για όλα τα keywords του άρθρου και όλα τα keywords του banner και συγκρίνουμε ανά ζεύγη αν είναι ίδια. Αν βρεθεί κοινό στοιχείο αυξάνουμε τον μετρητή pointsystem (γραμμή 8). Αφού ολοκληρωθούν οι δύο επαναλήψεις στη γραμμή 12 βάζουμε την τιμή του μετρητή pointsystem σε μία δομή πίνακα και έτσι συμπληρώνουμε με τις τιμές των κοινών keywords του κάθε banner με το άρθρο – commonkeywords[]. Τέλος στη γραμμή 14, αφού τελειώσει και η επανάληψη για όλα τα banners, εμφανίζουμε το banner με τη μέγιστη τιμή στον πίνακα commonkeywords[], δηλαδή αυτό με τα περισσότερα κοινά keywords.

Στο Σχήμα 4 βλέπουμε μία σχηματική απεικόνιση του αλγορίθμου ώστε να κατανοηθεί καλύτερα η λειτουργία του. Ο αλγόριθμος λαμβάνει ως είσοδο άρθρα και banners, τα μετατρέπει σε σύνολο από keywords, τα συγκρίνει μεταξύ τους βρίσκοντας τα κοινά και τέλος εμφανίζει το banner με τα περισσότερα κοινά keywords με αυτά του άρθρου.



Σχήμα 2. Σχηματική αναπαράσταση αλγορίθμου

4.3.2 Αλγόριθμος συνεργατικού φιλτραρίσματος

Ο αλγόριθμος συνεργατικού φιλτραρίσματος (collaborative filtering) προτείνει διαφημιστικά banners στον υπό εξέταση χρήστη σύμφωνα με τις προτιμήσεις παρόμοιων με αυτόν χρηστών. Από τη στιγμή που ένας χρήστης συνδεθεί στο σύστημα

AdsRec, καταγράφεται όλη του η συμπεριφορά και διαμορφώνεται το προφίλ του. Ο αλγόριθμος αυτός στηρίζεται στη λογική ότι τα άτομα με παρόμοια συμπεριφορά τείνουν να συμφωνούν στις προτιμήσεις μελλοντικών επιλογών τους . Η συμπεριφορά τους ορίζεται από τα άρθρα που έχουν επισκεφτεί και από τα banners στα οποία έχουν κάνει κλικ.

Πρόκειται για το δεύτερο banner που εμφανίζεται δεξιά από το κείμενο του άρθρου διαστάσεων 336x280. Αναλυτικά η λειτουργία του αλγορίθμου σε βήματα παρουσιάζεται στην Εικόνα 28. Ακολουθεί η περιγραφή της λειτουργίας του αλγορίθμου και η σχηματική του απεικόνιση στο Σχήμα 5.

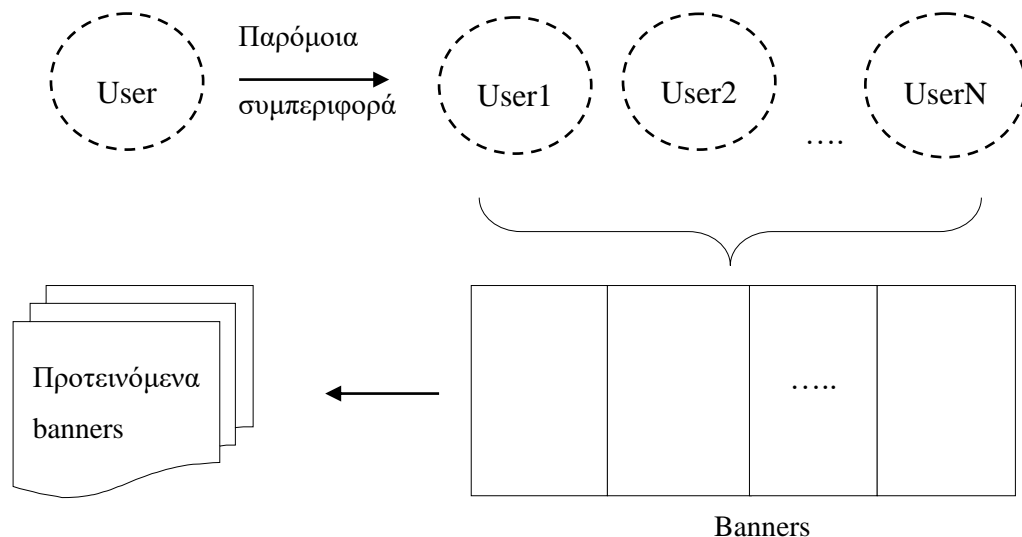
1. Βρες το id του συνδεδεμένου χρήστη (ερώτημα 1 του Παραρτήματος II)
2. Βρες τους χρήστες με κοινή συμπεριφορά με τον συνδεδεμένο χρήστη (ερώτημα 2 του Παραρτήματος II)
3. Βάλε τους χρήστες σε έναν πίνακα *collab[]*
4. Ξεκινάει επανάληψη για όλους τους κοινούς χρήστες
5. Βρες τα banners που έχει κάνει κλικ ο χρήστης (ερώτημα 3 του Παραρτήματος II)
6. Βάλε σε έναν πίνακα τα banners που έχει κάνει κλικ ο χρήστης *bancolab[]*
7. Τέλος επανάληψης
8. Βρες τα κοινά banners
9. Εμφάνισε τυχαία ένα από αυτά

Εικόνα 28. Ο αλγόριθμος συνεργατικού φιλτραρίσματος

Στην Εικόνα 28 βλέπουμε τη λειτουργία του αλγορίθμου συνεργατικού φιλτραρίσματος σε λογικά βήματα. Ξεκινάμε με το SQL Ερώτημα 1, το οποίο μπορείτε να δείτε στο Παράρτημα II και εκτελείται στη γραμμή 1 για να βρούμε το id του συνδεδεμένου χρήστη. Έχουμε αποθηκευμένη την πληροφορία για το ποια άρθρα έχει επισκεφτεί κάθε χρήστης, έτσι στη γραμμή 2 σύμφωνα με το αρκετά περίπλοκο ερώτημα 2 του Παραρτήματος II, βρίσκουμε τους χρήστες με παρόμοια συμπεριφορά (περισσότερα κοινά άρθρα) με αυτή του συνδεδεμένου χρήστη. Στη γραμμή 3 δημιουργείται μία δομή πίνακα (array) με τους χρήστες που έχουν κοινή συμπεριφορά. Στη θέση 0 της δομής τοποθετείται ο συνδεδεμένος χρήστης, ενώ στις υπόλοιπες οι χρήστες με παρόμοια

συμπεριφορά. Στη συνέχεια από τη γραμμή 4 μέχρι τη γραμμή 7 με τη χρήση μίας επαναληπτικής δομής, βρίσκουμε για κάθε χρήστη τα banners στα οποία έχει κάνει κλικ εκτελώντας στη γραμμή 5 το SQL ερώτημα 3 του Παραρτήματος II. Στη γραμμή 6 δημιουργείται για κάθε χρήστη μία δομή πίνακα (array) με το σύνολο των banners στα οποία έχει κάνει κλικ. Στη γραμμή 8 βρίσκουμε τα κοινά στοιχεία των πινάκων με τα banners που δημιουργήσαμε στη γραμμή 6 και επιλέγοντας τα κοινά τους στοιχεία συγκεντρώνουμε τα banners στα οποία έχουν κάνει κλικ οι περισσότεροι από τους χρήστες με παρόμοια συμπεριφορά. Τελικά στη γραμμή 9 εμφανίζεται ένα από αυτά τα banners με τυχαίο τρόπο.

