

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

Μεταπτυχιακό Πρόγραμμα Σπουδών
Κοινωνικά Πληροφοριακά Συστήματα

Μεταπτυχιακή Διατριβή



Σύστημα Συστάσεων Ταξιδιωτικών Προορισμών:
Προσέγγιση Συνεργατικού Φιλτραρίσματος με Βάση τα
Προφίλ και τις Προτιμήσεις των Χρηστών

Ιωάννα Μόκα

Επιβλέπων Καθηγητής
Κλήμης Νταλιάνης

Μάιος 2018

Ανοικτό Πανεπιστήμιο Κύπρου

Σχολή Θετικών και Εφαρμοσμένων Επιστημών

**Μεταπτυχιακό Πρόγραμμα Σπουδών
*Κοινωνικά Πληροφοριακά Συστήματα***

Μεταπτυχιακή Διατριβή

**Σύστημα Συστάσεων Ταξιδιωτικών Προορισμών:
Προσέγγιση Συνεργατικού Φιλτραρίσματος με Βάση τα
Προφίλ και τις Προτιμήσεις των Χρηστών**

Ιωάννα Μόκα

**Επιβλέπων Καθηγητής
Κλήμης Νταλιάνης**

Η παρούσα μεταπτυχιακή διατριβή υποβλήθηκε προς μερική εκπλήρωση των απαιτήσεων για απόκτηση μεταπτυχιακού τίτλου σπουδών στα Κοινωνικά Πληροφοριακά Συστήματα από τη Σχολή Θετικών και Εφαρμοσμένων Επιστημών του Ανοικτού Πανεπιστημίου Κύπρου.

Μάιος 2018

Περίληψη

Στην παρούσα μεταπτυχιακή διατριβή περιγράφεται ένα σύστημα συστάσεων τουριστικών προορισμών. Μετά από διεξοδική μελέτη της υπάρχουσας διεθνούς βιβλιογραφίας σχεδιάστηκε ένα σύστημα συστάσεων με την τεχνική του συνεργατικού φιλτραρίσματος (CF – Collaborative Filtering). Ενώ στα περισσότερα συστήματα συστάσεων το συνεργατικό φιλτράρισμα βασίζεται στις αξιολογήσεις των χρηστών προς αντικείμενα, στο παρόν σύστημα το φιλτράρισμα βασίζεται στα προφίλ των χρηστών και στις προτιμήσεις τους, οι οποίες δίνονται ρητά από τους χρήστες. Επιπλέον στα περισσότερα ταξιδιωτικά συστήματα συστάσεων εφαρμόζονται άλλες προσεγγίσεις όπως φιλτράρισμα βάσει περιεχομένου (content-based filtering) και φιλτράρισμα βασισμένο στη γνώση (knowledge-based filtering). Στο πρώτο μέρος της μεταπτυχιακής διατριβής παρουσιάζεται η βιβλιογραφική ανασκόπηση όπου αναφέρονται ορισμοί, δημοφιλείς τουριστικοί ιστότοποι, τεχνικές φιλτραρίσματος που εφαρμόζονται στα συστήματα συστάσεων, τεχνικές αναπαράστασης και προσαρμογής προφίλ χρηστών. Στη συνέχεια αναφέρονται τουριστικά συστήματα συστάσεων. Ακολουθεί η περιγραφή του συστήματος, όπου αναλύονται οι βασικές λειτουργίες του. Στο δεύτερο μέρος, παρουσιάζεται η έρευνα που διεξήχθη σχετικά με τη γνώμη του κοινού για ένα τέτοιο ταξιδιωτικό σύστημα συστάσεων. Στο τελευταίο μέρος αναφέρονται τα πλεονεκτήματα και μειονεκτήματα αυτής της προσέγγισης, γίνονται προτάσεις για μελλοντικές επεκτάσεις και αναφέρονται τα συμπεράσματα της έρευνας.

Summary

The present postgraduate dissertation describes a travel destination recommendation system. Following an in-depth study of the existent international literature a recommendation system was planned implementing the Collaborative Filtering Technique (CF). While in the majority of travel recommendation systems Collaborative Filtering is based upon the user-to-item ratings, in the presented system filtering is based on users' profiles and their preferences, explicitly expressed by the users themselves. Furthermore, different approaches are applied in most travel recommendation systems such as content-based filtering and knowledge-based filtering. In the first part of this paper the literature review including definitions, popular tourist websites, filtering techniques, initial profile generation techniques and profile adaptation techniques is presented. Next, tourist recommendation systems are reported. The description of the system and an analysis of its basic functions is given. In the second part, the survey carried out on the public opinion on such a travel recommendation system is presented. Finally, the advantages and disadvantages of the approach are discussed, complemented by suggestions for future extensions and the outline of the research findings.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέπων καθηγητή της διατριβής μου, Δρ. Κλήμη Νταλιάνη, για την πολύτιμη καθοδήγηση, διαρκή στήριξη και την άμεση ανταπόκρισή του σε οποιαδήποτε απορία μου κατά την διάρκεια εκπόνησης της διατριβής.

Την ακαδημαϊκή υπεύθυνη του τμήματός μου και καθηγήτριά μου Δρ. Jahna Otterbacher και τους υπόλοιπους καθηγητές για τις γνώσεις τους και την πολύτιμη βοήθειά τους κατά τη διάρκεια των σπουδών μου.

Επιπλέον να αφιερώσω την διατριβή στην οικογένειά μου για την υπομονή και ψυχική υποστήριξη.

Περιεχόμενα

1	Εισαγωγή	1
1.1	Σημαντικότητα του θέματος	1
1.2	Συνεισφορά	2
1.3	Δομή της διατριβής	4
2	Βιβλιογραφική Ανασκόπηση - Ορισμοί	5
2.1	Τουρισμός - Ορισμοί – Παράγοντες που παρακινούν τους τουρίστες	5
2.2	Ταξιδιωτικοί Ιστότοποι	6
2.2.1	TripAdvisor	6
2.2.2	Travelocity	8
2.2.3	Booking.com	8
2.2.4	Χάρτες Google (Google Maps)	8
2.2.5	FOURSQUARE	9
2.2.6	Skyscanner	9
2.2.7	Airtickets	9
2.3	Συστήματα Συστάσεων – Ορισμοί – Βασικές προσεγγίσεις	10
2.3.1	Αλγόριθμοι φιλτραρίσματος	11
2.3.1.1	Συνεργατικό φιλτράρισμα - Collaborative Filtering based Recommendation (CF)	12
2.3.1.2	Δημογραφικό φιλτράρισμα - Demographic Recommendation	15
2.3.1.3	Φιλτράρισμα βάσει περιεχομένου - Content-based Recommendation	15
2.3.1.4	Φιλτράρισμα βασιζόμενο στη γνώση - Knowledge-based Recommendation	15
2.3.1.5	Υβριδικά συστήματα	16
2.3.1.6	Geographic RS – GRS	16
2.3.1.7	Κοινωνικές πληροφορίες – Social Information	17
2.3.1.8	Συστάσεις σε ομάδες χρηστών	18
2.4	Εξηγώντας τις Συστάσεις	19
2.5	Ηθικά ζητήματα - Ζητήματα ιδιωτικού απόρρητου	20
2.6	Προφίλ χρήστη στα συστήματα συστάσεων	20
2.6.1	Παραγωγή αρχικού προφίλ – Τεχνικές προσαρμογής προφίλ	22
3	Επιστημονική Βιβλιογραφία	24
4	Περιγραφή συστήματος	27
4.1	Εισαγωγή	27
4.2	Σύστημα	28
4.2.1	Είσοδος	29
4.2.2	Ερωτήσεις συστήματος	29
4.3	Προφίλ χρηστών	30
4.4	Αλγόριθμος Συστάσεων	32
4.4.1	Ομοιότητα χρηστών	32
4.4.2	Υπολογισμός βάρους ομοιότητας	33
4.4.2.1	Ομοιότητα δημογραφικών χαρακτηριστικών	33
4.4.2.2	Ομοιότητα παραγόντων που επηρεάζουν τον χρήστη για την επιλογή ταξιδιού	35
4.4.2.3	Ομοιότητα προτιμήσεων των χρηστών	36
4.4.2.4	Τελική Ομοιότητα χρηστών	38
4.5	Επιλογή γειτόνων - Neighborhood Selection	38
4.6	Προσαρμογή προφίλ	39
4.7	Χρήστες που δεν απαντούν στις ερωτήσεις προτίμησης	40
4.8	Δημιουργία προφίλ από αναζητήσεις των χρηστών	40
4.9	Συστάσεις	40
5.	Πόσο πρόθυμοι είναι οι χρήστες να χρησιμοποιήσουν το προτεινόμενο σύστημα συστάσεων	42
5.1	Περιγραφή έρευνας - Ερευνητικά ερωτήματα	42
5.2	Παρουσίαση και ερμηνεία των αποτελεσμάτων	45
5.2.1	Δημογραφικά	45

5.2.2	Συχνότητα ταξιδιών στο εσωτερικό/εξωτερικό	47
5.2.3	Προέλευση συστάσεων	49
5.2.4	Συμπλήρωση προφίλ, ταξιδιωτικών προτιμήσεων και επιλογών	51
5.2.4.1	Συμπλήρωση προφίλ	51
5.2.4.2	Συμπλήρωση δημογραφικών στοιχείων	54
5.2.4.3	Συμπλήρωση ερωτήσεων σχετικά με τους παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές	58
5.2.4.4	Συμπλήρωση ερωτήσεων σχετικά με τις ταξιδιωτικές προτιμήσεις	59
5.2.4.5	Εκμάθηση προφίλ μέσω των αναζητήσεων των χρηστών και το ιστορικό των ταξιδιών τους	60
5.2.4.6	Ερωτήσεις συστήματος από ποιες ομάδες χρηστών θα θέλανε να τους γίνονται οι συστάσεις	62
5.2.4.7	Ερωτήσεις συστήματος για εύρεση ομοιότητας βάση της σύγκρισης με άλλους χρήστες	63
6.	Επίλογος – Συμπεράσματα – Μελλοντικές επεκτάσεις	65
6.1	Πλεονεκτήματα - Μειονεκτήματα συστήματος.....	65
6.2	Συμπεράσματα έρευνας.....	66
6.3	Επεκτάσεις	67

Παραρτήματα

A	Ερωτήσεις Συστήματος	69
A.1	Δημογραφικά στοιχεία.....	69
A.2	Παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές του χρήστη	70
A.3	Προτιμήσεις συστάσεων	72
A.4	Προτιμήσεις και συχνότητα επιλογής	74
B	Πίνακες συστήματος	76
B.1	Προφίλ χρήστη.....	76
B.2	Προφίλ προορισμού	77
Γ	Διάγραμμα	78
Γ.1	Διάγραμμα κύριων λειτουργιών του συστήματος	78
Δ.	Ερωτηματολόγιο	79
Δ.1	Ερωτηματολόγιο έρευνας	79
	Βιβλιογραφία	83

Κεφάλαιο 1

Εισαγωγή

Ο τουρισμός είναι ένας από τους σημαντικούς και αναπτυσσόμενους τομείς της οικονομίας παγκοσμίως. Σύμφωνα με τον Παγκόσμιο Οργανισμό Τουρισμού οι συνολικές αφίξεις εισερχόμενων τουριστών σε όλες τις χώρες του κόσμου υπολογίζεται ότι το 2014 έφθασαν τα 1.133 εκατομμύρια. Κατά τους πρώτους τέσσερις μήνες του 2017 οι παγκόσμιοι προορισμοί δέχτηκαν 369 εκατομμύρια διεθνείς τουρίστες¹. Στη χώρα μας το έτος 2015 οι αφίξεις έφθασαν τα 23,6 εκατομμύρια² ενώ για το 2017 οι τουριστικές αφίξεις προβλέπονται να φτάσουν πάνω από 28,5 εκατομμύρια³.

Η πρόοδος στην Τεχνολογία της Πληροφορίας έχει επηρεάσει και τον τουριστικό κλάδο. Η τουριστική βιομηχανία είναι από τις πρώτες που υιοθέτησαν ΤΠΕ στον τομέα των επιχειρήσεων με την εισαγωγή του συστήματος ηλεκτρονικών κρατήσεων (CRS) στην δεκαετία του '70 (Buhalis & Law, 2008). Σήμερα το Διαδίκτυο προσφέρει ποικιλία τουριστικών υπηρεσιών και προϊόντων. Οι τουρίστες μπορούν να αναζητήσουν ταξιδιωτικές πληροφορίες, να πραγματοποιήσουν on-line κρατήσεις αεροπορικών εισιτηρίων, δωματίων, εισόδου σε σημεία ενδιαφέροντος. Επιπλέον η κοινωνική υπολογιστική και οι on-line κοινότητες αλλάζουν τον τρόπο που οι άνθρωποι μοιράζονται πληροφορίες και επικοινωνούν, επηρεάζουν σημαντικά την παγκόσμια οικονομία, τις κοινωνικές αλληλεπιδράσεις, και κάθε πτυχή της ζωής μας (Wang, Zeng, M. Carley, & Mao, 2007).