Στο Σχήμα 5 βλέπουμε τη σχηματική απεικόνιση του αλγορίθμου συνεργατικού φιλτραρίσματος. Αρχικά βρίσκουμε τους χρήστες με παρόμοια συμπεριφορά με αυτή του συνδεδεμένου χρήστη και αυτό γίνεται βρίσκοντας κοινά άρθρα που έχουν επισκεφτεί. Κρατάμε τους χρήστες με τα περισσότερα κοινά άρθρα. Αυτό γίνεται με ερώτημα στη βάση δεδομένων (μπορείτε να το βρείτε στο Παράρτημα II). Αφού βρεθούν οι χρήστες με παρόμοια συμπεριφορά, παίρνουμε για κάθε χρήστη το σύνολο των banners που έχουν κάνει κλικ. Τέλος, βρίσκουμε τα κοινά στοιχεία αυτών των συνόλων και τελικά εμφανίζουμε ένα από αυτά.



Σχήμα 3. Σχηματική απεικόνιση του μηχανισμού συνεργατικού φιλτραρίσματος

5 Μελέτη χρηστών

Στο κεφάλαιο αυτό παρουσιάζεται μία μελέτη χρηστών που πραγματοποιήθηκε με σκοπό την αξιολόγηση των συστάσεων του συστήματος AdsRec. Οι συστάσεις γίνονται για κάθε χρήστη ξεχωριστά, οπότε δε θα υπήρχε κάποιο νόημα αν δεν μπορούσε ο ίδιος να αξιολογήσει τις συστάσεις που του γίνονται. Με τον τρόπο αυτό μας δίνεται η δυνατότητα να διαπιστώσουμε την αξία που έχει η κάθε διαφήμιση για τον ίδιο τον χρήστη.

Έτσι ένα σύνολο χρηστών επιλέγονται τυχαία και καλούνται να χρησιμοποιήσουν το σύστημα συστάσεων AdsRec για την ενημέρωσή τους. Καθώς αυτοί ενημερώνονται, το σύστημα τους προτείνει κάποια διαφημιστικά banners από δύο διαφορετικούς αλγορίθμους σύστασης. Τα banners αυτά καλούνται να αξιολογήσουν οι χρήστες σύμφωνα με κάποια κριτήρια και να τα βαθμολογήσουν.

Συλλέγοντας αυτές τις πληροφορίες, επεξεργαζόμαστε τα δεδομένα που έχουμε στα χέρια μας για να βγάλουμε κάποια χρήσιμα συμπεράσματα για το σύστημα μας. Θα χρησιμοποιήσουμε τη μέθοδο a/b testing για να συγκρίνουμε τις δύο ομάδες συστάσεων μεταξύ τους και να δούμε ποια από τις δύο υπερτερεί. Όλα αυτά θα αναλυθούν και θα παρουσιαστούν εκτενέστερα στη συνέχεια.

5.1 Συλλογή δεδομένων

Το σύστημα συστάσεων AdsRec παρέχει τη δυνατότητα βαθμολόγησης των συστάσεων που γίνονται στους χρήστες του. Κάθε εγγεγραμμένος χρήστης κάνοντας κλικ σε κάποιο διαφημιστικό banner, εμφανίζεται σε αυτόν ένα νέο παράθυρο όπου μπορεί να καταχωρήσει δύο βαθμολογίες (σε κλίμακα από το 1 μέχρι 10), οι οποίες αφορούν τα παρακάτω:

- Πόσο σχετικό είναι το διαφημιστικό banner με το περιεχόμενο της σελίδας (Ερώτηση 1) και
- Πόσο σχετικό είναι το διαφημιστικό banner με τα ενδιαφέροντα του χρήστη (Ερώτηση 2).

Στην Εικόνα 29 βλέπουμε τις βαθμολογίες που καταχωρεί ο χρήστης. Οι βαθμολογίες αυτές αποθηκεύονται σε έναν πίνακα με διάφορα στοιχεία, τον οποίο πίνακα μπορούμε να δούμε στην Εικόνα 30. Ο πίνακας έχει στο σύνολο 189 γραμμές, που σημαίνει ότι τόσες ήταν οι βαθμολογίες που καταχωρήθηκαν. Εδώ θα πρέπει να τονιστεί ότι ένα banner δε μπορεί να βαθμολογηθεί από έναν χρήστη παραπάνω από μία φορά, ενώ ένας χρήστης μπορεί να βαθμολογήσει όσα banners επιθυμεί.

[Κλείσιμο](#)

Πόσο σχετικό είναι το διαφημιστικό banner με το
περιεχόμενο του/των άρθρ-ου/-ων της σελίδας μας?
(σε κλίμακα 1 μέχρι 10)

1 ▼

Πόσο σχετικό είναι το διαφημιστικό banner με τα
δικά σας ενδιαφέροντα/προτιμήσεις σε προϊόντα/
υπηρεσίες? (σε κλίμακα 1 μέχρι 10)

1 ▼

Αποστολή βαθμολογίας

Εικόνα 29. Οι βαθμολογίες που καταχωρεί ο χρήστης για το banner

Συνολικά συμμετείχαν 20 άτομα στην αξιολόγηση ύστερα από δική μας παρότρυνση. Η επιλογή έγινε τυχαία, οι 12 προήλθαν από άμεση προσέγγιση και 8 από τα social media. Οι χρήστες κλήθηκαν να κάνουν εγγραφή στο σύστημα μας, να διαβάσουν τουλάχιστον δέκα άρθρα που τους ενδιαφέρουν και να κάνουν κλικ σε τουλάχιστον πέντε διαφημιστικά banners που τους φαίνονται χρήσιμα και κοντά στα ενδιαφέροντα τους. Κάνοντας κλικ στο διαφημιστικό banner συμμετέχουν στη μελέτη χρηστών καταχωρώντας τρεις βαθμολογίες για το κάθε banner. Η πληροφορία αυτή αποθηκεύεται στον πίνακα *mods_opinions*, την οποία καλούμαστε να αξιολογήσουμε στη συνέχεια. Υπολογίζεται ότι κατά μέσο όρο κάθε χρήστης καταχώρησε περίπου 9 βαθμολογίες.

Στην πρώτη στήλη του πίνακα που φαίνεται στην Εικόνα 30 αποθηκεύεται το username του χρήστη και στη δεύτερη το banner στο οποίο έκανε κλικ. Στη στήλη opinion1 καταχωρείται η βαθμολογία για την Ερώτηση 1 και στη στήλη opinion 2 η βαθμολογία για την Ερώτηση 2. Στη στήλη article έχουμε το id του άρθρου στο οποίο εμφανίστηκε το συγκεκριμένο banner και στη τελευταία στήλη έχουμε τον αλγόριθμο από τον οποίο προήλθε ως σύσταση. Επειδή τα δεδομένα είναι αριθμητικά βάζουμε 0 για να δηλώσουμε ότι προήλθε από το αλγόριθμο σύστασης βάσει περιεχομένου και 1 για τον αλγόριθμο συνεργατικού φιλτραρίσματος.

username	banurl	opinion1	opinion2	article	action
elena	banner31.jpg	1	5	4	0
elena	banner29.jpg	10	10	9	1
elena	banner15.jpg	9	9	6	1
elena	banner2.jpg	6	10	2	1
elena	banner26.jpg	10	8	18	1
elena	banner5.jpg	8	9	10	1
elena	banner12.jpg	10	8	17	1
elena	banner36.jpg	10	9	8	0
elena	banner80.jpg	10	7	11	0
elena	banner100.jpg	9	9	7	0
elena	banner102.jpg	10	9	3	1
elena	banner71.jpg	9	9	16	0
elena	banner68.jpg	7	9	23	1
elena	banner9.jpg	10	10	15	0

Εικόνα 30. Ο πίνακας με τις βαθμολογίες των χρηστών

5.2 Στατιστική ανάλυση

Αφού συλλέξαμε τα δεδομένα που επιθυμούσαμε για τα διαφημιστικά banners, πρέπει να τα επεξεργαστούμε ώστε να εξάγουμε χρήσιμα συμπεράσματα για το σύστημα AdsRec και τις συστάσεις του. Για να επιτευχθεί αυτό θα χρησιμοποιήσουμε την τεχνική A/B Testing, η οποία χρησιμοποιείται από πολλά online συστήματα προκειμένου να αξιολογήσει αποτελέσματα και να πετύχει μέγιστη απόδοση για το σύστημα.

Σύμφωνα με τη διαδικασία αυτή, δοκιμάζουμε δύο διαφορετικές εκδοχές ενός στοιχείου για ένα χρονικό διάστημα, μετράμε τα αποτελέσματα και καταλήγουμε στην καλύτερη επιλογή. Το στοιχείο αυτό μπορεί να είναι ένα banner, το subject ενός newsletter, ένα call-to-action κουμπί σε μία landing page, ένας τίτλος σε μία σελίδα κτλ. Χρησιμοποιείται ευρέως στο ηλεκτρονικό εμπόριο, σε καμπάνιες marketing, σε συστήματα συστάσεων και στα κοινωνικά δίκτυα για να μετρηθεί η προτίμηση των στοιχείων και η απήχηση που έχουν στο κοινό.

Θα πρέπει να οριστούν δύο διαφορετικά σύνολα ώστε να τα συγκρίνουμε μεταξύ τους και να δούμε ποιο από τα δύο υπερτερεί έναντι του άλλου. Στο δικό μας σύστημα συστάσεων έχουμε δύο διαφορετικές προτάσεις που προκύπτουν από διαφορετικούς αλγορίθμους συστάσεων και αυτούς τους δύο θα συγκρίνουμε μεταξύ τους. Έτσι θα δούμε ποιος από τους δύο είναι περισσότερο αποτελεσματικός.

5.2.1 Η διαδικασία

Για την ανάλυση μας θα χρησιμοποιήσουμε το SPSS, το πιο διαδεδομένο πρόγραμμα για τη στατιστική ανάλυση δεδομένων [14]. Ξεκινάμε εισάγοντας τα δεδομένα μας που είναι ο πίνακας `mods_opinions`. Έχουμε όλα τα δεδομένα όπως ακριβώς τα χρειαζόμαστε στον πίνακα αυτό, οπότε δε θα χρειαστεί κάποια προεργασία.

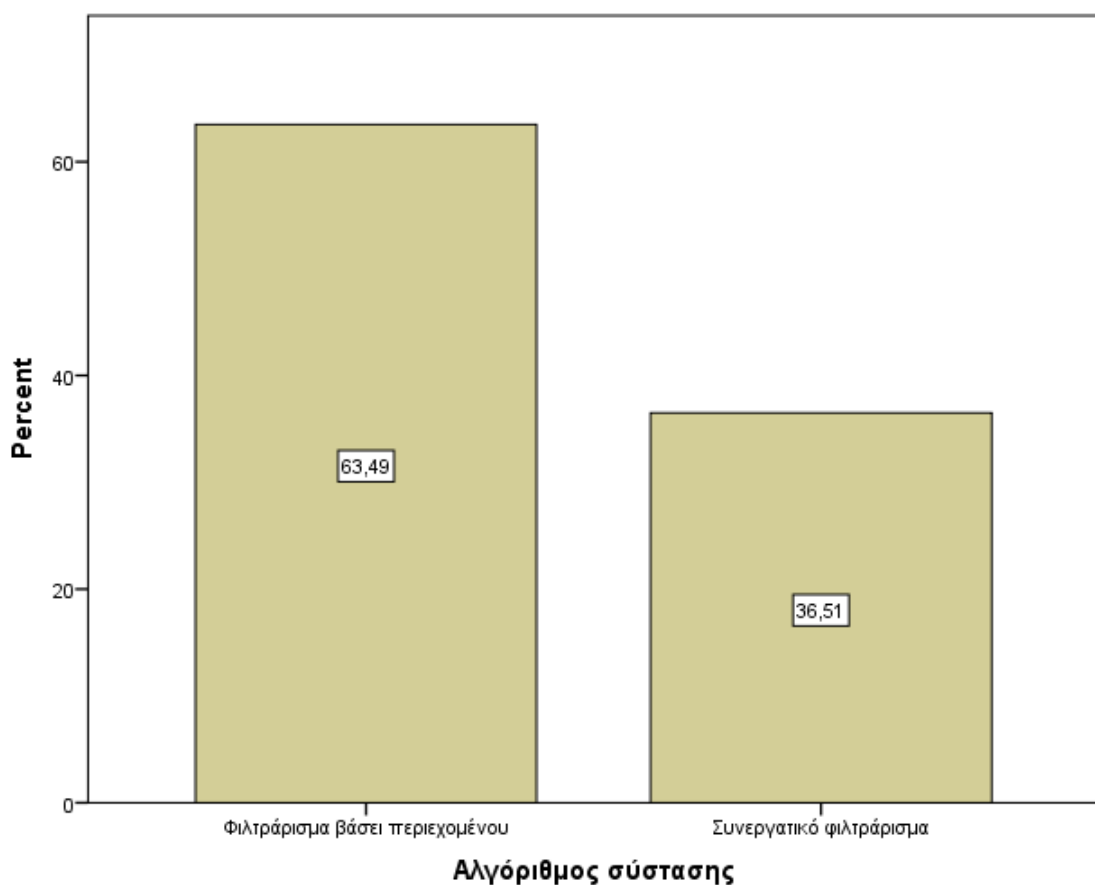
Η διάκριση των δεδομένων μας σε δύο ομάδες θα γίνει με βάση την τελευταία στήλη `Type`, η οποία δηλώνει από ποιον αλγόριθμο έχουν προέλθει οι συστάσεις. Επειδή όλα τα δεδομένα πρέπει να είναι αριθμητικά, χρησιμοποιώ μία κωδικοποίηση όπου βάζω 0 για τα banners που έχουν προέλθει από τον αλγόριθμο συστάσεων βάσει περιεχομένου και 1 για αυτά που έχουν προέλθει από τον αλγόριθμο συνεργατικού φιλτραρίσματος.

Για να ελέγξουμε τις μέσες τιμές των δύο δειγμάτων θα χρησιμοποιήσουμε το t-test. Θα ξεκινήσουμε από τη μηδενική υπόθεση, όπου οι μέσοι των δύο δειγμάτων θεωρούνται ίσοι και στο τέλος είτε τη δεχόμαστε είτε την απορρίπτουμε. Υποθέτουμε ότι τα δείγματα είναι ανεξάρτητα μεταξύ τους, δηλαδή δεν υπάρχουν συσχετίσεις των βαθμολογιών που βάζουν οι χρήστες για τις συστάσεις και από τους δύο αλγορίθμους.

Υπάρχει μία ακόμη προϋπόθεση που απαιτείται για την εκτέλεση του t-test, αυτή της κανονικότητας των δεδομένων. Αυτή η προϋπόθεση μπορεί να ξεπεραστεί με το κεντρικό οριακό θεώρημα, το οποίο μας λέει ότι όσο το μέγεθος του δείγματος τείνει στο άπειρο, η κατανομή των δεδομένων τείνει στην κανονική. Η θεωρία αναφέρει πως

ένα δείγμα των 30 παρατηρήσεων είναι ικανοποιητικό για να μας εξασφαλίσει την κανονικότητα που επιθυμούμε.

Στο Διάγραμμα 1 βλέπουμε τα ποσοστά που συγκεντρώνουν οι παρατηρήσεις μας ανάλογα με τον αλγόριθμο από τον οποίο προήλθαν. Παρατηρούμε ότι οι περισσότερες παρατηρήσεις συγκεντρώνονται στην πρώτη κατηγορία των διαφημιστικών banners, που είναι αυτά που προκύπτουν από τον αλγόριθμο σύστασης βάσει περιεχομένου. Το ποσοστό αυτό φτάνει το 63,49% έναντι του 36,51% που συγκεντρώνουν τα banners που προήλθαν από τον αλγόριθμο συνεργατικού φιλτραρίσματος. Οι περισσότεροι χρήστες δηλαδή έκαναν κλικ στα banners από την πρώτη κατηγορία.