1.1 Σημαντικότητα του θέματος

Πριν λίγο καιρό οι ταξιδιωτικές επιλογές των ατόμων βασιζότανε σε συστάσεις φίλων και συγγενών, σε αναζήτηση πληροφοριών μέσα από τουριστικούς οδηγούς και

¹ http://cf.cdn.unwto.org/sites/all/files/pdf/unwto_barom17_03_june_excerpt_1.pdf

² http://www.gnto.gov.gr/sites/default/files/files_basic_pages/ELSTAT2015.pdf

³ <http://news.travelling.gr/2017/09/o-σετκε-για-tην-παγκόσμια-ημέρα-τουρισ/>

πληροφορίες από ταξιδιωτικά πρακτορεία. Με την εμφάνιση του Διαδικτύου ο χρήστης έχει πρόσβαση σε έναν τεράστιο όγκο πληροφοριών. Επιπλέον, σήμερα με την γενικευμένη χρήση τεχνολογιών Web 2.0 ο χρήστης εκτός από καταναλωτής πληροφοριών είναι και ο ίδιος παραγωγός περιεχομένου. Όλα αυτά έχουν οδηγήσει στο φαινόμενο της υπερπληροφόρησης (information overload) με αποτέλεσμα οι χρήστες να δυσκολεύονται να εντοπίσουν τις πληροφορίες που τους ενδιαφέρουν και είναι σημαντικές για αυτούς. Σύμφωνα με τον Hui et al. (2007), η ανακάλυψη των κατάλληλων τουριστικών υπηρεσιών είναι ένα από τα κύρια μειονεκτήματα όταν ταξιδεύουν οι άνθρωποι. Για αυτό το λόγο, τα συστήματα συστάσεων για τον τουρισμό έχουν προσελκύσει πολλή ερευνητική ενέργεια και ενδιαφέρον (Kabassi, 2010). Το χαρακτηριστικό των συστημάτων αυτών είναι ότι με βάση τα προφίλ των χρηστών και τη συμπεριφορά τους μπορούν να προσωποποιήσουν τις συστάσεις τους σε κάθε χρήστη και να του προτείνουν προϊόντα ή/και υπηρεσίες τα οποία είτε είναι παρόμοια με αυτά που έχουν ήδη αξιολογήσει, είτε είναι προϊόντα/υπηρεσίες που προτιμούν άλλοι χρήστες με παρόμοιο προφίλ.

1.2 Συνεισφορά

Τα συστήματα συστάσεων σήμερα ενσωματώνονται σε ένα ευρύ φάσμα εφαρμογών εμπορίου και περιεχομένου. Παρέχουν συστάσεις μουσικής (LastFm⁴, Pandora Radio⁵), βιβλίων (Reedgreek⁶), προϊόντων ηλεκτρονικού εμπορίου (Amazon.com, eBay), ταινιών (IMDb⁷, Netflix), βίντεο (YouTube), ειδήσεων, εφημερίδων, εργασιών. Επίσης, τα κοινωνικά μέσα δικτύωσης χρησιμοποιούν συστήματα συστάσεων για να προτείνουν νέους φίλους, ακόλουθους, ομάδες στους χρήστες τους (Facebook, Twitter, LinkedIn). Τα κινητά συστήματα συστάσεων (mobile recommender systems) εκμεταλλεύονται την αυξανόμενη χρήση έξυπνων τηλεφώνων για να προτείνουν διαδρομές (π.χ. googlemaps), και σημεία ενδιαφέροντος (π.χ. FourSquare).

Στον τομέα του τουρισμού η κατάσταση είναι πιο σύνθετη. Ο σχεδιασμός ενός ταξιδιού είναι μια περίπλοκη διαδικασία λήψης αποφάσεων με πολλούς και διαφορετικούς παράγοντες να επηρεάζουν την επιλογή προορισμού. Προσωπικά χαρακτηριστικά (όπως ηλικία, εισόδημα, εκπαίδευση, προσωπικότητα), ταξιδιωτικά χαρακτηριστικά

⁴ <https://en.wikipedia.org/wiki/Last.fm>

⁵ https://en.wikipedia.org/wiki/Pandora_Radio

⁶ https://en.wikipedia.org/wiki/Recommender_system

⁷ <https://en.wikipedia.org/wiki/IMDb>

(όπως σκοπός, διάρκεια ταξιδιού, απόσταση) επηρεάζουν τις αποφάσεις του ταξιδιώτη (Ricci, 2002). Υπάρχουν πολλά κίνητρα που ωθούν κάποιον να ταξιδέψει ικανοποιώντας ένα πλήθος αναγκών-λόγων που για τον κάθε χρήστη είναι διαφορετικοί.

Το τουριστικό προϊόν είναι ένα ιδιαίτερο προϊόν το οποίο είναι αδύνατο να αξιολογηθεί πριν την κατανάλωσή του. Η επιλογή του γίνεται μετά από μια διαδικασία λήψης απόφασης η οποία βασίζεται σε πληροφορίες από διάφορες πηγές, όπως από στόμα σε στόμα διάδοση, διαδίκτυο, τηλεόραση, φυλλάδια. Η πληροφορία λοιπόν αποτελεί σημαντικό στοιχείο για τον τουρισμό και η παροχή χρήσιμης και κατάλληλης πληροφορίας την κατάλληλη στιγμή και στον κατάλληλο αγοραστή αποτελεί στοιχείο που μπορεί να διαφοροποιήσει και να αυξήσει την ανταγωνιστικότητα των επιχειρήσεων που δραστηριοποιούνται στον τουριστικό τομέα (Σιγάλα, 2008).

Προς το παρόν δεν υπάρχει ένα σύστημα το οποίο να υποστηρίζει τον χρήστη για την οργάνωση ενός ολοκληρωμένου ταξιδιού. Υπάρχουν συστήματα τα οποία συστήνουν στον χρήστη προορισμούς, άλλα συστήματα κάνουν προτάσεις για σημεία ενδιαφέροντος, άλλα για καταλύματα, δραστηριότητες ή αξιοθέατα. Κάποια από αυτά δημιουργούν και διατηρούν προφίλ για τον κάθε χρήστη τους και με βάση αυτό το προφίλ και τις προτιμήσεις του πραγματοποιούν συστάσεις, άλλα χρησιμοποιούν τεχνολογία εντοπισμού θέσης και προτείνουν σημεία ενδιαφέροντος με βάση την απόσταση από την τρέχουσα θέση του χρήστη και άλλα κάνουν προτάσεις με βάση τις αξιολογήσεις των υπόλοιπων χρηστών του συστήματος.

Στην παρούσα εργασία, προτείνεται ένα διαφορετικό σύστημα συστάσεων, το οποίο διερευνά τους πιο ισχυρούς παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές ενός χρήστη και τη σημαντικότητα που δίνει ο χρήστης σε κάθε έναν από αυτούς τους παράγοντες. Επιπρόσθετα ο χρήστης μπορεί να δηλώσει στο σύστημα από ποιους χρήστες προτιμά να του γίνονται συστάσεις. Λαμβάνοντας υπόψη τα δημογραφικά χαρακτηριστικά των χρηστών, τις προτιμήσεις τους και τις επιλογές τους αλλά και τη σημαντικότητα που προσδίδει ο χρήστης σε κάθε μια από αυτές τις παραμέτρους γίνεται ο υπολογισμός ομοιότητας των χρηστών με προσέγγιση συνεργατικού φιλτραρίσματος. Οι προτιμήσεις δηλώνονται ρητά από τον χρήστη, απαντώντας σε ερωτήσεις του συστήματος και επομένως είναι πιο αξιόπιστες από αυτές που εξάγονται

σιωπηρά από τη συμπεριφορά του χρήστη ή από τις αξιολογήσεις όπως γίνεται σε άλλα συστήματα συστάσεων.

1.3 Δομή της διατριβής

Η παρούσα μεταπτυχιακή διατριβή αποτελείται από έξι κεφάλαια. Στο παρόν κεφάλαιο αναφέρεται η εισαγωγή, η σημαντικότητα του θέματος που ερευνάται, η συνεισφορά της εργασίας μας και παρουσιάζεται περιληπτικά το περιεχόμενο του κάθε κεφαλαίου.

Στο δεύτερο κεφάλαιο, έπειτα από βιβλιογραφική ανασκόπηση, αναφέρεται ο ορισμός του τουρισμού και παρουσιάζονται οι παράγοντες που παρακινούν τους τουρίστες να επισκεφτούν έναν προορισμό. Στη συνέχεια παρουσιάζονται δημοφιλείς τουριστικοί ιστότοποι. Αναφέρονται αναλυτικά οι τεχνικές φιλτραρίσματος που χρησιμοποιούν τα συστήματα συστάσεων (όπως collaborative filtering, content based). Έπειτα αναφέρονται τεχνικές αναπαράστασης, παραγωγής και προσαρμογής προφίλ χρήστη. Και τέλος, τα είδη παροχής εξηγήσεων που δικαιολογούν τις συστάσεις καθώς και ηθικά ζητήματα που προκύπτουν στα συστήματα συστάσεων.

Στο τρίτο κεφάλαιο αναφέρονται διάφορες τεχνολογίες τουριστικών συστημάτων συστάσεων που έχουν προταθεί από την επιστημονική κοινότητα.

Στο τέταρτο κεφάλαιο παρουσιάζεται το σύστημα συστάσεων που προτείνεται στην παρούσα εργασία με προσέγγιση συνεργατικού φιλτραρίσματος με βάση τα προφίλ των χρηστών και τις ταξιδιωτικές τους προτιμήσεις.

Στο πέμπτο κεφάλαιο παρουσιάζεται η έρευνα και τα αποτελέσματα σχετικά με τη γνώμη του κοινού για το σύστημα συστάσεων που περιγράφηκε στο κεφάλαιο 4.

Τέλος, στο έκτο κεφάλαιο αναφέρονται τα συμπεράσματα και μελλοντικές επεκτάσεις.

Κεφάλαιο 2

Βιβλιογραφική Ανασκόπηση - Ορισμοί

«Το ταξίδι αναφέρεται στη δραστηριότητα των ταξιδιωτών. Ένας ταξιδιώτης είναι κάποιος που κινείται μεταξύ διαφορετικών γεωγραφικών τοποθεσιών για οποιοδήποτε σκοπό και διάρκεια (United Nations, 2010⁸).»

2.1 Τουρισμός - Ορισμοί - Παράγοντες που παρακινούν τους τουρίστες

Στην Ελληνική και Διεθνή βιβλιογραφία υπάρχει πληθώρα ορισμών της έννοιας του τουρισμού. Σύμφωνα με τον Παγκόσμιο Οργανισμό Τουρισμού⁹ «Ο τουρισμός περιλαμβάνει τις δραστηριότητες των ανθρώπων που ταξιδεύουν και διαμένουν σε προορισμούς και περιοχές άλλες εκτός από αυτές που αποτελούν το συνηθισμένο περιβάλλον τους και για χρονικό διάστημα που δεν ξεπερνά τον έναν χρόνο, με σκοπό την αναψυχή, την ικανοποίηση των επαγγελματικών τους αναγκών κ.α.».

Ο όρος e-Tourism σύμφωνα με τον Buhalis (2003) αντικατοπτρίζει την ψηφιοποίηση όλων των διαδικασιών και αλυσίδων αξίας του τουρισμού, των ταξιδιών, της φιλοξενίας και της τροφοδοσίας. «Η έννοια του ηλεκτρονικού τουρισμού περιλαμβάνει όλες τις επιχειρηματικές λειτουργίες (π.χ. ηλεκτρονικό εμπόριο, ηλεκτρονικό μάρκετινγκ, ηλεκτρονική λογιστική, eHRM, ηλεκτρονικές δημόσιες συμβάσεις, eR & D, ηλεκτρονική παραγωγή) καθώς και ηλεκτρονική στρατηγική, το σχεδιασμό και την ηλεκτρονική διαχείριση για όλους τους τομείς της τουριστικής βιομηχανίας, συμπεριλαμβανομένου του τουρισμού, των μετακινήσεων, των μεταφορών, του ελεύθερου χρόνου, της φιλοξενίας, των διευθυντών, των μεσαζόντων και των οργανώσεων του δημόσιου τομέα» (Buhalis, 2003).

⁸ Recommendations on Tourism Statistics, Statistical Papers, Series M, No.83 United Nations Publication Sales No. E.94.XVII.

⁹ <http://www2.unwto.org/>

Ο τουρισμός είναι ένα πολυδιάστατο οικονομικοκοινωνικό φαινόμενο. Σύμφωνα με τον Ηγουμενάκη κ.α. (1999) τέσσερις είναι οι παράγοντες που παρακινούν τους τουρίστες να επισκεφθούν έναν προορισμό: (α) τα θέλγητρα, (β) οι εγκαταστάσεις και οι υπηρεσίες, (γ) η ευκολία προσέγγισης και (δ) το κόστος επίσκεψης ενός τουριστικού προορισμού. Όσον αφορά τα θέλγητρα ενός προορισμού, αυτά διακρίνονται σε τέσσερις κατηγορίες: (α) φυσικά, δηλ. το φυσικό περιβάλλον όπως βουνά, λίμνες σπήλαια, ακρογιαλιές, κλιματολογικές συνθήκες κ.α., (β) κατασκευασμένα, όπως αρχαιολογικοί χώροι, κτίρια αρχιτεκτονικής σημασίας, αθλητικά κέντρα, πάρκα, ενυδρεία κ.α. (γ) πολιτιστικά, όπως μουσεία, πινακοθήκες, θέατρα, χώροι τέχνης, θρησκείας, ιστορίας κ.α. και (δ) κοινωνικά, όπως τα ήθη και έθιμα του ντόπιου πληθυσμού, οι παραδόσεις του, η γλώσσα του, κοινωνικές αντιθέσεις κ.α. (Ηγουμενάκης, Κραβαρίτης, & Λύτρας, 1999).

Μια ταξινόμηση με βάση τον κύριο σκοπό επίσκεψης ενός προορισμού είναι η εξής¹⁰:

- Ελεύθερος χρόνος, αναψυχή και διακοπές
- Επίσκεψη φίλων και συγγενών
- Επαγγελματικοί λόγοι (συνέδρια, εκθέσεις, αγορές – πωλήσεις κ.α.)
- Λόγοι υγείας (ιαματικές πηγές, γυμναστήρια, θαλασσοθεραπεία, θέρετρα υγείας, θεραπείες κ.α.)
- Θρησκεία / Προσκυνήματα (συμμετοχή σε θρησκευτικές εκδηλώσεις, προσκυνήματα)
- Άλλοι λόγοι

2.2 Ταξιδιωτικοί Ιστότοποι

Τα τελευταία χρόνια υπάρχουν αρκετοί ιστότοποι στον τομέα του τουρισμού. Παρακάτω παρουσιάζονται οι πιο δημοφιλείς.