Διάγραμμα 1. Συχνότητων των δύο κατηγοριών αλγορίθμων

5.2.2 Σχετικότητα ως προς το περιεχόμενο

Στην ενότητα αυτή θα μελετήσουμε τις απαντήσεις των χρηστών στην Ερώτηση 1 του ηλεκτρονικού ερωτηματολογίου μας, το οποίο μπορείτε να δείτε στο Παράρτημα Ι, και αφορούν στις βαθμολογίες που καταχώρησαν σύμφωνα με το πόσο σχετικό θεωρούν το

banner με το περιεχόμενο της σελίδας. Έτσι διαμορφώνονται τα δύο δείγματα με τις απαντήσεις από τους δύο αλγορίθμους. Αυτό που θέλουμε να ελέγξουμε είναι αν οι μέσες τιμές των δύο δειγμάτων είναι ίσες ή όχι για να δούμε αν η διαφορά τους είναι στατιστικά σημαντική. Οι υποθέσεις διαμορφώνονται ως εξής:

H_0 : Η διαφορά των μέσων όρων των δύο δειγμάτων μ_1 , μ_2 δεν είναι στατιστικά σημαντική και άρα θεωρούνται ίσοι

H_1 : Η διαφορά των μέσων όρων των δύο δειγμάτων μ_1 , μ_2 είναι στατιστικά σημαντική και άρα δε θεωρούνται ίσοι

Αφού ορίσαμε τις δύο υποθέσεις που χρησιμοποιούμε, τη μηδενική και την εναλλακτική, θα διαπιστώσουμε ποια από τις δύο ισχύει χρησιμοποιώντας τη μέθοδο t-test. Στο συμπέρασμα αυτό οδηγούμαστε από την τιμή του Sig στο Levene τεστ στον Πίνακα 2 στη δεύτερη στήλη των αποτελεσμάτων και από την τιμή του sig (2-tailed) στην πέμπτη στήλη. Στον Πίνακα 1 στην πρώτη στήλη βλέπουμε τον αριθμό των παρατηρήσεων που συγκεντρώνουν οι δύο κατηγορίες, στη δεύτερη στήλη τον μέσο όρο των δύο δειγμάτων, στην τρίτη την τυπική τους απόκλιση και στην τέταρτη το τυπικό σφάλμα μέσου.

Group Statistics

		N	Mean	Std. Deviation	Std. Error Mean
Ερώτηση 1	Αλγόριθμος σύστασης	120	8,35	2,372	,216
	Φιλτράρισμα βάσει περιεχομένου	69	5,71	2,521	,304

Πίνακας 1. Περιγραφικά μέτρα των δύο δειγμάτων

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Ερώτηση 1	Equal variances assumed	,641	,425	7,199	187	,000
	Equal variances not assumed			7,081	134,854	,000

Πίνακας 2. Αποτελέσματα t-test

Για να διαπιστώσουμε αν η διαφορά των μέσων όρων των δύο δειγμάτων που βλέπουμε στον Πίνακα 1 στη δεύτερη στήλη είναι στατιστικά σημαντική, ελέγχουμε τα αποτελέσματα στον Πίνακα 2. Ο πίνακας αυτός έχει δύο γραμμές αποτελεσμάτων, η

πρώτη αναφέρεται [14] στην περίπτωση που μπορούμε να υποθέσουμε ισότητα των δύο διακυμάνσεων και η δεύτερη στην περίπτωση που δε μπορούμε να υποθέσουμε ισότητα των διακυμάνσεων. Ο πίνακας είναι χωρισμένος σε δύο κατηγορίες αποτελεσμάτων, η μία αφορά το Levene test για την ισότητα των διακυμάνσεων και η άλλη περιέχει τα αποτελέσματα του t-test. Έτσι ελέγχουμε στην τρίτη γραμμή την τιμή του Sig και διαπιστώνουμε ότι είναι ίσο με $0,425 > 0,05$ που σημαίνει ότι το τεστ είναι μη σημαντικό και οι διακυμάνσεις είναι ομοιογενείς. Έτσι ελέγχουμε το sig (2-tailed) στην πρώτη γραμμή, το οποίο ισούται με $0,00 < 0,05$ άρα απορρίπτεται η μηδενική υπόθεση (H_0) και δεχόμαστε την εναλλακτική (H_1) που σημαίνει ότι η διαφορά των μέσων όρων των δύο δειγμάτων θεωρείται στατιστικά σημαντική.

5.2.3 Σχετικότητα ως προς τα ενδιαφέροντα του χρήστη

Στην ενότητα αυτή θα εξετάσουμε τις απαντήσεις των χρηστών και τις βαθμολογίες που έδωσαν στην ερώτηση 2 του ηλεκτρονικού ερωτηματολογίου, όπως αυτό φαίνεται στο Παράρτημα 1. Με τις απαντήσεις τους μετράμε τον βαθμό συσχέτισης του banner που επέλεξε ένας χρήστης με τα ενδιαφέροντα του. Χρησιμοποιούμε ξανά τις ίδιες υποθέσεις και πραγματοποιούμε τον ίδιο έλεγχο t-test και για τα δύο banners.

H_0 : Η διαφορά των μέσων όρων των δύο δειγμάτων μ_1 , μ_2 δεν είναι στατιστικά σημαντική και άρα θεωρούνται ίσοι

H_1 : Η διαφορά των μέσων όρων των δύο δειγμάτων μ_1 , μ_2 είναι στατιστικά σημαντική και άρα δε θεωρούνται ίσοι

Στον Πίνακα 3 βλέπουμε τα περιγραφικά μέτρα για τις βαθμολογίες των χρηστών που αφορούν την ερώτηση 2 του ερωτηματολογίου μας που βρίσκεται στο Παράρτημα I. Όπως παρατηρούμε στη δεύτερη γραμμή του Πίνακα 3 η μέση τιμή βαθμολογίας των χρηστών στην Ερώτηση 2 για τα banners που προτάθηκαν από τον αλγόριθμο σύστασης βάσει περιεχομένου είναι 6,70. Στην τρίτη γραμμή του πίνακα διαπιστώνουμε ότι η μέση βαθμολογία για τα banners που προέκυψαν από τον αλγόριθμο συνεργατικού φιλτραρίσματος είναι 8,87.

Group Statistics

		N	Mean	Std. Deviation	Std. Error Mean
Ερώτηση 2	Αλγόριθμος σύστασης	120	6,70	2,465	,225
	Βάσει περιεχομένου	69	8,87	1,806	,217

Πίνακας 3. Περιγραφικά μέτρα

Στον Πίνακα 4 βλέπουμε τα αποτελέσματα του t-test και για τα δύο δείγματα. Τονίζεται ότι το πρώτο δείγμα (120 παρατηρήσεις) αφορά τις βαθμολογίες χρηστών στα διαφημιστικά banners που προκύπτουν από τον αλγόριθμο σύστασης βάσει περιεχομένου, ενώ το δεύτερο δείγμα (69 παρατηρήσεις) αφορά τις βαθμολογίες των χρηστών στα banners που προκύπτουν από τον αλγόριθμο συνεργατικού φίλτραρίσματος.