2.2.1 TripAdvisor

Σύμφωνα με την ιστοσελίδα, το TripAdvisor είναι ο μεγαλύτερος ταξιδιωτικός ιστότοπος με 500 εκ. κριτικές και γνώμες, διαθέτει τη μεγαλύτερη συλλογή ταξιδιωτικών καταχωρήσεων παγκοσμίως (ξενοδοχεία, αεροπορικές εταιρείες,

¹⁰ https://unstats.un.org/unsd/publication/seriesm/seriesm_83e.pdf

εστιατόρια, αξιοθέατα). Το πρώτο τρίμηνο του 2017 σύμφωνα με τα αρχεία καταγραφής της ιστοσελίδας TripAdvisor είχε 390 εκ. μοναδικούς επισκέπτες το μήνα¹¹.

Τα μέλη αυτής της κοινότητας αξιολογούν τουριστικά προϊόντα και υπηρεσίες όπως ξενοδοχεία, εστιατόρια, δραστηριότητες, αεροπορικές εταιρείες. Βαθμολογούν σε μια κλίμακα από 1 έως 5, δίνουν έναν τίτλο στην κριτική τους και στη συνέχεια γράφουν ένα κείμενο με ελάχιστο αριθμό χαρακτήρων τους 100. Έπειτα επιλέγουν το είδος της επίσκεψης (ζευγάρια, οικογένειες...), ημερομηνία επίσκεψης και προαιρετικά μπορούν να δώσουν και άλλες πληροφορίες, να βαθμολογήσουν και άλλες υπηρεσίες (εξυπηρέτηση, καθαριότητα, αξία).

Επιπλέον μπορούν να ψηφίσουν (θετικά) κριτική άλλου χρήστη ή να επισημάνουν κάποιο πρόβλημα σε μια κριτική επιλέγοντας από μια λίστα που τους προτείνει το σύστημα ή να αναφέρουν κάτι άλλο αν δεν τους ικανοποιούν οι διαθέσιμες επιλογές.

Ως αναγνώριση της συνεισφοράς των χρηστών η εταιρεία διαθέτει το πρόγραμμα συμβολής TripCollective¹². Κάθε φορά που ο χρήστης προσθέτει κάτι στην κοινότητα του TripAdvisor (κριτική, φωτογραφία, βίντεο, δημοσίευση στο forum, βαθμολογία) συλλέγει πόντους. Όσο περισσότερους πόντους συγκεντρώνει τόσο ανεβαίνει επίπεδο και κερδίζει ταξιδιωτικά σήματα. Το επίπεδο του χρήστη, οι συνεισφορές του καθώς και ο αριθμός των χρήσιμων ψήφων που έχει συγκεντρώσει εμφανίζονται στο ταξιδιωτικό προφίλ του.

Όλα τα παραπάνω προϋποθέτουν την εγγραφή του χρήστη στον ιστότοπο. Υπάρχει η δυνατότητα να συνδεθεί ο χρήστης μέσω του λογαριασμού που διαθέτει στο Facebook. Τους δίνεται η δυνατότητα να επεξεργαστούν το προφίλ τους, όπου τους ζητείται να συμπληρώσουν την ηλικία τους, το φύλλο και την τοποθεσία τους. Επιπλέον μπορούν να επιλέξουν 3 και άνω ετικέτες από τις 19, σχετικές με τις προτιμήσεις τους όπως «Λάτρης των οικογενειακών διακοπών», «Φυσιολάτρης», «Χορτοφάγος» κλπ.

¹¹ <https://tripadvisor.mediaroom.com/GR-about-us>

¹² <https://www.tripadvisor.com.gr/TripCollective>

Επιπλέον το TripAdvisor παρέχει στους χρήστες του εξατομικευμένες προτάσεις για ξενοδοχεία βάσει των προσωπικών προτιμήσεών τους και της αναζήτησης ταξιδιών στον ιστότοπο μέσω της δυνατότητας «Μόνο για εσάς». Χρησιμοποιώντας προηγούμενες αναζητήσεις και συνεισφορές βοηθάει τους χρήστες να επιλέξουν ένα ξενοδοχείο κοντά στις ανάγκες τους και τα ενδιαφέροντά τους. Όσο περισσότερο χρησιμοποιεί και αλληλεπιδρά ο χρήστης με τον ιστότοπο, τόσο πιο πολύ το σύστημα αυτό προσαρμόζεται στις προσωπικές του ταξιδιωτικές προτιμήσεις.¹³

2.2.2 Travelocity

Είναι ένα online γραφείο ταξιδιών. Παρέχει την δυνατότητα για κρατήσεις αεροπορικών πτήσεων, ξενοδοχείων, ενοικιάσεων αυτοκινήτων, πακέτα διακοπών, δραστηριότητες. Παρέχει εξατομικευμένες συστάσεις προορισμών στους χρήστες του.

2.2.3 Booking.com

Στην ιστοσελίδα της εταιρείας αναφέρει ότι «*Ο ιστοχώρος και οι εφαρμογές για κινητές συσκευές της Booking.com διατίθενται σε πάνω από 40 γλώσσες, προσφέρουν 1.479.147 καταλύματα, και καλύπτουν πάνω από 121.448 προορισμούς σε 228 χώρες και επικράτειες σε όλο τον κόσμο. Καθημερινά γίνονται κρατήσεις για πάνω από 1.500.000 διανυκτερεύσεις στην πλατφόρμα μας.*¹⁴» Παρέχει υπηρεσίες on line κρατήσεων για διάφορα προϊόντα και υπηρεσίες όπως για καταλύματα, πτήσεις, ενοικίαση αυτοκινήτου, ταξί αεροδρομίου. Η ιστοσελίδα συλλέγει τις απόψεις των χρηστών καθώς και τις αξιολογήσεις που κάνουν σε σχόλια άλλων χρηστών, πληροφορίες σχετικές με προορισμούς που μοιράζονται οι χρήστες ώστε να βοηθήσουν άλλους ταξιδιώτες σαν αυτούς τους χρήστες να βρουν τον ιδανικό προορισμό και το κατάλληλο κατάλυμα για αυτούς. Επίσης ο χρήστης μπορεί να συμπεριλάβει στο προφίλ του τις ταξιδιωτικές προτιμήσεις του ώστε ο ιστότοπος να προσαρμόζει την αναζήτησή του με αυτές.

2.2.4 Χάρτες Google (Google Maps)

Η υπηρεσία Χάρτες Google είναι μια υπηρεσία χαρτογράφησης στο Διαδίκτυο¹⁵.

¹³ <https://www.tripadvisor.com/hc/el/articles/203825386-Τι-είναι-η-δυνατότητα-Μόνο-για-εσάς->

¹⁴ <https://www.booking.com/content/about.el.html?aid=376387;label=bookings-name-4QYQPZsQi8OR2tqLpbgQKwS72583353961%3Apl%3Aata%3Ap1%3Ap2%3Aac%3Aap1t1%3Aneg%3Afi%3Atiaud-285284110726%3Akwd-65526620%3Alp20363%3Ali%3Adec%3Adm;sid=c2ff3cf98f30dde807cd120e8059586c>

¹⁵ https://el.wikipedia.org/wiki/Χάρτες_Google

Προσφέρει πολλές υπηρεσίες όπως χάρτες δρόμων, σχεδιαστή διαδρομών και μεταξύ αυτών οι χρήστες μπορούν να γράφουν αξιολογήσεις – κριτικές για τα μέρη που έχουν επισκεφθεί, να προσθέτουν φωτογραφίες καθώς και να επισημαίνουν κριτικές άλλων ως χρήσιμες ή ακατάλληλες.

2.2.5 FOURSQUARE

Σύμφωνα με τον ιστότοπο της FOURSQUARE 50 εκ. άνθρωποι χρησιμοποιούν την εφαρμογή Foursquare City Guide και Foursquare Swarm. Περισσότεροι από 105 εκ. χώροι είναι καταχωρημένοι στην πλατφόρμα.

Παρέχει και εφαρμογή για κινητά. Το Foursquare City Guide χρησιμοποιεί τον εντοπισμό θέσης - location-based social networking (LBSN) - για να βοηθήσει τους χρήστες του να ανακαλύψουν εύκολα, σημεία ενδιαφέροντος όπως μέρη για καφέ, φαγητό, νυχτερινή διασκέδαση, ψώνια. Όπως και στα προηγούμενα οι χρήστες αφήνουν τα σχόλιά τους, αξιολογούν τα μέρη που έχουν επισκεφτεί, ανεβάζουν φωτογραφίες, ψηφίζουν θετικά ή αρνητικά κριτικές άλλων μελών. Παρέχει εξατομικευμένες συστάσεις στους χρήστες του με βάση το ιστορικό αναζήτησης, αγορών, επισκέψεων και κρατήσεων.

2.2.6 Skyscanner

Ταξιδιωτικός ιστότοπος αναζήτησης και μηχανή σύγκρισης τιμών αεροπορικών εισιτηρίων, ενοικίασης αυτοκινήτων και ξενοδοχείων. Κάθε μήνα χρησιμοποιούν τον ιστότοπο 60 εκ. άνθρωποι. Οι υπηρεσίες τους είναι διαθέσιμες σε πάνω από 30 γλώσσες και 79 νομίσματα. Η εφαρμογή τους για κινητά την έχουν κατεβάσει πάνω από 60 εκ. φορές.¹⁶ Προτείνει προορισμούς στους χρήστες με βάση τις αναζητήσεις τους στον ιστότοπο.

2.2.7 Airtickets

Το μεγαλύτερο εξειδικευμένο ελληνικό site κράτησης αεροπορικών εισιτηρίων και άλλων ταξιδιωτικών υπηρεσιών, με περισσότερες από 70.000 επισκέψεις ημερησίως.¹⁷ Δεν παρέχει συστάσεις στους χρήστες του.

¹⁶ <https://gr.skyscanner.com/aboutskyscanner.aspx>

¹⁷ <http://www.airtickets.gr/about-us>

2.3 Συστήματα Συστάσεων - Ορισμοί - Βασικές προσεγγίσεις

“Τα συστήματα συστάσεων είναι η τεχνική απάντηση στο γεγονός ότι βασιζόμαστε συχνά στην εμπειρία και τις συστάσεις άλλων ανθρώπων όταν αντιμετωπίζουμε μια νέα πείρα, όπου δεν έχουμε ευρεία γνώση όλων των γεγονότων ή όπου η γνώση αυτή θα υπερβεί το ποσό των πληροφοριών που οι άνθρωποι μπορούν να ασχοληθούν με τη γνώση” (Berka & Plöbzig, 2004).

Σύμφωνα με το ACM Recommender Systems conference (RecSys)¹⁸ «*Η σύσταση είναι μια ιδιαίτερη μορφή φιλτραρίσματος πληροφοριών, που εκμεταλλεύεται προηγούμενες συμπεριφορές και ομοιότητες χρηστών για να δημιουργήσει μια λίστα στοιχείων πληροφοριών που προσαρμόζονται προσωπικά στις προτιμήσεις ενός τελικού χρήστη*».

Τα συστήματα συστάσεων είναι ένα χρήσιμο εργαλείο για να αντιμετωπιστεί εν μέρει το φαινόμενο της υπερφόρτωσης των πληροφοριών. Εξελίσσονται παράλληλα με την εξέλιξη του ιστού. Η **πρώτη** γενιά συστημάτων συστάσεων χρησιμοποίησε παραδοσιακούς ιστοτόπους για τη συλλογή πληροφοριών από τις ακόλουθες τρεις πηγές:

(α) δεδομένα βάσει περιεχομένου (content-based data) από προϊόντα που αγοράστηκαν ή χρησιμοποιήθηκαν,

(β) δημογραφικά δεδομένα (demographic data) που συλλέγονται σε αρχεία χρηστών όπως ηλικία, εθνικότητα, φύλο, και

(γ) δεδομένα που βασίζονται στη μνήμη (memory-based data) που συλλέγονται από τις προτιμήσεις σε αντικείμενα των χρηστών.

Η **δεύτερη** γενιά συστημάτων συστάσεων χρησιμοποιεί εκτενώς το web 2.0 συλλέγοντας κοινωνικές πληροφορίες (social information) (π.χ. φίλους, οπαδούς, ακολουθούμενους χρήστες, αξιόπιστους χρήστες, μη αξιόπιστους χρήστες, ετικέτες, σχόλια).

Η **τρίτη** γενιά συστημάτων συστάσεων θα χρησιμοποιεί το Web 3.0 μέσω των πληροφοριών που παρέχονται από τις ενσωματωμένες συσκευές στο Internet. Η χρήση

¹⁸ <https://recsys.acm.org/recsys17/>

των πληροφοριών θέσης (location information) που έχουν ήδη ενσωματωθεί σε πολλά συστήματα συστάσεων θα ακολουθούνται από δεδομένα από συσκευές και αισθητήρες, τα οποία θα χρησιμοποιηθούν ευρέως (π.χ. σήματα υγείας σε πραγματικό χρόνο, RFID, διατροφικές συνήθειες, online τοπικές παραμέτρους καιρού όπως θερμοκρασία και πίεση).

Σύμφωνα με τους Bobadilla et al. (2013), η διαδικασία για τη δημιουργία ενός συστήματος συστάσεων βασίζεται στον συνδυασμό των παρακάτω εκτιμήσεων:

- Τύπος δεδομένων που είναι διαθέσιμος στη βάση δεδομένων (π.χ. βαθμολογίες, πληροφορίες εγγραφής χρηστών, χαρακτηριστικά και περιεχόμενο για στοιχεία που μπορούν να ταξινομηθούν, κοινωνικές σχέσεις μεταξύ χρηστών και πληροφορίες που αφορούν την τοποθεσία).
- Αλγόριθμος φιλτραρίσματος π.χ. δημογραφικός (demographic), βασισμένος στο περιεχόμενο (content-based), συνεργατικός (collaborative), κοινωνικός (social-based) και υβριδικός (hybrid).
- Η μέθοδος, π.χ. μέθοδος που βασίζεται στη μνήμη (Memory-based methods) ή μέθοδος που βασίζεται σε μοντέλα (Model-based methods).
- Χρησιμοποιούμενες τεχνικές: πιθανολογικές προσεγγίσεις (probabilistic approaches), δίκτυα Bayesian (Bayesian networks), αλγόριθμος πλησιέστερων γειτόνων (nearest neighbors algorithm), βιο-εμπνευσμένους αλγορίθμους όπως τα νευρωνικά δίκτυα και οι γενετικοί αλγόριθμοι, ασαφή μοντέλα κ.λπ.
- Το επίπεδο έλλειψης στοιχείων της βάσης δεδομένων και η δυνατότητα διεύρυνσης.
- Απόδοση του συστήματος (κατανάλωση χρόνου και μνήμης).
- Ο επιθυμητός στόχος (π.χ. προβλέψεις και κορυφαίες συστάσεις).
- Η επιδιωκόμενη ποιότητα των αποτελεσμάτων όπως για παράδειγμα η καινοτομία (novelty), η κάλυψη (coverage) και η ακρίβεια (precision).

2.3.1 Αλγόριθμοι φιλτραρίσματος

Οι εσωτερικές λειτουργίες ενός συστήματος συστάσεων χαρακτηρίζονται από τον αλγόριθμο φιλτραρίσματος που χρησιμοποιεί. Οι πιο διαδεδομένοι αλγόριθμοι παρουσιάζονται παρακάτω:

2.3.1.1 Συνεργατικό φιλτράρισμα - Collaborative filtering based Recommendation (CF)

Οι αλγόριθμοι αυτής της κατηγορίας βρίσκουν χρήστες με παρόμοια ενδιαφέροντα και παράγουν συστάσεις στον χρήστη. Αθροίζονται οι αξιολογήσεις των χρηστών και δημιουργούνται συστάσεις βάσει ομοιότητας χρηστών. Ως είσοδο παίρνει τις αξιολογήσεις του χρήστη προς το αντικείμενο και εντοπίζει παρόμοιους χρήστες από τις αξιολογήσεις τους. Το CF αυτοματοποιεί τη διαδικασία της πληροφόρησης «από στόμα σε στόμα» ("word of mouth") (Ricci, Rokach, Shapira, & Kantor, 2011). Όπως στην πραγματική ζωή ζητάμε συμβουλές από άτομα της εμπιστοσύνης μας (φίλους, συναδέλφους, ειδικούς) έτσι και η προσέγγιση αυτή σταθμίζει όλους τους χρήστες σε σχέση με την ομοιότητα με τον ενεργό χρήστη. Ως ομοιότητα νοείται να έχουν τα ίδια ενδιαφέροντα, παρόμοιες προτιμήσεις, να βαθμολογούν και να αξιολογούν με παρόμοιο τρόπο. Οι αλγόριθμοι αυτής της προσέγγισης στηρίζονται στην αρχή της ομοφιλίας (Homophily). Άνθρωποι με παρόμοια χαρακτηριστικά έχουν την τάση να γίνονται φίλοι (Easley & Kleinberg, 2010) και να ενδιαφέρονται για παρόμοια αντικείμενα. Οι αλγόριθμοι επιλέγουν ένα υποσύνολο από παρόμοιους χρήστες (γείτονες) και βασιζόμενοι στις αξιολογήσεις αυτών προβλέπουν την αξιολόγηση του ενεργού χρήστη για συγκεκριμένα αντικείμενα. Τέλος συστήνονται στον ενεργό χρήστη τα αντικείμενα με τη μεγαλύτερη πρόβλεψη.

Στις τεχνικές αυτής της προσέγγισης υπάρχει μια λίστα χρηστών $u \{u_1, u_2 \dots u_n\}$, μια λίστα με αντικείμενα $i \{i_1, i_2 \dots i_n\}$ και μια λίστα με αντικείμενα lu_i τα οποία έχει αξιολογήσει ο χρήστης. Οι αξιολογήσεις του χρήστη μπορούν να είναι ρητές, π.χ. ο χρήστης αξιολογεί ένα στοιχείο σε μια κλίμακα, είτε σιωπηρές όπως αναζητήσεις, ή αγορές (Xiaojuan & Taghi, 2009).

Δύο είναι οι γενικές κατηγορίες μεθόδων στις συνεργατικές μεθόδους φιλτραρίσματος:

- η μέθοδος που βασίζεται στην μνήμη (memory-based), χρησιμοποιεί τα δεδομένα αξιολόγησης χρήστη για τον υπολογισμό της ομοιότητας ή του βάρους μεταξύ χρηστών ή αντικειμένων και πραγματοποιούν προβλέψεις ή συστάσεις σύμφωνα με αυτές τις τιμές ομοιότητας (Xiaojuan & Taghi, 2009). Τέτοια συστήματα βρίσκουν εφαρμογή σε εμπορικά site όπως η Amazon.com.

Αυτό γίνεται με δύο τρόπους:

- πρόβλεψη βασιζόμενη στον χρήστη (user-based). Ένας χρήστης u θα ενδιαφερθεί πιθανότατα για ένα στοιχείο αν και οι γείτονές του ενδιαφέρονται για αυτό το στοιχείο. Ως γείτονες θεωρούνται χρήστες που αξιολόγησαν τα ίδια αντικείμενα με παρόμοιο τρόπο με τον χρήστη u .
- πρόβλεψη βασιζόμενη στα αντικείμενα (item-based). Σε αυτή την προσέγγιση η βαθμολογία ενός χρήστη για ένα στοιχείο i υπολογίζεται βάσει των αξιολογήσεων του χρήστη αυτού για στοιχεία παρόμοια με το i . Δύο στοιχεία είναι παρόμοια αν αρκετοί χρήστες του συστήματος έχουν αξιολογήσει αυτά τα στοιχεία με παρόμοιο τρόπο.
- η μέθοδος μοντέλου (model-based) χρησιμοποιεί τις αξιολογήσεις για να δημιουργήσει ένα μοντέλο πρόβλεψης. Η γενική ιδέα είναι να μοντελοποιηθούν οι αλληλεπιδράσεις χρήστη-στοιχείου. Το μοντέλο μπορεί να είναι ένας αλγόριθμος εξόρυξης δεδομένων ή μηχανής μάθησης (Xiaoouan & Taghi, 2009). Το μοντέλο αυτό εκπαιδεύεται χρησιμοποιώντας τα διαθέσιμα δεδομένα και αργότερα χρησιμοποιείται για να προβλέψει τις αξιολογήσεις των χρηστών για νέα στοιχεία. (Ricci, Rokach, Shapira, & Kantor, 2011). Τέτοια μοντέλα είναι Bayesian belief nets (BNs) CF models, clustering CF models και latent semantic CF models.

Ο πιο ευρέως χρησιμοποιούμενος αλγόριθμος για συνεργατικό φιλτράρισμα είναι ο αλγόριθμος συστάσεων πλησιέστερων γειτόνων - k nearest neighbors recommendation algorithm (kNN) (Bobadilla et al., 2013). Εννοιολογικά είναι πολύ κοντά με την προσέγγιση του συνεργατικού φιλτραρίσματος, η εύρεση παρόμοιων χρηστών είναι ισοδύναμη με την εύρεση γειτόνων για έναν συγκεκριμένο χρήστη (Ricci, Rokach, Shapira, & Kantor, 2011). Οι αλγόριθμοι User-Based Top-N Recommendation προσδιορίζουν τους k παρόμοιους χρήστες (γείτονες) με τον ενεργό χρήστη u (Xiaoouan & Taghi, 2009). Η επιλογή του N θα πρέπει να γίνεται προσεκτικά διότι εάν το N είναι πολύ μεγάλο, θα χρειαστεί υπερβολική ποσότητα μνήμης για την αποθήκευση των λιστών γειτονιάς και η πρόβλεψη θα είναι αργή (Ricci, Rokach, Shapira, & Kantor, 2011, p. 130). Ενώ από την άλλη η επιλογή μιας υπερβολικά μικρής τιμής για το N μπορεί να οδηγήσει σε μη ακριβείς συστάσεις.

Τα πλεονεκτήματά του είναι η απλότητά του και οι καλής ποιότητας προβλέψεις και συστάσεις, ενώ τα μειονεκτήματα είναι τα υψηλά επίπεδα έλλειψης δεδομένων (sparsity) (Bobadilla et al., 2013). Στις τεχνικές που βασίζονται στη μνήμη (memory-based) ο υπολογισμός ομοιότητας βασίζεται σε αξιολογήσεις κοινών αντικειμένων και όταν τα δεδομένα που έχουμε στη διάθεσή μας είναι αραιά τότε τα κοινά αντικείμενα είναι λίγα (Xiaoouyan & Taghi, 2009). Στην περίπτωση λοιπόν που δεν υπάρχουν επαρκείς αξιολογήσεις προτιμήσεων από τους χρήστες, τότε τα δεδομένα αξιολόγησης είναι ελλιπή και αυτό μπορεί να οδηγήσει σε λανθασμένες συσχετίσεις και κατ' επέκταση σε μη αξιόπιστες συστάσεις (Bobadilla et al., 2013). Ελλιπή δεδομένα συναντάμε στην περίπτωση της ψυχρής εκκίνησης (cold start problem), και όπως αναφέρθηκε προηγουμένως, σε αντικείμενα με χαμηλό αριθμό αξιολογήσεων (Bobadilla et al., 2013) όταν εισέρχεται ένα νέο στοιχείο ή νέος χρήστης. Στην πρώτη περίπτωση ένα νέο στοιχείο όταν εισέρχεται δεν έχει αρχικές αξιολογήσεις και έτσι δεν συστήνεται στους χρήστες. Στη συνέχεια ένα νέο στοιχείο που δεν προτείνεται μένει απαρατήρητο από ένα μεγάλο μέρος χρηστών καθώς δεν το γνωρίζουν και κατά συνέπεια δεν το βαθμολογούν. Αυτό οδηγεί σε ένα φαύλο κύκλο όπου μια ομάδα στοιχείων παραλείπεται από τη διαδικασία αξιολόγησης/συστάσεων (Bobadilla et al., 2013). Παρόμοια και στην δεύτερη περίπτωση, ένας νέος χρήστης αφού δεν έχει αξιολογήσει κάποιο στοιχείο δεν μπορεί να λάβει και κάποια σύσταση. Επιπλέον ο αριθμός των αξιολογήσεων των χρηστών μπορεί να μην επαρκεί για να γίνουν αξιόπιστες συστάσεις και οι χρήστες νιώθοντας ότι η υπηρεσία δεν ανταποκρίνεται στις προτιμήσεις τους σταματάνε να τη χρησιμοποιούν. Επομένως, μία από τις μεγαλύτερες προκλήσεις για τον σχεδιασμό ενός τέτοιου συστήματος φιλτραρίσματος είναι ο τρόπος παροχής συστάσεων με τα αραιά δεδομένα προφίλ χρήστη (Havaragi & Prathibha, 2016).

Άλλο μειονέκτημα είναι η χαμηλή επεκτασιμότητα (scalability). Καθώς οι βάσεις δεδομένων αυξάνονται με εκατοντάδες χιλιάδες χρήστες και δεκάδες χιλιάδες αξιολογήσεις ο υπολογισμός της τέλειας γειτνίασης είναι ακριβός και χρονοβόρος.

Σημαντικό ρόλο σε αυτούς τους αλγορίθμους που βασίζονται στους χρήστες κατέχει η μέτρηση ομοιότητας των χρηστών (Havaragi & Prathibha, 2016). Οι παραδοσιακές μέθοδοι υπολογισμού ομοιότητας εμφανίζουν δύο σημαντικές ελλείψεις:

- όλα τα αντικείμενα προς αξιολόγηση αντιμετωπίζονται το ίδιο

- δεν μπορεί να υπολογιστεί η ομοιότητα δύο χρηστών εάν δεν έχουν αξιολογήσει ένα ίδιο αντικείμενο (Havaragi & Prathibha, 2016).

2.3.1.2 Δημογραφικό φιλτράρισμα - Demographic Recommendation

Τα συστήματα αυτά πραγματοποιούν συστάσεις με βάση τη δημογραφική κατηγοριοποίηση (Pazzani, 1999). Ο αλγόριθμος ως είσοδο δέχεται δημογραφικές πληροφορίες του χρήστη (ηλικία, φύλο, εισόδημα, εκπαίδευση κ.λ.π.), τον κατατάσσει σε κάποια δημογραφική τάξη/κατηγορία και στη συνέχεια δημιουργεί συστάσεις με βάση τις δημογραφικές τάξεις. Το πλεονέκτημα αυτής της τεχνικής σε σύγκριση με τις άλλες είναι ότι παρακάμπτεται το πρόβλημα της ψυχρής εκκίνησης, για να γίνουν συστάσεις δεν χρειάζεται ιστορικό αξιολογήσεων οπότε ένας νέος χρήστης μπορεί να πάρει μια σύσταση πριν αξιολογήσει αντικείμενα (Wang, Chan, & Ngai, 2012).