Όπως φαίνεται στη γραμμή 3 του Πίνακα 4, η παράμετρος Sig είναι $0,00 < 0,05$, οπότε απορρίπτεται η υπόθεση της ισότητας των διακυμάνσεων όπως αυτή διατυπώθηκε στην ενότητα 5.5.2 Στη συνέχεια, ελέγχουμε την παράμετρο sig (2-tailed) στη γραμμή 4 του Πίνακα 4, η οποία ισούται με $0,00 < 0,05$, άρα η μηδενική υπόθεση H_0 απορρίπτεται και η διαφορά των μέσων όρων των βαθμολογιών θεωρείται στατιστικά σημαντική.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Ερώτηση 2	Equal variances assumed	17,116	,000	-6,388	187	,000
	Equal variances not assumed			-6,933	176,187	,000

Πίνακας 4. Αποτελέσματα t-test

5.3 Παρουσίαση στοιχείων

Στην ενότητα αυτή, θα παρουσιάσουμε κάποια στοιχεία των δύο αλγορίθμων σε πίνακες και διαγράμματα προκειμένου να κατανοήσουμε καλύτερα τα αποτελέσματα της στατιστικής μας ανάλυσης. Ξεκινάμε από τις βαθμολογίες των χρηστών στην Ερώτηση 1 του ερωτηματολογίου που βρίσκεται στο Παράρτημα Ι. Στον Πίνακα 5

βλέπουμε τις συχνότητες εμφάνισης των βαθμολογιών στην Ερώτηση 1 για τους δύο αλγορίθμους σύστασης.

Ερώτηση 1 * Αλγόριθμος σύστασης Crosstabulation

Count		Αλγόριθμος σύστασης		Total
		Βάσει περιεχομένου	Συνεργατικό φιλτράρισμα	
Ερώτηση 1	1	2	4	6
	2	4	3	7
	3	3	7	10
	4	4	6	10
	5	3	14	17
	6	4	10	14
	7	10	12	22
	8	8	2	10
	9	26	1	27
	10	56	10	66
Total		120	69	189

Πίνακας 5. Συχνότητες εμφάνισης βαθμολογιών σε σχέση με περιεχόμενο

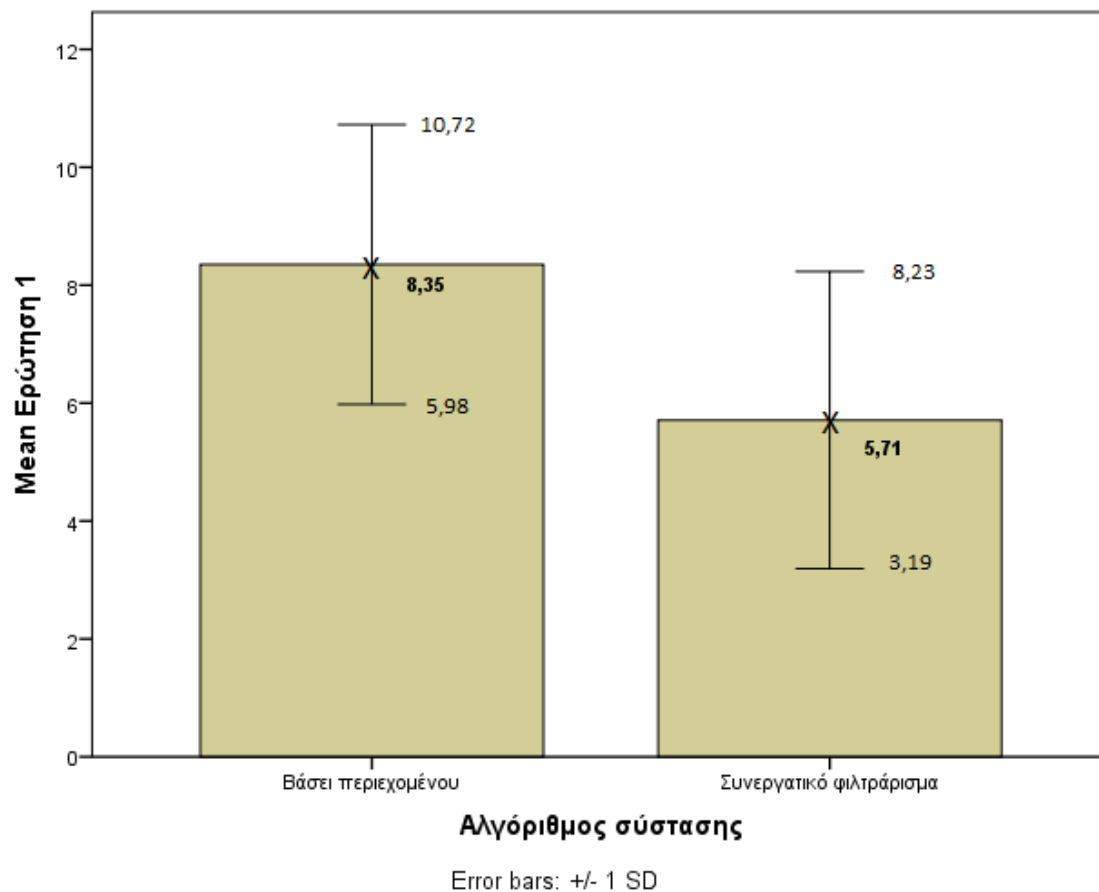
Στον Πίνακα 6 βλέπουμε τις τιμές του μέσου όρου(μ) και της τυπικής απόκλισης(σ) των βαθμολογιών που καταχώρησαν οι χρήστες για κάθε ένα από τους δύο αλγορίθμους σύστασης στην Ερώτηση 1 του ερωτηματολογίου. Όπως βλέπουμε στη δεύτερη γραμμή του πίνακα, η μέση τιμή για τον αλγόριθμο σύστασης βάσει περιεχομένου είναι $\mu_0=8,35$ και η τυπική απόκλιση $\sigma_0=2,372$. Για τον αλγόριθμο συνεργατικού φιλτραρίσματος βλέπουμε στην ίδια γραμμή ότι η μέση τιμή $\mu_1=5,71$ και η τυπική απόκλιση $\sigma_1=2,521$.

	μ_0	σ_0	μ_1	σ_1	μ_d	σ_d
Ερώτημα 1	8,35	2,372	5,71	2,521	2,64	0,149

Πίνακας 6. Αποτελέσματα βαθμολογιών σε σχέση με περιεχόμενο

Διαπιστώνουμε και στο Διάγραμμα 2 ότι μεγαλύτερες βαθμολογίες συγκεντρώνουν τα διαφημιστικά banners που προκύπτουν από τον αλγόριθμο σύστασης βάσει περιεχομένου με μέση τιμή ίση με 8,35 έναντι του 5,71 που συγκεντρώνουν αυτές που προκύπτουν από τον αλγόριθμο συνεργατικού φιλτραρίσματος. Άρα τα banners που προκύπτουν από τον αλγόριθμο βάσει περιεχομένου έχουν μεγαλύτερη απήχηση στο

καταναλωτικό κοινό και συγκεντρώνουν υψηλότερες βαθμολογίες ως προς το περιεχόμενο σε σχέση με τα banners που προκύπτουν από τον αλγόριθμο συνεργατικού φιλτραρίσματος. Τα αποτελέσματα αυτά μπορούν να θεωρηθούν αναμενόμενα και λογικά.



Διάγραμμα 2. Μέσης τιμές και τυπικής απόκλισης με βάση το περιεχόμενο

Στη συνέχεια θα μελετήσουμε τις βαθμολογίες που έδωσαν οι χρήστες στην Ερώτηση 2 του ερωτηματολογίου και αφορούν τη σχετικότητα ως προς τα ενδιαφέροντα τους. Βλέπουμε στον Πίνακα 7 τις συχνότητες των βαθμολογιών αυτών και για τους δύο αλγορίθμους σύστασης των διαφημιστικών banners.