2.3.1.3 Φιλτράρισμα βάσει περιεχομένου - Content-based Recommendation

Μια άλλη κατηγορία συστημάτων συστάσεων είναι το φιλτράρισμα βάσει περιεχομένου. Τα συστήματα αυτά κάνουν συστάσεις αναλύοντας το περιεχόμενο των πληροφοριών κειμένου, και βασίζονται στα χαρακτηριστικά των χρηστών και των αντικειμένων (Xiaoouyan & Taghi, 2009). Ο αλγόριθμος συστήνει αντικείμενα που είναι παρόμοια με αντικείμενα που αρέσανε στον χρήστη κατά το παρελθόν (Ricci, Rokach, Shapira, & Kantor, 2011). Η γενική λειτουργία αυτών των συστημάτων είναι να κατασκευάζουν ένα προφίλ χρήστη βάσει των χαρακτηριστικών των αντικειμένων που έχει αξιολογήσει ή/και καταναλώσει ο χρήστης. Αυτό το προφίλ χρησιμοποιείται για τον εντοπισμό νέων στοιχείων τα οποία ταιριάζουν με το προφίλ του και ενδεχομένως ενδιαφέρουν τον χρήστη. Ως είσοδο δέχεται αξιολογήσεις του χρήστη προς τα αντικείμενα. Δημιουργεί έναν ταξινομητή βασιζόμενο σε αυτές τις αξιολογήσεις, τον οποίο χρησιμοποιεί για να ταξινομήσει νέα αντικείμενα. Παράδειγμα, εάν ένας χρήστης αξιολογήσει θετικά μια κωμική ταινία, τότε το σύστημα μαθαίνει και του συστήνει ταινίες που τις έχει ταξινομήσει στην κατηγορία κωμικές. Και σε αυτά τα συστήματα συναντάμε το πρόβλημα της ψυχρής εκκίνησης που περιγράφηκε σε προηγούμενη ενότητα, τα ελλιπή δηλ. δεδομένα στην περίπτωση νέων χρηστών ή νέων αντικειμένων.

2.3.1.4 Φιλτράρισμα βασιζόμενο στη γνώση - Knowledge-based Recommendation

Τα συστήματα που βασίζονται στη γνώση παρέχουν συστάσεις σύμφωνα με τις

γνώσεις σχετικά με τον τρόπο με τον οποίο ένα στοιχείο ικανοποιεί τις απαιτήσεις ενός χρήστη (Wang, Chan, & Ngai, 2012). Ένα τέτοιο σύστημα δηλ. προτείνει προϊόντα βασισμένα σε συμπεράσματα σχετικά με τις ανάγκες και τις προτιμήσεις του χρήστη (Burke, 2007). Υπολογίζει τη χρησιμότητα κάθε στοιχείου στον χρήστη και στις ανάγκες των χρηστών. Συνάγει μια αντιστοιχία μεταξύ των στοιχείων και τις ανάγκες του χρήστη.

2.3.1.5 Υβριδικά συστήματα

Κάθε μία από τις προηγούμενες τεχνικές έχει μειονεκτήματα, όπως το πρόβλημα της ψυχρής εκκίνησης σε συστήματα συνεργατικού φιλτραρίσματος (collaborative) και συστήματα βάσει περιεχομένου (content-based). Ένα υβριδικό σύστημα συστάσεων είναι ένα σύστημα που συνδυάζει πολλαπλές τεχνικές μαζί για να επιτευχθεί κάποια συνέργεια μεταξύ τους (Burke, 2007).

Για παράδειγμα μπορεί να χρησιμοποιεί συνδυασμό συνεργατικού φιλτραρίσματος με δημογραφικό φιλτράρισμα ή συνεργατικό φιλτράρισμα με φιλτράρισμα βάσει περιεχομένου για να εκμεταλλευτεί τα πλεονεκτήματα κάθε μιας από αυτές τις τεχνικές. Βασίζεται συνήθως σε εμπνευσμένες από την επιστήμη της βιολογίας ή πιθανοτικές μεθόδους όπως οι γενετικοί αλγόριθμοι, fuzzy genetic, neural networks, Bayesian networks, latent features (Bobadilla, Ortega, Hernando, & Gutiérrez, 2013).

Ένας αυξανόμενος αριθμός δημοσιεύσεων απευθύνεται σε υβριδικές προσεγγίσεις που χρησιμοποιούν τρέχουσες βάσεις δεδομένων για την ταυτόχρονη ενσωμάτωση πληροφοριών βασισμένων στη μνήμη, κοινωνικών πληροφοριών και πληροφοριών περιεχομένου (Bobadilla et al., 2013). Όπως υποστηρίζουν οι Berka και Plöbzig (2004) λόγω των ετερογενών πληροφοριών στον τομέα τουρισμού αλλά και το ότι και οι τρεις βασικοί τύποι συστημάτων συστάσεων είναι κατάλληλοι, ευνοείται η χρήση υβριδικών συστημάτων στον τομέα του τουρισμού.

2.3.1.6 Geographic RS - GRS

Για τα συστήματα αυτής της κατηγορίας, οι συστάσεις παράγονται λαμβάνοντας υπόψη τη γεωγραφική θέση του χρήστη που λαμβάνει τη σύσταση (Bobadilla et al., 2013). Οι χρήστες ενός τέτοιου συστήματος αναμένουν να τους προτείνει το σύστημα ένα σημείο ενδιαφέροντος (π.χ. εστιατόριο, δραστηριότητες, κινηματογράφο κλπ.) όχι

μόνο των καλών αξιολογήσεων από άλλους παρόμοιους χρήστες αλλά και ανάλογα την απόσταση της τρέχουσας θέσης του χρήστη και της θέσης του σημείου ενδιαφέροντος (Bobadilla, Ortega, Hernando, & Gutiérrez, 2013).

2.3.1.7 Κοινωνικές πληροφορίες – Social Information

Η αυξανόμενη χρήση του Διαδικτύου και των Κοινωνικών Μέσων Δικτύωσης έχει οδηγήσει στο φαινόμενο κοινωνικής υπερφόρτωσης (Social overload). Διακρίνονται δύο κατηγορίες υπερφόρτωσης:

(α) Information overload από δημοσιεύσεις στα κοινωνικά μέσα όπως στο facebook και twitter, από uploads στο youtube. Πιο συγκεκριμένα:

- facebook: 2,07 δις ενεργοί χρήστες το 1^ο τρίμηνο του 2017, 2 τρις δημοσιεύσεις Ιούλιος του 2016¹⁹,
- twitter: 317 εκ. χρήστες μηνιαίως (Νοέμβριος του 2016), 1,6 δις μοναδικοί επισκέπτες μηνιαίως βλέπουν tweets. (Ιανουάριος 2017)²⁰, δημοσιεύονται κατά μέσο όρο περίπου 6.000 tweets το δευτερόλεπτο²¹,
- youtube: 1.5 δις συνδεδεμένοι χρήστες μηνιαίως (Ιούνιος 2017), κάθε λεπτό ανεβαίνουν βίντεο 400 ωρών (Αύγουστος 2017)²² και

(β) Interaction overload: φιλίες (friends), followers²³, following²⁴, like,

- facebook: ο μέσος χρήστης έχει 155 φίλους (Ιανουάριος 2016), 800 εκ. χρήστες έκαναν like σε μια ημέρα (Ιούνιος 2017)²⁵,
- twitter: περίπου 80 εκ. followers στους 10 πιο δημοφιλείς λογαριασμούς.

Τα συστήματα συστάσεων στοχεύουν στην αντιμετώπιση αυτού του φαινομένου παρουσιάζοντας πιο ελκυστικό και πιο σχετικό περιεχόμενο στους χρήστες.

Με την ανάπτυξη του Web 2.0 τα συστήματα συστάσεων ενσωματώνουν ολοένα και περισσότερες κοινωνικές πληροφορίες όπως αξιόπιστους και μη αξιόπιστους χρήστες, ακόλουθοι και ακολουθούμενοι, λίστες φίλων, δημοσιεύσεις, blogs-ιστολόγια, tags-ετικέτες κ.α. Χρήστες που είναι συνδεδεμένοι σε ένα δίκτυο εμπιστοσύνης παρουσιάζουν σημαντικά υψηλότερη ομοιότητα στα αντικείμενα και στα μεταδεδομένα

¹⁹ <https://expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats/>

²⁰ <https://expandedramblings.com/index.php/march-2013-by-the-numbers-a-few-amazing-twitter-stats/>

²¹ <http://www.internetlivestats.com/twitter-statistics/>

²² <https://expandedramblings.com/index.php/youtube-statistics/>

²³ Ο συνολικός αριθμός των followers ενός συγκεκριμένου προφίλ Twitter

²⁴ Ο συνολικός αριθμός προφίλ που ακολουθήθηκε από ένα συγκεκριμένο προφίλ Twitter

²⁵ <https://expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats/>

από τους μη συνδεδεμένους χρήστες. Οι πληροφορίες αυτές βελτιώνουν το πρόβλημα της ανεπάρκειας δεδομένων (sparsity).

2.3.1.8 Συστάσεις σε ομάδες χρηστών

Οι ομάδες χρηστών αρχίζουν να χρησιμοποιούνται σε διάφορους τομείς μεταξύ αυτών είναι και ο τομέας του τουρισμού. Πολλά αντικείμενα για τα οποία δημιουργούνται συστάσεις χρησιμοποιούνται τόσο από άτομα όσο και από ομάδες (Jameson, Baldes, & Kleinbauer, 2004). Στον τομέα του τουρισμού για παράδειγμα συστάσεις για επισκέψεις σε αξιοθέατα/μουσεία, πακέτα διακοπών, επιλογή εστιατορίων μπορεί να γίνονται τόσο σε μεμονωμένα άτομα όσο και σε ομάδες όπως μια παρέα φίλων ή ένα τουριστικό γκρουπ. Επιπλέον είναι πολύ πιθανό τα μέλη μιας ομάδας να χρησιμοποιούν φορητές συσκευές και να αλληλεπιδρούν μέσω αυτών των συσκευών.

Σύμφωνα με τους Jameson και Smyth (2004) διακρίνονται 4 βασικά ζητήματα στα συστήματα συστάσεων σε ομάδες:

- 1.** Απόκτηση πληροφοριών σχετικά με τις προτιμήσεις των μελών της ομάδας. Αυτό μπορεί να γίνει άμεσα ή έμμεσα.
 - a.** Άμεση απόκτηση όταν το σύστημα ζητά από το χρήστη να δηλώσει τις προτιμήσεις του.
 - b.** Έμμεση απόκτηση όταν το σύστημα συλλέγει πληροφορίες από την συμπεριφορά και τις ενέργειες του χρήστη για παράδειγμα πόσες φορές άκουσε ένα μουσικό κομμάτι.
- 2.** Προδιαγραφές προτιμήσεων που θα χρησιμοποιηθούν για την παραγωγή συστάσεων.
 - a.** Εστίαση στις αρνητικές προτιμήσεις. Σε κάποιες περιπτώσεις είναι πιο εύκολο και λιγότερο χρονοβόρο να λαμβάνονται υπόψη το τι δεν προτιμά ο χρήστης (αρνητικές προτιμήσεις) από αυτά που προτιμά, π.χ. για τις μουσικές προτιμήσεις χρηστών είναι χρονοβόρο να συγκεντρωθούν οι προτιμήσεις του κάθε χρήστη της ομάδας και γίνεται σπατάλη ενέργειας αφού πολλά από τα τραγούδια που προτιμά ο ένας δεν θα τα προτιμά άλλος χρήστης.
 - b.** Κοινή χρήση πληροφοριών σχετικά με τις καθορισμένες προτιμήσεις - συνεργατική μέθοδος συγκέντρωσης προτιμήσεων. Ένα μέλος της ομάδας μπορεί να έχει ενδιαφέρον για τις προτιμήσεις των άλλων μελών π.χ. για

κάποιο προορισμό διακοπών για τον οποίο δεν έχει προηγούμενη εμπειρία.

3. Δημιουργία συστάσεων.

Μια ευρέως εφαρμοσμένη προσέγγιση συγκέντρωσης προτιμήσεων είναι το σύστημα να χρησιμοποιεί τις πληροφορίες σχετικά με τις προτιμήσεις των μεμονωμένων μελών της ομάδας για να φτάσει σε ένα πρότυπο προτιμήσεων όλης της ομάδας. Μια τέτοια προσέγγιση είναι το σύστημα INTRIGUE το οποίο έχει σχεδιαστεί για να βοηθήσει τους ξεναγούς που σχεδιάζουν περιηγήσεις για ετερογενείς ομάδες τουριστών που περιλαμβάνουν σχετικά ομοιογενείς υποομάδες (Ardissono, Goy, Petrone, Segnan, & Torasso, 2003).

4. Εξηγώντας τις συστάσεις που προτάθηκαν.

Δεδομένων των πολλών τρόπων με τους οποίους μπορούν να παραχθούν συστάσεις για μια ομάδα - και των συχνά αντικρουόμενων στόχων που μπορούν να επιδιώκουν - είναι φυσικό τα μέλη της ομάδας να θέλουν να κατανοήσουν σε κάποιο βαθμό τον τρόπο με τον οποίο συστήνεται ένα αντικείμενο - και ειδικότερα, πόσο ελκυστικό μπορεί να είναι ένα συνιστώμενο στοιχείο για κάθε μεμονωμένο μέλος της ομάδας.

2.4 Εξηγώντας τις Συστάσεις

Ένα σημαντικό ερευνητικό θέμα στον τομέα συστημάτων συστάσεων επικεντρώνεται στην παροχή εξηγήσεων που δικαιολογούν τις συστάσεις που έχει λάβει ο χρήστης. Αυτή είναι μια σημαντική πτυχή ενός συστήματος επειδή βοηθά στη διατήρηση ενός υψηλότερου βαθμού εμπιστοσύνης των χρηστών στα αποτελέσματα που παράγονται από το σύστημα (Bobadilla et al., 2013). Ιδιαίτερως, όταν έχουμε συστάσεις σε ετερογενείς ομάδες χρηστών (tourist group) η δυνατότητα εξήγησης είναι απαραίτητη διότι τα αντικρουόμενα συμφέροντα των χρηστών μπορεί να οδηγήσουν σε συγκεχυμένες προτάσεις (Ardissono et al., 2003).