Ερώτηση 2 * Αλγόριθμος σύστασης Crosstabulation

Count		Αλγόριθμος σύστασης		Total
		Βάσει περιεχομένου	Συνεργατικό φιλτράρισμα	
Ερώτηση 2	1	4	1	5
	2	5	0	5
	3	9	2	11
	4	4	1	5
	5	15	0	15
	6	10	0	10
	7	18	4	22
	8	20	11	31
	9	24	15	39
	10	11	35	46
Total		120	69	189

Πίνακας 7. Συχνότητες εμφάνισης βαθμολογιών σε σχέση με ενδιαφέροντα

Στον Πίνακα 8 βλέπουμε τη μέση τιμή(μ) και την τυπική απόκλιση(σ) των βαθμολογιών που συγκεντρώνουν τα διαφημιστικά banners σύμφωνα με την Ερώτηση 2 του ερωτηματολογίου. Στην δεύτερη γραμμή διαπιστώνουμε ότι η μέση τιμή των βαθμολογιών για την Ερώτηση 2 των διαφημιστικών banners που προκύπτουν από τον αλγόριθμο σύστασης βάσει περιεχομένου είναι $\mu_0=6,7$ και η τυπική απόκλιση $\sigma_0=2,465$. Στην ίδια γραμμή του πίνακα βλέπουμε ότι η μέση τιμή των βαθμολογιών των διαφημιστικών banners που προκύπτουν από τον αλγόριθμο συνεργατικού φιλτραρίσματος είναι $\mu_1=8,87$ και η τυπική απόκλιση $\sigma_1=1,806$.

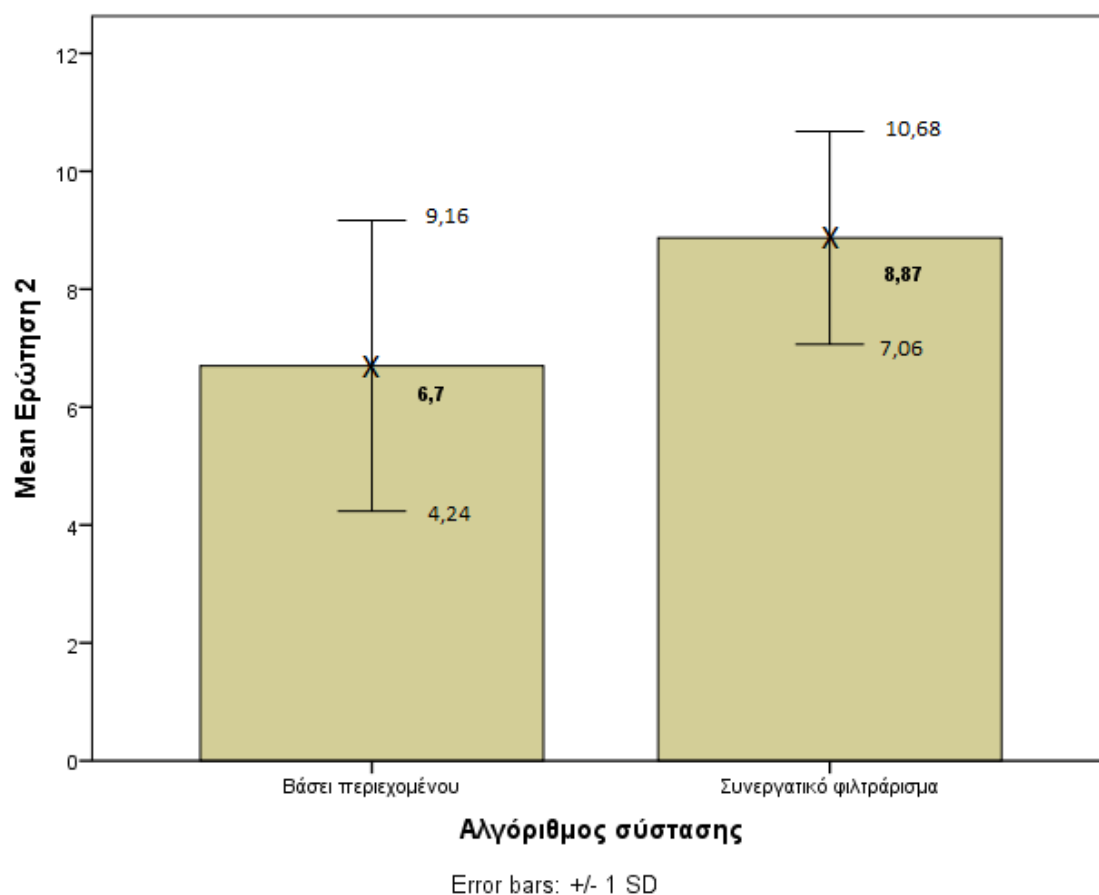
	μ_0	σ_0	μ_1	σ_1	μ_d	σ_d
Ερώτημα 2	6,7	2,465	8,87	1,806	2,17	0,659

Πίνακας 8. Αποτελέσματα βαθμολογιών σε σχέση με ενδιαφέροντα

Παρατηρούμε στον Διάγραμμα 4 ότι οι βαθμολογίες στην Ερώτηση 2 είναι υψηλότερες στα banners που προέρχονται από τον αλγόριθμο συνεργατικού φιλτραρίσματος. Οι βαθμολογίες αυτές έχουν να κάνουν με το πόσο σχετικό θεωρεί ο χρήστης ένα banner με τα ενδιαφέροντα του. Το αποτέλεσμα αυτό είναι άκρως ενθαρρυντικό καθώς

σημαίνει ότι τα banners αυτά σχετίζονται άμεσα με τα ενδιαφέροντα των χρηστών. Ιδιαίτερο ενδιαφέρον παρουσιάζει η υψηλή μέση τιμή των βαθμολογιών που ισούται με 8,87 κάτι το οποίο δείχνει ότι οι συστάσεις αυτές συνδέονται άμεσα με τα πραγματικά ενδιαφέροντα του χρήστη.

Η μέση τιμή των βαθμολογιών για την Ερώτηση 2 στα banners που έχουν προκύψει από τον αλγόριθμο σύστασης βάσει περιεχομένου είναι 6,7. Διαπιστώνουμε ότι οι βαθμολογίες είναι αρκετά χαμηλές αλλά αυτό είναι λογικό καθώς οι συστάσεις αυτές προκύπτουν από τα ενδιαφέροντα του χρήστη και τις προτιμήσεις παρόμοιων σε συμπεριφορά χρηστών και δεν έχουν καμία σχέση με το περιεχόμενο.



Διάγραμμα 3. Μέσης τιμής και τυπικής απόκλισης σε σχέση με ενδιαφέροντα

6 Συμπεράσματα

Στα πλαίσια της διπλωματικής εργασίας αναπτύχθηκε ένα σύστημα συστάσεων διαφημιστικών banners, που ονομάζεται AdsRec. Ευελπιστούμε αρχικά να αποσπάσει το ενδιαφέρον των αναγνωστών και να φανεί ένα χρήσιμο εργαλείο για διαφημίσεις στο διαδίκτυο. Πάνω σε αυτή τη φιλοδοξία μας θα αναπτύξουμε και την πιθανή εμπορική αξιοποίηση που θα μπορούσε να έχει το σύστημα AdsRec.

6.1 Γενικά συμπεράσματα

Τα οφέλη της διαφήμισης μέσω του Διαδικτύου είναι πολυάριθμα. Οι διαφημίσεις βρίσκονται παντού διασκορπισμένες μέσα στον παγκόσμιο ιστό. Φιλοξενούνται σε δικτυακούς τόπους, σε μηχανές αναζήτησης και στα Μέσα Κοινωνικής Δικτύωσης. Κοινός άξονας στον οποίο πρέπει να κινούνται όλοι οι διαφημιζόμενοι αλλά και αυτοί που φιλοξενούν τις διαφημίσεις, είναι η στόχευση της διαφήμισης στο κατάλληλο κοινό.

Με αυτό τον τρόπο τα οφέλη είναι πολλαπλά και για τις δύο πλευρές. Οι μεν διαφημιζόμενοι έχουν κέρδος καθώς το κοινό που θα δει τη διαφήμιση τους είναι αυτοί που πραγματικά ενδιαφέρονται γι' αυτήν, αλλά και οι δε που φιλοξενούν τις διαφημίσεις κερδίζουν καθώς αυξάνεται η εμπιστοσύνη των χρηστών προς το δικτυακό τους τόπο όπου προβλήθηκαν οι συγκεκριμένες διαφημίσεις.