Είδη εξηγήσεων:

- Επεξηγήσεις στυλ αντικειμένου (προσέγγιση αντικείμενο σε αντικείμενο). Για παράδειγμα, συνιστούμε τον προορισμό διακοπών *i* επειδή σας άρεσαν οι προορισμοί διακοπών *g, c, r, ...* (*g, c, r, ...* είναι προορισμοί διακοπών παρόμοιοι με το *i* και βαθμολογούνται καλά από τον ενεργό χρήστη).
- Επεξηγήσεις στυλ χαρακτηριστικών (συνιστάται με βάση τα χαρακτηριστικά των αντικειμένων). Για παράδειγμα, συνιστούμε την ταινία *i* επειδή αυτή

σκηνοθετήθηκε από τον σκηνοθέτη d, παίζουν οι ηθοποιοί a, b, και ανήκει στο είδος ταινίας g (d, a, b, g είναι χαρακτηριστικά που ενδιαφέρουν τον ενεργό χρήστη).

- Υβριδικές μέθοδοι. Αυτή η κατηγορία περιλαμβάνει κατά κύριο λόγο τα εξής: άνθρωπος / αντικείμενο, άνθρωπος / χαρακτηριστικό, χαρακτηριστικό / αντικείμενο και άνθρωπος / χαρακτηριστικό / αντικείμενο.

2.5 Ηθικά ζητήματα - Ζητήματα ιδιωτικού απόρρητου

Η προστασία του ιδιωτικού απορρήτου είναι ένα από τα σημαντικά ζητήματα των συστάσεων. Για να μπορέσει ένα σύστημα να παράγει ακριβείς συστάσεις χρειάζεται πρόσβαση σε ιδιωτικά δεδομένα. Αυτό είναι ένα κρίσιμο ζήτημα γιατί η άρνηση των χρηστών να παρέχουν τέτοιου είδους δεδομένα μπορεί να οδηγήσει στη μη χρησιμότητα των εξατομικευμένων συστημάτων.

Σύμφωνα με τους Lam et al. (2006), υπάρχουν τρεις κατηγορίες χρηστών σχετικά με τις αποφάσεις τους και τις επιλογές τους όσον αφορά τις ιδιωτικές πληροφορίες που παρέχουν:

1. Χρήστες που παρέχουν κάθε είδους πληροφορία σε ένα σύστημα συστάσεων με αντάλλαγμα περισσότερο εξατομικευμένο περιεχόμενο.
2. Χρήστες που παρέχουν κάποιες πληροφορίες για ένα σύστημα συστάσεων, προκειμένου να πάρουν βελτιωμένες προτάσεις.
3. Χρήστες που δεν θα δώσουν καμία πληροφορία σε ένα σύστημα συστάσεων, λόγω των ανησυχιών τους για την ιδιωτικότητά τους.

Η δεύτερη κατηγορία χρηστών περιορίζεται σε πληροφορίες όπως το φύλλο, η ηλικία, και η εκπαίδευση. Τέτοιου είδους πληροφορίες δίνονται ευκολότερα από τις πιο προσωπικές (Jeckmans, et al., 2013). Όμως τα συστήματα συστάσεων χρειάζονται περισσότερα δεδομένα. Οι χρήστες θα δώσουν πιο προσωπικά δεδομένα αν πεισθούν ότι αυτά θα είναι ασφαλή. Υπάρχουν διάφορες μέθοδοι διαφύλαξης της ιδιωτικότητας στα συστήματα συστάσεων.

2.6 Προφίλ χρήστη στα συστήματα συστάσεων

Σύμφωνα με τους Montaner et al. (2003), η κατασκευή ενός ακριβούς προφίλ χρήστη

είναι ζωτικής σημασίας για το σύστημα, γιατί θα μπορεί να αναπαριστά με ακρίβεια τα ενδιαφέροντα του χρήστη. Για παράδειγμα στα συνεργατικά συστήματα εξασφαλίζει ότι χρήστες με παρόμοια προφίλ είναι πράγματι παρόμοιοι.

Υπάρχουν διάφορες τεχνικές αναπαράστασης (Montaner, López, & de la Rosa, 2003):

- Μοντέλο βάσει ιστορικού (History-based model): διατηρείται μια λίστα με τις αγορές, το ιστορικό πλοήγησης στο web ή το περιεχόμενο των θυρίδων ηλεκτρονικού ταχυδρομείου ως προφίλ χρήστη. Επιπλέον, διατηρείται και η σχετική ανατροφοδότηση του χρήστη που σχετίζεται με κάθε στοιχείο στο ιστορικό.
- Μοντέλο διανυσματικού χώρου (Vector space model): σε αυτό το μοντέλο τα αντικείμενα αντιπροσωπεύονται με ένα διάνυσμα χαρακτηριστικών, (λέξεων ή εννοιών), με μια σχετική τιμή. Η τιμή αυτή μπορεί να είναι ένας αριθμός Boolean ή ένας πραγματικός αριθμός.
- Σταθμισμένα n-grams (Weighted n-grams): τα στοιχεία αντιπροσωπεύονται με ένα δίκτυο λέξεων με βάρη στους κόμβους και τις άκρες.
- Σταθμισμένα σημασιολογικά δίκτυα (Weighted associative networks): αποθηκεύουν τις έννοιες των λέξεων, έτσι ώστε είναι δυνατή η ανθρώπινη χρήση των εννοιών αυτών. Στο σύστημα ifWeb (Asnicar & Tasso, 1997) το μοντέλο χρησιμοποιεί ένα σταθμισμένο σημασιολογικό δίκτυο του οποίου οι κόμβοι αντιστοιχούν σε όρους (έννοιες) που βρίσκονται σε έγγραφα του Web και τόξα συνδέουν τους όρους μεταξύ τους. Μια σημασιολογική βάση δικτύου περιέχει μια συλλογή σημασιολογικών δικτύων που περιγράφει ένα τυπικό μοτίβο θεμάτων ενδιαφέροντος για τον χρήστη.
- Μοντέλα που βασίζονται σε ταξινομητές (Classifier-based models): χρησιμοποιούν έναν ταξινομητή ως τεχνική μάθησης προφίλ χρήστη και διατηρούν τη δομή του ταξινομητή ως προφίλ. Όπως στην περίπτωση των νευρωνικών δικτύων, των δέντρων αποφάσεων και των δικτύων Bayesian.
- Μήτρα αξιολογήσεων χρήστη-στοιχείου (User-item ratings matrix): συναντάται σε ορισμένα συνεργατικά συστήματα, τα οποία διατηρούν ένα πίνακα βαθμολογίας χρήστη-στοιχείου ως προφίλ χρήστη. Ο πίνακας αυτός περιέχει το ιστορικό των αξιολογήσεων του χρήστη. Κάθε κελί (u, i) περιέχει μια βαθμολογία που αντιπροσωπεύει την αξιολόγηση του χρήστη u στο στοιχείο i. Κενή τιμή σημαίνει ότι δεν υπάρχει αξιολόγηση.

- Δημογραφικά χαρακτηριστικά (Demographic features): η αναπαράσταση προφίλ χρήστη είναι μια λίστα δημογραφικών χαρακτηριστικών που αντιπροσωπεύουν το είδος του χρήστη.

2.6.1 Παραγωγή αρχικού προφίλ – Τεχνικές προσαρμογής προφίλ

Όσο περισσότερα στοιχεία έχουμε για το χρήστη, τόσο καλύτερα αποτελέσματα θα δώσουν τα συστήματα συστάσεων. Όμως οι χρήστες δεν είναι διατεθειμένοι να αφιερώσουν χρόνο για τη δημιουργία του προφίλ τους και επιπλέον οι προτιμήσεις και τα ενδιαφέροντά τους μεταβάλλονται με την πάροδο του χρόνου. Παρακάτω παρουσιάζονται τεχνικές παραγωγής αρχικού προφίλ (Montaner, López, & de la Rosa, 2003):

Τεχνικές παραγωγής αρχικού προφίλ:

- Κενό (Empty): δεν γίνεται καμία αρχικοποίηση του προφίλ, το προφίλ δομείται καθώς ο χρήστης αλληλεπιδρά με το σύστημα.
- Χειροκίνητο (Manual): το σύστημα ζητά από τους χρήστες να καταχωρήσουν τις προτιμήσεις τους. Το πλεονέκτημα είναι η διαφάνεια της συμπεριφοράς του συστήματος. Το μειονέκτημα είναι ότι απαιτεί προσπάθεια από τη μεριά του χρήστη. Επιπλέον οι χρήστες ίσως να μην μπορούν να καθορίσουν τις προτιμήσεις τους.
- Στερεότυπο (Stereotyping): οι χρήστες καλούνται να συμπληρώσουν τα δημογραφικά τους στοιχεία και μετά το σύστημα τους ταξινομεί σύμφωνα με τα στοιχεία αυτά. Το μειονέκτημα αυτής της μεθόδου είναι ότι δύσκολα οι χρήστες θα δώσουν τα προσωπικά τους στοιχεία λόγω έλλειψης εμπιστοσύνης στην πολιτική προστασίας προσωπικών δεδομένων.
- Σύνολο εκπαίδευσης (Training set): το προφίλ δημιουργείται από την αλληλεπίδραση του χρήστη με κάποια παραδείγματα, τα οποία αξιολογεί ως σχετικά ή άσχετα με τις προτιμήσεις του. Το μειονέκτημα σε αυτή την μέθοδο είναι ότι τα παραδείγματα που θα επιλεγούν μπορεί να μην είναι πάντα αντιπροσωπευτικά και αυτό να οδηγήσει σε λιγότερο ακριβή αποτελέσματα.

Τεχνικές προσαρμογής προφίλ:

Οι προτιμήσεις και τα ενδιαφέροντα των χρηστών δεν παραμένουν σταθερά με την πάροδο του χρόνου. Οι πιο πρόσφατες παρατηρήσεις αντιπροσωπεύουν καλύτερα τα

τρέχοντα ενδιαφέρονται του χρήστη από τα παλιότερα. Υπάρχει λοιπόν η ανάγκη για μια τεχνική η οποία θα προσαρμόζει το προφίλ χρήστη στα νέα ενδιαφέροντά του και θα ξεχνάει τα πιο παλιά. Παρουσιάζονται αυτές οι τεχνικές προσαρμογής προφίλ (Montaner, López, & de la Rosa, 2003):

- Χειροκίνητα (Manual): ο χρήστης τροποποιεί το προφίλ του ο ίδιος, προσθέτοντας ή αποκλείοντας προτιμήσεις/ενδιαφέροντα. Ισχύουν τα ίδια μειονεκτήματα που περιγράφηκαν προηγουμένως στην χειροκίνητη τεχνική παραγωγής αρχικού προφίλ.
- Προσθέτοντας νέες πληροφορίες (Add new information): η πιο συνηθισμένη προσέγγιση. Η ενημέρωση του προφίλ γίνεται με την προσθήκη νέων πληροφοριών από την αλληλεπίδραση του χρήστη με το σύστημα. Έτσι το προφίλ προσαρμόζεται στα νέα ενδιαφέροντα του χρήστη. Το κύριο μειονέκτημα είναι ότι τα παλιά ενδιαφέροντα/προτιμήσεις δεν ξεχνιούνται.
- Βαθμιαία λειτουργία λήθης (Gradual forgetting function): Οι Webb και Kuzmycz εισάγουν το 1966 την έννοια της βαθμιαίας λήθης. Προτείνουν έναν μηχανισμό γήρανσης των δεδομένων, ο οποίος σε κάθε παρατήρηση δίνει ένα αρχικό βάρος 1. Κάθε φορά που ενσωματώνεται στο μοντέλο μια νέα παρατήρηση αυτό το αρχικό βάρος των προηγούμενων παρατηρήσεων μειώνεται. Έτσι οι πιο πρόσφατες παρατηρήσεις είναι πιο «σημαντικές» με την έννοια ότι αντιπροσωπεύουν καλύτερα τα ενδιαφέροντα του χρήστη.
- Φυσική επιλογή (Natural selection): αυτή η τεχνική εφαρμόζεται σε συστήματα που βασίζονται στους γενετικούς αλγορίθμους. Οι πράκτορες που παράγουν καλύτερες συστάσεις αναπαράγονται ενώ οι άλλοι καταστρέφονται.

Κεφάλαιο 3

Επιστημονική Βιβλιογραφία

Στον τομέα του τουρισμού, έχουν αναπτυχθεί πολλά συστήματα συστάσεων, εκ των οποίων τα περισσότερα από αυτά είναι βασισμένα στο περιεχόμενο (content-based) και στη γνώση (knowledge-based) ενώ περιορίζονται στην παραγωγή συστάσεων σε επίπεδο προορισμού (Wang, Chan, & Ngai, 2012).

Ωστόσο, τα συστήματα συστάσεων τουρισμού βρίσκονται ακόμη στο στάδιο της ανάπτυξης, χωρίς να επιτυγχάνεται το επίπεδο επιτυχίας όπως και στους τομείς των βιβλίων ή των ταινιών. Ως εκ τούτου, εξακολουθεί να υπάρχει περιθώριο βελτίωσης (Wang, Chan, & Ngai, 2012).

Δύο από τις πιο επιτυχημένες τεχνολογίες συστάσεων είναι η Triplehop's TripMatcher και Vacation Coach's expert advice platform (Ricci, 2002), χρησιμοποιούν την μέθοδο φιλτραρίσματος βάσει περιεχομένου για να παράγουν συστάσεις τουριστικών προορισμών (Wang, Chan, & Ngai, 2012). Προσπαθούν να μιμηθούν την διαδικασία κατά την οποία ένας χρήστης ζητάει συμβουλές από ένα ταξιδιωτικό πρακτορείο για να επισκεφθεί έναν τουριστικό προορισμό. Οι δύο αυτές τεχνολογίες χρησιμοποιούν όπως αναφέρθηκε, κυρίως την προσέγγιση βάση περιεχομένου (Content-based) όπου οι χρήστες εκφράζουν τις προτιμήσεις τους (ανάγκες, οφέλη, περιορισμούς) και το σύστημα αντιστοιχεί αυτές τις προτιμήσεις με μια λίστα προορισμών με την ίδια περιγραφή. Η Vacation Coach δημιουργεί προφίλ χρηστών είτε ζητώντας από τους χρήστες να συμπληρώσουν κατάλληλη φόρμα, είτε να επιλέξουν οι ίδιοι να ταξινομηθούν σε ένα από τα προτεινόμενα προφίλ ταξιδιώτη. Ενώ η Triplehop's υποθέτει τα χαρακτηριστικά του χρήστη λαμβάνοντας υπόψη ερωτήματα/αναζητήσεις του χρήστη.

Στην εργασία των García-Crespo et al. (2009) παρουσιάζεται το σύστημα SPETA (Social pervasive e-Tourism advisor) το οποίο χρησιμοποιεί την τρέχουσα θέση του χρήστη, τις

προτιμήσεις του καθώς και το ιστορικό των τοποθεσιών που επισκέφθηκε κατά το παρελθόν, ώστε να παράγει συστάσεις. Ενσωματώνει τεχνολογίες επίγνωσης περιβάλλοντος (context-aware), συστήματα GIS, κοινωνικά και σημασιολογικά δίκτυα. Προσφέρει έναν εξατομικευμένο οδηγό, ο αλγόριθμος του οποίου χειρίζεται όλες τις τεχνολογίες που αναφέρθηκαν ώστε να προσφέρει στον χρήστη τις καλύτερες συστάσεις. Το σύστημα είναι προσανατολισμένο προς τους τουρίστες οι οποίοι επισκέπτονται μια πόλη που δεν γνωρίζουν και χρειάζονται καθοδήγηση. Με βάση το προφίλ του χρήστη, τις προτιμήσεις του, το ιστορικό επισκέψεών του, την τοποθεσία σε πραγματικό χρόνο, την πρόγνωση καιρού, την ώρα, τις συστάσεις των φίλων του προτείνει συστάσεις για νέους προορισμούς.

Συστήματα συστάσεων σε ομάδες όπως το POCKET RESTAURANTFINDER (McCarthy, 2002) το οποίο παράγει συστάσεις σε μια ομάδα ανθρώπων που θέλουν να επισκεφθούν μαζί κάποιο εστιατόριο, με βάση την τοποθεσία των ατόμων και των γαστρονομικών τους προτιμήσεων. Οι χρήστες αυτού του συστήματος δηλώνουν την προτίμησή τους (μου αρέσει/δεν μου αρέσει) για κάθε ένα από 15 τύπους κουζίνας, για παροχές των εστιατορίων, για το εύρος τιμής για την απόσταση από την τρέχουσα τοποθεσία. Όταν συγκεντρωθεί μια ομάδα ατόμων το σύστημα συγκεντρώνει και επεξεργάζεται τις προτιμήσεις τους και παρουσιάζει μια λίστα εστιατορίων ταξινομημένη κατά σειρά αναμενόμενης επιθυμίας για την ομάδα (McCarthy, 2002).

Ένα άλλο σύστημα, το TRAVEL DECISION FORUM, βοηθά μια ομάδα χρηστών, να συμφωνήσουν για τα επιθυμητά χαρακτηριστικά των διακοπών που σχεδιάζουν και να καταλήξουν σε συμφωνία για μια συνολική λύση (Jameson, Baldes, & Kleinbauer, 2004). Το σύστημα πρέπει να γνωρίζει ποιες είναι οι προτιμήσεις κάθε χρήστη για δεκάδες χαρακτηριστικά των προορισμών διακοπών, τα οποία κυμαίνονται από τις εγκαταστάσεις που είναι διαθέσιμες στα δωμάτιά τους έως τα αξιοθέατα που είναι διαθέσιμα στη γύρω περιοχή. Ιδιαίτερη προσοχή δίνεται στη στήριξη των χρηστών οι οποίοι δεν είναι σε θέση να επικοινωνήσουν άμεσα (Jameson, 2004).

Στην εργασία των Ardissono et al. (Ardissono, Goy, Petrone, Segnan, & Torasso, 2003) παρουσιάζεται το INTRIGUE (INteractive TouRist Information GUIDe), ένας ηλεκτρονικός τουριστικός οδηγός που παρέχει εξατομικευμένες πληροφορίες για τα τουριστικά αξιοθέατα στην περιοχή του Torino της Ιταλίας. Το σύστημα αυτό

δημιουργεί συστάσεις προορισμών και δρομολογίων όχι μόνο σε μεμονωμένους χρήστες αλλά και σε ετερογενείς ομάδες χρηστών-τουριστών όπως οικογένειες με παιδιά και ηλικιωμένοι με προβλήματα κινητικότητας. Χωρίζει την ομάδα σε υποομάδες μελών με παρόμοια χαρακτηριστικά και ιδιότητες και εξετάζει τις συνολικές προτιμήσεις των υποομάδων χωρίς να εξετάζει το κάθε μεμονωμένο μέλος ξεχωριστά. Επίσης δίνει στις υποομάδες και διαφορετικό βαθμό επιρροής στις προτιμήσεις. Και αυτό γιατί μπορεί τα μέλη μιας υποομάδας να έχουν ιδιαίτερες ανάγκες όπως τα παιδιά ή τα άτομα με κινητικά προβλήματα.

Άλλες έρευνες προσανατολίζονται σε συστήματα συστάσεων για σημεία ενδιαφέροντος (POI) σε ομάδες χρησιμοποιώντας δεδομένα κοινωνικής δικτύωσης βάσει τοποθεσίας. Στην εργασία των Gottapu και Sriram Monangi (2017) παρουσιάζεται ένα τέτοιο σύστημα συστάσεων, που προτείνει σημεία ενδιαφέροντος, όπως εστιατόρια, σε ομάδες χρηστών με βάση το μέγεθος της ομάδας, της τρέχουσας θέσης, της ώρας και άλλων παραμέτρων. Το σύστημα ανακαλύπτει ποιοι χρήστες επισκέπτονται ένα σημείο ενδιαφέροντος ως ομάδα, και δημιουργεί τα χαρακτηριστικά αυτής της ομάδας. Στη συνέχεια δημιουργεί υπογραφές για όλες τις ομάδες που επισκέφθηκαν αυτό το σημείο. Μια υπογραφή POI περιγράφει ιδιότητες της ομάδας, όπως ο αριθμός των χρηστών, ο αριθμός των σχέσεων μεταξύ τους, το πόσο ισχυρές είναι αυτές οι σχέσεις. Για μια νέα ομάδα που ζητάει συστάσεις για POI, υπολογίζεται η υπογραφή για αυτήν την ομάδα και παράγεται μια σύσταση χρησιμοποιώντας τους K-πλησιέστερους γείτονες (KNN).

Κεφάλαιο 4

Περιγραφή συστήματος

4.1 Εισαγωγή

Στο παρόν κεφάλαιο περιγράφεται ο σχεδιασμός ενός ταξιδιωτικού συστήματος συστάσεων το οποίο διαφοροποιείται από τα ήδη υπάρχοντα συστήματα ως προς τα εξής:

- η μέτρηση ομοιότητας των χρηστών υπολογίζεται από τις ταξιδιωτικές επιλογές και προτιμήσεις των χρηστών και από τη σημαντικότητα που δίνουν σε αυτές και όχι από τις αξιολογήσεις των προορισμών,
- οι ίδιοι οι χρήστες δηλώνουν από ποιες κατηγορίες χρηστών θα θέλανε να προέρχονται οι συστάσεις και τη σημαντικότητα που δίνουν σε κάθε κατηγορία, κάτι που δεν γίνεται σε άλλα υπάρχοντα συστήματα.

Αυτές οι ρητές δηλώσεις του χρήστη προς το σύστημα οδηγούν σε αξιόπιστα συμπεράσματα σχετικά με τις προτιμήσεις του χρήστη σε σχέση με άλλα συστήματα τα οποία υποθέτουν τις προτιμήσεις του χρήστη από την συμπεριφορά του και τις αξιολογήσεις του. Το αποτέλεσμα είναι να παρέχονται στον χρήστη πιο ακριβείς εξατομικευμένες συστάσεις οι οποίες ανταποκρίνονται στα προσωπικά του χαρακτηριστικά. Επιπλέον το σύστημα μπορεί να δικαιολογήσει στον χρήστη για ποιο λόγο του γίνονται οι συγκεκριμένες συστάσεις, πράγμα το οποίο αυξάνει τον βαθμό εμπιστοσύνης του χρήστη προς το σύστημα.

Κανένας από τους δημοφιλείς ταξιδιωτικούς ιστότοπους (tripadvisor, booking κλπ.) που παρουσιάστηκαν στο Κεφάλαιο 3 δεν παρέχει εξηγήσεις που να δικαιολογούν τις συστάσεις που λαμβάνει ο χρήστης. Οι συστάσεις που παρέχονται βασίζονται είτε σε αξιολογήσεις, είτε σε κάποιες γενικές ταξιδιωτικές προτιμήσεις που μπορεί να συμπεριλάβει ο χρήστης στο προφίλ του. Τέλος, σε κανένα από όλα τα συστήματα που αναφέρθηκαν ο χρήστης δεν έχει τη δυνατότητα (α) να δηλώσει τη σημαντικότητα των

παραγόντων που τον επηρεάζουν στην επιλογή ενός προορισμού και (β) να επιλέξει από ποιους χρήστες θα προέρχονται οι συστάσεις.

Σε ένα CF-based memory οι χρήστες ομαδοποιούνται σε κατηγορίες με βάση το ιστορικό των αξιολογήσεών τους, έπειτα υπολογίζεται η πιθανότητα για έναν χρήστη να του αρέσει ένα συγκεκριμένο στοιχείο ώστε να του προταθεί από το σύστημα. Τα μέτρα ομοιότητας που εφαρμόζονται σε αυτά τα συστήματα (π.χ. cosine, Pearson correlation coefficient) υπολογίζουν την ομοιότητα σε στοιχεία τα οποία έχουν αξιολογηθεί από δύο χρήστες. Σε αυτή την εργασία εξετάζεται ένα διαφορετικό μοντέλο μέτρησης ομοιότητας χρηστών. Η ομοιότητα μεταξύ των χρηστών υπολογίζεται από τα δημογραφικά τους στοιχεία και τη σημαντικότητα που δίνουν σε αυτά για την παραγωγή συστάσεων, από τους παράγοντες που τους επηρεάζουν για την επιλογή των ταξιδιών, και τέλος από τις επιλογές των ταξιδιών τους και τη σημαντικότητα που δίνουν σε αυτές για την παραγωγή συστάσεων. Το σύστημα λοιπόν δεν βασίζεται σε αξιολογήσεις αντικειμένων για να υπολογίσει την ομοιότητα των χρηστών αλλά βασίζεται στις προτιμήσεις και επιλογές τους πιο συγκεκριμένα, στα κριτήρια τα οποία θεωρούν οι ίδιοι σημαντικά σε ένα ταξίδι.

Το σύστημα λοιπόν, παράγει συστάσεις στον χρήστη u για νέους προορισμούς i που έχουν επισκεφθεί οι πλησιέστεροι γείτονες προς αυτόν. Ως πλησιέστεροι γείτονες εννοούνται χρήστες που μοιάζουν περισσότερο με τον χρήστη u . Κατά την είσοδο στο σύστημα ο χρήστης καλείται να απαντήσει σε ερωτήσεις σχετικά με το ποιοι παράγοντες είναι αυτοί που τον επηρεάζουν στην επιλογή ενός ταξιδιωτικού προορισμού.

Στο Παράρτημα Α παρουσιάζονται οι ερωτήσεις του συστήματος, στο Παράρτημα Β οι πίνακες προφίλ χρήστη και προορισμού και στο Παράρτημα Γ το διάγραμμα βασικών λειτουργιών του συστήματος.

4.2 Σύστημα

Το σύστημα διατηρεί προφίλ για κάθε χρήστη του συστήματος. Το προφίλ δημιουργείται απαντώντας σε ερωτήσεις του συστήματος κατά την εγγραφή του, ή οποιαδήποτε στιγμή θελήσει ο χρήστης. Για τους χρήστες που δεν θέλουν να απαντήσουν στις ερωτήσεις προτιμήσεων γίνεται δημογραφικό φιλτράρισμα, ενώ σε

αυτούς που δίνουν ελλιπείς απαντήσεις, το σύστημα τους αποδίδει τον μέσο όρο απαντήσεων των υπόλοιπων χρηστών. Επιπλέον διατηρείται για τον κάθε χρήστη ιστορικό τουριστικών προορισμών που επισκέφθηκε καθώς και αξιολόγηση των προορισμών σε κλίμακα 1 έως 5 από τον χρήστη. Στις επόμενες παραγράφους γίνεται αναλυτική περιγραφή.

4.2.1 Είσοδος

Κατά την είσοδό του στο σύστημα, ο χρήστης μπορεί να εγγραφεί ή να συνδεθεί αν είναι ήδη εγγεγραμμένος. Γίνεται έλεγχος στοιχείων λογαριασμού, σε περίπτωση λάθους εμφανίζει αντίστοιχο μήνυμα.

4.2.2 Ερωτήσεις συστήματος

Έπειτα καλείται να απαντήσει σε ερωτήσεις οι οποίες παρουσιάζονται στο Παράρτημα A.1 της εργασίας και χωρίζονται στα εξής βασικά μέρη: α) Δημογραφικά Στοιχεία, (β) Παράγοντες που επηρεάζουν τον χρήστη-τουρίστα ως προς την επιλογή ενός προορισμού, Προτιμήσεις χρηστών-τουριστών ως προς τα ταξίδια (γ) και (δ) Προτιμήσεις χρηστών ως προς τις συστάσεις.

Οθόνη πρώτη – Δημογραφικά στοιχεία

Συμπληρώνει δημογραφικά στοιχεία, αυτό γίνεται κατά την 1^η είσοδο του χρήστη στο σύστημα (Παράρτημα A.1). Τις επόμενες φορές η οθόνη αυτή δεν εμφανίζεται, η αλλαγή των στοιχείων γίνεται από κατάλληλο μενού π.χ. Ο λογαριασμός μου. Οι απαντήσεις αποθηκεύονται σε έναν πίνακα data_D. Οι ερωτήσεις είναι υποχρεωτικές ώστε αν ο χρήστης δεν θελήσει να απαντήσει στις επόμενες ερωτήσεις, το σύστημα να μπορέσει βάσει των δημογραφικών του στοιχείων να υπολογίσει την ομοιότητα με άλλους χρήστες και να του προτείνει προορισμούς.