Στα πλαίσια της διπλωματικής εργασίας αποφασίσαμε να υλοποιήσουμε ένα τέτοιο σύστημα, το οποίο υποστηρίζει αυτή την ιδέα. Έτσι δημιουργήθηκε το Σύστημα Συστάσεων AdsRec, το οποίο προσπαθεί να προτείνει σε κάθε χρήστη όσο το δυνατό πιο σχετικές προς αυτόν διαφημίσεις. Το μόνο που έχει να κάνει κάθε χρήστης είναι να συνδεθεί στο σύστημα και να ενημερωθεί για οτιδήποτε τον ενδιαφέρει. Έτσι το σύστημα καταγράφει τη συμπεριφορά του και προσπαθεί να του προτείνει τα κατάλληλα γι' αυτόν διαφημιστικά banners.

Με σκοπό να έχουμε όσο το δυνατό μεγαλύτερη αποτελεσματικότητα του συστήματος, προτείνουμε σε κάθε χρήστη διαφημιστικά banners βάσει δύο διαφορετικών αλγορίθμων σύστασης. Το ένα προκύπτει από το περιεχόμενο το οποίο έχει επισκεφτεί

ο χρήστης και το άλλο από το δικό του προφίλ σε συνδυασμό με το προφίλ των άλλων χρηστών, το οποίο διαμορφώνεται από τη συνολική του συμπεριφορά. Όπως διαπιστώθηκε στην ανάλυση που πραγματοποιήθηκε η διαφορά των δύο κατηγοριών είναι στατιστικά σημαντική, και όπως φαίνεται η μία συμπληρώνει την άλλη.

Η πρώτη κατηγορία των διαφημίσεων που έχει να κάνει με το περιεχόμενο της σελίδας, αποτελεί μία ασφαλή επιλογή διαφήμισης. Η δεύτερη κατηγορία στηρίζεται στη λογική των παρόμοιων χρηστών, που σημαίνει ότι χρήστες με παρόμοια συμπεριφορά τείνουν να συμφωνούν γενικότερα. Οπότε του προτείνεται κάτι το οποίο θα μπορούσε να μην έχει σκεφτεί ούτε καν ο ίδιος, και τελικά όντως να του αρέσει και να του φανεί χρήσιμο. Αυτή η επιλογή είναι και η πιο «ριψοκίνδυνη».

Σύμφωνα με τη μελέτη χρηστών διαπιστώθηκε ότι οι περισσότεροι από αυτούς επέλεξαν την πρώτη κατηγορία διαφημιστικών banners σε ποσοστό 63,49%. Αυτό είναι πολύ πιθανό να σχετίζεται με τη θέση του banner και όχι τόσο με πραγματικό ενδιαφέρον του χρήστη στο περιεχόμενο του. Αυτό το συμπέρασμα εξάγεται και από τις βαθμολογίες που καταχώρησαν οι χρήστες.

Βλέπουμε ότι οι χρήστες ενώ έχουν επιλέξει την πρώτη κατηγορία διαφημιστικών banners, που έχουν να κάνουν με τον αλγόριθμο σύστασης βάσει περιεχομένου, όταν καλούνται να τις αξιολογήσουν βάζουν υψηλότερες βαθμολογίες ως προς τη σχετικότητα τους με το περιεχόμενο. Όταν βαθμολογούν τα διαφημιστικά banners σχετικά με το πόσο κοντά είναι στα δικά τους ενδιαφέροντα, εκεί τα πράγματα αλλάζουν και οι βαθμολογίες είναι χαμηλότερες.

Σε αντίθεση, τα διαφημιστικά banners που προκύπτουν από τον αλγόριθμο συνεργατικού φιλτραρίσματος τα οποία συνδέονται άμεσα με τις δικές τους προτιμήσεις αλλά τις προτιμήσεις παρόμοιων σε συμπεριφορά χρηστών, συγκεντρώνουν πολύ υψηλές βαθμολογίες ως προς τη σχετικότητα τους με τα ενδιαφέροντα των χρηστών.

Έτσι προτείνουμε στους χρήστες που δεν έχουν αλληλεπιδράσει με το σύστημα τα διαφημιστικά banners που προκύπτουν βάσει περιεχομένου διότι το σύστημα δεν έχει στοιχεία για το προφίλ τους. Για τους υπόλοιπους που χρησιμοποιούν το σύστημα για κάποιο διάστημα προτείνουμε τα banners που προκύπτουν από το συνεργατικό φιλτράρισμα καθώς αναμένεται να είναι πιο κοντά στα ενδιαφέροντα τους.

6.2 Εμπορική αξιοποίηση συστήματος

Υπάρχουν αρκετές δυνατότητες εξέλιξης των αλγορίθμων που χρησιμοποιήθηκαν στο σύστημα AdsRec. Σε αυτή την εργασία υλοποιήθηκαν δύο αλγόριθμοι, ένας από την κατηγορία αλγορίθμων συστάσεων βάσει περιεχομένου και ένας από τους αλγορίθμους συνεργατικού φιλτραρίσματος. Μπορούν να εξελιχθούν οι ίδιοι, να επεκταθούν ακόμη και να προστεθούν καινούργιοι. Επιπλέον υπάρχουν και τα υβριδικά μοντέλα που υλοποιούν πάνω από έναν αλγόριθμο συνδυαστικά, κάτι το οποίο παρουσιάζει ιδιαίτερο ενδιαφέρον.

Για λόγους απλούστευσης στη διπλωματική μας δημιουργήσαμε ένα ιστολόγιο άρθρων. Το συγκεκριμένο σύστημα συστάσεων θα μπορούσε να φιλοξενηθεί σε οποιοδήποτε δικτυακό με κάποιο συγκεκριμένο περιεχόμενο. Ο μηχανισμός συστάσεων θα παραμείνει ίδιος, αλλάζοντας τα δεδομένα εισόδου. Πάρα πολλά ιστολόγια (blogs), ηλεκτρονικά καταστήματα (eshops) και μέσα κοινωνικής δικτύωσης, όπως το Facebook, το Instagram κλπ. χρησιμοποιούν παρόμοιους μηχανισμούς συστάσεων.

Η τελευταία τάση της διαφήμισης στο διαδίκτυο είναι τα κοινωνικά δίκτυα και μέσα τα οποία χρησιμοποιούνται ευρέως από τους χρήστες του διαδικτύου. Μερικά από τα πιο διαδεδομένα μέσα είναι το Facebook, το Instagram, το Viber, το YouTube και οι εφαρμογές σε κινητά. Είναι δηλαδή μέσα τα οποία βρίσκονται στην καθημερινότητα του χρήστη και έτσι μπορούν να εκλάβουν πληροφορίες για τους ίδιους και να τους παρέχουν απόλυτα εξατομικευμένες διαφημίσεις. Με κάποιες τροποποιήσεις οι αλγόριθμοι μας μπορούν να χρησιμοποιηθούν σε όλα αυτά τα μέσα.

Βέβαια όλα αυτά τα συστήματα ξοδεύουν τεράστιους όγκους χρόνου και χρήματος για την καταγραφή της συμπεριφοράς των εκατομμυρίων χρηστών του διαδικτύου, αντιμετωπίζοντας και τα ανάλογα προβλήματα. Το διαδικτυακό marketing θεωρείται από τους πιο διαδεδομένους και κερδοφόρους κλάδους στην προώθησης μιας επιχείρησης ή ενός προϊόντος. Διαθέτει πάρα πολλές πτυχές και απαιτεί πολύ καλή γνώση του αντικειμένου καθώς δεν είναι καθόλου εύκολη υπόθεση λόγω του τεράστιου ανταγωνισμού που υπάρχει.

Τα κυριότερα προβλήματα της διαφήμισης μέσω διαδικτύου που πρέπει να αντιμετωπιστούν είναι ο φόβος των χρηστών να μοιραστούν πληροφορίες που αφορούν τους ίδιους καθώς επίσης και ο τρόπος που αντιμετωπίζουν οι ίδιοι τον καταιγισμό από τις διαφημίσεις που δέχονται καθημερινά. Η λύση και για τα δύο προβλήματα αυτά είναι να παρουσιάζονται στους χρήστες μόνο διαφημίσεις που τους αφορούν. Με αυτό

τον τρόπο αλλάζει η στάση τους απέναντι στις διαφημίσεις και η προθυμία τους για να μοιραστούν πληροφορίες που τους αφορούν μεγαλώνει. Φαίνεται ότι η παροχή εξατομικευμένων διαφημίσεων αποτελεί μονόδρομο.

Τέλος, θέλουμε να επισημάνουμε ότι ο πρωταρχικός μας στόχος έχει επιτευχθεί. Το σύνολο των χρηστών που επισκέφτηκαν τον δικτυακό τόπο διαβάζοντας τα άρθρα που τους ενδιαφέρουν βαθμολόγησαν θετικά τα διαφημιστικά banners που τους προτάθηκαν. Μελλοντικά προσδοκούμε το Σύστημα Συστάσεων AdsRec να αποτελέσει αφορμή για περεταίρω μελέτη και επέκταση του σε ένα εμπορικό σύστημα διαφημίσεων και θα ήμασταν περήφανοι να έχουμε συμβάλει στην επίτευξη αυτού του σκοπού.

Βιβλιογραφία

- [1] <https://el.wikipedia.org/wiki/%CE%94%CE%B9%CE%B1%CF%86%CE%AE%CE%BC%CE%B9%CF%83%CE%B7> (προσπελάσθηκε 17/11/2015)
- [2] Ζώτος Γ. Χ. (2008). “Διαφήμιση”, Θεσσαλονίκη: University Studio Press (5η Έκδοση)
- [3] Ducoffe R. H., “Advertising value and advertising on the web.” Journal of advertising Research, October 1996, Vol. 36, September-October, pp. 21-36
- [4] Βλαχόπουλου Μ. & Πατσιούρα Φ., (2000), “Η έννοια και λειτουργία της On line Διαφήμισης και η σχέση με την παραδοσιακή διαφήμιση”, εκδόσεις Rossili
- [5] Ιωάννου Σ., (2008), “Το Διαδίκτυο κερδίζει μερίδια στη διαφημιστική δαπάνη”, Ναυτεμπορική
- [6] <http://www.internetinfo.gr/internetadvertising/index.html> (προσπελάσθηκε 20/11/2015)
- [7] <https://advmodels.wordpress.com/category/> (προσπελάσθηκε 17/11/2015)
- [8] Shahabi C., Chen Y. S., “An Adaptive Recommendation System without Explicit Acquisition of User Relevance Feedback”, Kluwer Academic Publishers, Netherlands, 2003
- [9] Ricci F., Venturini A., Cavada D., Mirzadeh N., Blaas D. and Nones M., “Product Recommendation with Interactive Query Management and Twofold Similarity”, eCommerce and Tourism Research Laboratory ITC-irst via Sommarive 18, 38050 Povo, Italy
- [10] Dianshuang W., Jie L., Guangquan Z., Hua L., “A fuzzy matching based recommendation approach for mobile products/services”, 10th International Conference on Intelligent Systems Design and Applications, 2010
- [11] Reischach F. V, Michahelles F., Schmidt A., “The Design Space of Ubiquitous Product Recommendation Systems”, Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia, 1-10, 2009.

- [12] Hiralall M., “Recommender systems for e-shops”, Vrije Universiteit, Amsterdam, 2011
- [13] Schafer J. B., Konstan J. A., Riedl J., “E-Commerce Recommendation Applications”, GroupLens Research Project, Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN 55455
- [14] Πραμαγγιούλης Π. (2008), “Οδηγός Ανάλυσης Δεδομένων με τη χρήση SPSS”, Οικονομικό Πανεπιστήμιο Αθηνών, Τμήμα Στατιστικής
- [15] Matic A., Pielot M., Oliver N., “The value of highly-personalized advertising”, ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015
- [16] Vargiu E., Giuliani A. and Armano, G. 2013. Improving contextual advertising by adopting collaborative filtering. ACM Trans. Web 7,3, Article 13 (September 2013)
- [17] Yan J., Liu N., Wang G., Zhang W., Jiang Y. and Chen Z. “How much can behavioral targeting help online advertising?” In Proc WWW '09, ACM (2009), 261-270
- [18] Vargiu E. And Urru M. 2013 “Exploiting web scraping in a collaborative filtering-based approach to web advertising.”
- [19] <http://www.pathfinder.gr/stories/3970485/san-shmera-emfanisthke-to-protodiafhmistiko-banner-sto-internet/> (προσπελάσθηκε 21/11/2015)
- [20] Leskovec J., “Mining of Massive Datasets”, Stanford Univ. Anand Rajaraman.
- [21] Vargiu E., Giuliani A. and Armano G., “Improving Contextual Advertising by Adopting Collaborative Filtering”
- [22] Symeonidis P., Technical Report: “Stream Mining of Social Networks for Recommender Systems”, 2013.
- [23] Yang W. S., Cheng H. C., Dia J. B., Lin H. T., “Mining Social Networks for Targeted Advertising”, HICSS 2006
- [24] Gupta M., “Predicting Click Through Rate for Job Listings”, 2009.
- [25] Dembczynski K. Kotlowski W. Weiss D., “Predicting Ads’ Click-Through Rate with Decision Rules”, Workshop on Targeting and Ranking in Online Advertising 2008
- [26] Mehta A., Saberi A. Umesh V. Vazirani V., “Adwords and Generalized On-line Matching”, FOCS 2005: 264-273. [c10]

Παράρτημα I

Παραθέτουμε το ηλεκτρονικό ερωτηματολόγιο του συστήματος AdsRec, στο οποίο καταχωρούν τις βαθμολογίες τους οι χρήστες για τα διαφημιστικά banners που τους προτείνονται. Οι ερωτήσεις είναι δύο, η πρώτη αφορά το πόσο σχετικό θεωρούν οι χρήστες το διαφημιστικό banner με το περιεχόμενο της σελίδας και η δεύτερη πόσο σχετικό το θεωρούν με τα ενδιαφέροντα τους. Κάθε χρήστης καταχωρεί τις βαθμολογίες σύμφωνα με το πόσο σχετικό θεωρεί το banner σε κλίμακα από το 1 έως το 10. Το 1 είναι η μικρότερη τιμή και το 10 η μεγαλύτερη. Αφού καταχωρήσει τις βαθμολογίες του, πατάει το κουμπί «Αποστολή βαθμολογίας».

[Κλείσιμο](#)

Πόσο σχετικό είναι το διαφημιστικό banner με το περιεχόμενο του/των άρθρ-ου/-ων της σελίδας μας? (σε κλίμακα 1 μέχρι 10)

1 ▼

Πόσο σχετικό είναι το διαφημιστικό banner με τα δικά σας ενδιαφέροντα/προτιμήσεις σε προϊόντα/υπηρεσίες? (σε κλίμακα 1 μέχρι 10)

1 ▼

Αποστολή βαθμολογίας

Παράρτημα II

Στο Παράρτημα II θα βρείτε τα ερωτήματα SQL που γίνονται στη βάση.

Ερώτημα 1:

```
SELECT id
FROM mods_users
WHERE username='X';
```

Ερώτημα 2:

```
SELECT count(us.user_id),us.user_id
FROM (SELECT distinct user_id, activity
      FROM mods_users_collaborate ) us
INNER JOIN ( SELECT activity
            FROM mods_users_collaborate
            WHERE user_id ='X]. '
            GROUP BY activity) act
on act.activity = us.activity
GROUP BY us.user_id
order by count(us.user_id)
Desc LIMIT 6;
```

Ερώτημα 3:

```
SELECT user_id, banner
FROM mods_banners_collaborations
WHERE user_id='.$collaborations[$p];
```