Οθόνη δεύτερη – Παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές

Οι ερωτήσεις (Παράρτημα A.2) σε αυτή την οθόνη διερευνούν τους ισχυρούς παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές του χρήστη. Παράδειγμα, «Πόσο σημαντικό είναι για εσάς ο προορισμός να διαθέτει φυσικά θέλητρα;», ή «Πόσο σημαντικό είναι για εσάς το κλίμα ενός προορισμού για την επιλογή του;», όπου οι απαντήσεις είναι σε κλίμακα 1 (Καθόλου σημαντικό) έως 10 (Πολύ σημαντικό). Τέλος το σύστημα ζητάει να προτείνουν και οι ίδιοι χρήστες σημαντικούς παράγοντες για

αυτούς για την επιλογή ενός προορισμού καθώς και να προσδιορίσουν το επίπεδο σημαντικότητας αυτών. Οι απαντήσεις αποθηκεύονται στον πίνακα `data_pref`.

Οθόνη τρίτη – Προτιμήσεις συστάσεων

Εδώ ο χρήστης απαντά σε ερωτήσεις σχετικές με την προέλευση των συστάσεων (Παράρτημα A.3). Από ποιους χρήστες δηλ. θα ήθελε να προέρχονται οι συστάσεις. Επιπλέον το σύστημα τους δίνει τη δυνατότητα να προτείνουν και οι χρήστες από ποια κατηγορία χρηστών θα θέλανε συστάσεις. Οι απαντήσεις των πρώτων οκτώ ερωτήσεων (με βάση δημογραφικά χαρακτηριστικά) αποθηκεύονται στον πίνακα `pref_rec_dem`, ενώ οι υπόλοιπες (με βάση ίδιων προτιμήσεων) στον πίνακα `pref_rec`.

Οθόνη τέταρτη– Προτιμήσεις ταξιδιού των χρηστών

Ο χρήστης απαντάει σε ερωτήσεις (Παράρτημα A.4) σχετικές με τις προτιμήσεις του και την συχνότητα επιλογής αυτών των προτιμήσεων σε πέντε κατηγορίες, σκοπό ταξιδιού, μεταφορικό μέσο, με ποιον ταξιδεύει, προορισμό εσωτερικό ή εξωτερικό και διάρκεια ταξιδιού. Οι απαντήσεις σε αυτές τις ερωτήσεις συνδυάζονται με τις αντίστοιχες απαντήσεις στις ερωτήσεις σημαντικότητας (οθόνη δεύτερη) για κάθε ένα από αυτά τα πέντε κριτήρια.

4.3 Προφίλ χρηστών

Το προφίλ κάθε χρήστη αποτελείται από πέντε πίνακες, `data_D`, `data_pref`, `pref_rec_dem`, `pref_rec`, `travel_history`. Παρακάτω παρουσιάζονται αναλυτικά οι πίνακες:

- **data_D**: στον πίνακα αποθηκεύονται τα δημογραφικά στοιχεία που καταχωρεί ο χρήστης κατά την εγγραφή του στο σύστημα. Τα πεδία αυτού του πίνακα είναι *user_id* (αναγνωριστικό χρήστη), *gender* (φύλο), *age* (ηλικία), *nationality* (εθνικότητα), *denomination* (θρήσκευμα), *education* (εκπαίδευση), *profession* (επάγγελμα), *income* (εισόδημα), *disabilities* (δυσκολίες). Οι τιμές αυτού του πίνακα χρησιμοποιούνται σε δύο περιπτώσεις:
 - για τον υπολογισμό ομοιότητας δημογραφικών χαρακτηριστικών μεταξύ ενός χρήστη *u* με τους υπόλοιπους *v* χρήστες του συστήματος που έχουν απαντήσει στις υπόλοιπες ερωτήσεις του συστήματος.
 - για δημογραφικό φιλτράρισμα στην περίπτωση που ο χρήστης *u* δεν θέλει να απαντήσει στις υπόλοιπες ερωτήσεις του συστήματος.

- **data_pref:** στον πίνακα αποθηκεύονται οι ταξιδιωτικές προτιμήσεις του χρήστη, τη σημαντικότητα που δίνει ο χρήστης σε κάποιους παράγοντες για την επιλογή ενός τουριστικού προορισμού. Οι τιμές είναι από 1 έως 10, και τα πεδία αυτού του πίνακα είναι *user_id* (αναγνωριστικό χρήστη), *nature* (φυσικά θέλγητρα), *monuments* (κατασκευασμένα θέλγητρα), *culture* (πολιτιστικά θέλγητρα), *society* (κοινωνικά θέλγητρα), *cost* (κόστος ταξιδιού), *transport* (μεταφορικό μέσο), *access* (ευκολία πρόσβασης), *safety* (αίσθημα ασφάλειας), *r_tr_network* (οδικό δίκτυο και δίκτυο μεταφορών), *facilities_services* (εγκαταστάσεις και υπηρεσίες τουρισμού), *clime* (κλίμα-καιρικές συνθήκες), *free_field* (ελεύθερο πεδίο για να καταγράψει ο χρήστης τι άλλο είναι σημαντικό για αυτόν).
- **pref_rec_dem:** στον πίνακα αποθηκεύονται οι προτιμήσεις του χρήστη για τις συστάσεις που θα του γίνουν με βάση τα ίδια δημογραφικά χαρακτηριστικά που έχει με άλλους χρήστες. Οι τιμές είναι από 1 έως 10 και τα πεδία του πίνακα είναι: *user_id* (αναγνωριστικό χρήστη), *rec_gender* (συστάσεις βάσει φύλου), *rec_age* (συστάσεις βάσει ηλικίας), *rec_nation* (συστάσεις βάσει εθνικότητας), *rec_denom* (συστάσεις βάσει θρησκειώματος), *rec_edu* (συστάσεις βάσει εκπαίδευσης), *rec_prof* (συστάσεις βάσει επαγγέλματος), *rec_incom* (συστάσεις βάσει εισοδήματος), *rec_status* (συστάσεις βάσει οικογενειακής κατάστασης).
- **pref_rec:** στον πίνακα αποθηκεύονται οι προτιμήσεις του χρήστη για συστάσεις που θα του γίνουν με βάση ίδιες επιλογές ως προς τις ταξιδιωτικές επιλογές. *user_id* (αναγνωριστικό χρήστη), *rec_purpose* (ίδιος σκοπός ταξιδιού), *rec_destination* (ίδιος προορισμός), *rec_durat* (ίδια διάρκεια), *free_field* (ελεύθερο πεδίο για να καταγράψει ο χρήστης τι άλλο είναι σημαντικό για αυτόν).
- **travel_history:** στον πίνακα αποθηκεύεται το ιστορικό ταξιδιών του χρήστη, με πεδία: *user_id* (αναγνωριστικό χρήστη), *country* (χώρα προορισμού), *city* (πόλεις που επισκέφθηκε), *purpose* (σκοπός ταξιδιού), *company* (με ποιόν ταξίδεψε), *transport* (με τι μεταφορικό μέσο ταξίδεψε), *destination* (προορισμός εσωτερικό/εξωτερικό), *duration* (διάρκεια ταξιδιού), *date_tr* (ημερομηνία ταξιδιού), *rate* (αξιολόγηση).

Τέλος υπάρχει και ο πίνακας **travel**, με πεδία *destination_id* (αναγνωριστικό προορισμού), *country* (χώρα), *city* (πόλεις), *purpose* (σκοπός), *company* (με ποιον ταξίδεψε), *transport* (μεταφορικό μέσο), *destination* (εσωτερικό/εξωτερικό), *duration* (διάρκεια). Τα πεδία ενημερώνονται από το ιστορικό ταξιδιών των χρηστών (πίνακας

travel_history), οι τιμές που αποθηκεύονται είναι οι απαντήσεις της πλειοψηφίας των χρηστών. Για παράδειγμα για τη χώρα Γερμανία, και την πόλη Βερολίνο οι περισσότεροι χρήστες ταξίδεψαν εκεί για Αναψυχή/Διακοπές, με παιδιά, επέλεξαν το αεροπλάνο, ήταν προορισμός εκτός της χώρας διαμονής τους (εξωτερικό) και η διάρκεια 2-7 διανυκτερεύσεις.

4.4 Αλγόριθμος Συστάσεων

Το σύστημα συστάσεων βασίστηκε στην λογική της προσέγγισης συνεργατικού φιλτραρίσματος (Collaborative Filtering) με τη μέθοδο της γειτνίασης (neighborhood) βασιζόμενη στον χρήστη (user-based) με κάποιες παραλλαγές. Οι διαφοροποιήσεις αυτές εντοπίζονται στους παράγοντες που λαμβάνονται υπόψη για την μέτρηση της ομοιότητας. Εδώ δεν χρησιμοποιούνται οι αξιολογήσεις των χρηστών για κάποιο προορισμό αλλά σύμφωνα με τη γνώμη τους για το ποιοι αυτοί θεωρούν ισχυρούς παράγοντες κατά την επιλογή ενός ταξιδιωτικού προορισμού.

4.4.1 Ομοιότητα χρηστών

Οι k -πλησιέστεροι γείτονες (k -NN) του χρήστη u , που δηλώνεται με $N(u)$, είναι οι k χρήστες με την υψηλότερη ομοιότητα w_{uv} με τον χρήστη u . Ωστόσο, μόνο οι n χρήστες που έχουν απαντήσεις τις ίδιες ερωτήσεις με τον χρήστη u θα ληφθούν υπόψη. Στα παρακάτω λοιπόν όπου αναφέρονται οι n χρήστες εννοούνται οι χρήστες οι οποίοι έχουν απαντήσει στις ίδιες ερωτήσεις με τον χρήστη u . Επιπλέον κάθε φορά που ένας νέος χρήστης εγγράφεται στο σύστημα, ο αλγόριθμος επαναυπολογίζει την ομοιότητα των χρηστών μεταξύ τους.

Ο αλγόριθμος υπολογίζει τρεις μετρικές ομοιότητας, $d1$, $d2$ και $d3$ για τον ενεργό χρήστη u σε σχέση με τους υπόλοιπους n χρήστες του συστήματος. Πιο συγκεκριμένα:

- Η μετρική $d1$ υπολογίζει την ομοιότητα των δημογραφικών χαρακτηριστικών των χρηστών και τη σημαντικότητα που δίνουν σε καθένα από αυτά.
- Η μετρική $d2$ υπολογίζει την ομοιότητα των παραγόντων που επηρεάζουν τις ταξιδιωτικές επιλογές των χρηστών.
- Η μετρική $d3$ υπολογίζει την ομοιότητα στις προτιμήσεις ταξιδιών των χρηστών και τη σημαντικότητα που δίνουν σε αυτές.

Η τελική ομοιότητα είναι το άθροισμα των τριών αυτών μετρικών και το εύρος τιμών είναι $[-3, 3]$ όπου η τιμή 3 σημαίνει πλήρη ομοιότητα των χρηστών, ενώ η τιμή -3 πλήρη αναντιστοιχία.

Ακολουθεί η αναλυτική περιγραφή των παραπάνω:

4.4.2 Υπολογισμός βάρους ομοιότητας

Τα βάρη ομοιότητας διαδραματίζουν διπλό ρόλο στις μεθόδους σύστασης που βασίζονται στη γειτονιά:

- 1) επιτρέπουν την επιλογή αξιόπιστων γειτόνων των οποίων οι αξιολογήσεις χρησιμοποιούνται στην πρόβλεψη και
- 2) παρέχουν τα μέσα για να δώσουν μεγαλύτερη ή μικρότερη σημασία σε αυτούς τους γείτονες στην πρόβλεψη.

Ο υπολογισμός των σταθμίσεων ομοιότητας είναι μια από τις πιο κρίσιμες πτυχές της οικοδόμησης ενός συστήματος συστάσεων γειτονιάς, καθώς μπορεί να έχει σημαντικό αντίκτυπο τόσο στην ακρίβεια όσο και στην απόδοσή του (Ricci, Rokach, Shapira, & Kantor, 2011, p. 124).

4.4.2.1 Ομοιότητα δημογραφικών χαρακτηριστικών

Παίρνοντας υπόψη τα δημογραφικά χαρακτηριστικά των χρηστών (φύλλο, ηλικία, εθνικότητα, θρήσκευμα, εκπαίδευση, επάγγελμα, εισόδημα, ΑμΕΑ) υπολογίζουμε το βαθμό ομοιότητας μεταξύ δύο χρηστών u και v . Για τον χρήστη u συγκρίνονται οι τιμές των πεδίων του πίνακα, $data_D$, με τους υπόλοιπους v χρήστες, και προκύπτει ο πίνακας $aq1_{(u,v)}$. Ίδια απάντηση τιμή 1, διαφορετική απάντηση τιμή 0.

Στη συνέχεια πολλαπλασιάζουμε την κάθε μια από αυτές τις τιμές με ένα βάρος το οποίο προσδιορίζεται από τις απαντήσεις του χρήστη u ($aq2$) οι οποίες είναι αποθηκευμένες στον πίνακα $pref_rec_dem$, για τις προτιμήσεις του, όσον αφορά τη σημαντικότητα που δίνει σε κάθε ένα από τα δημογραφικά χαρακτηριστικά. Οι τιμές κυμαίνονται από 1 έως 10 (1 – Καθόλου σημαντικό, 10 – Πολύ σημαντικό) και πολλαπλασιάζουμε τις τιμές του πίνακα $aq1$ με έναν αριθμό από 0,1 (Καθόλου σημαντικό) έως 1 (Πολύ σημαντικό), επομένως οι απαντήσεις $aq1$ υπολογίζονται ως εξής:

$$(1) \quad d_{(u,v)} = aq1_{(u,v)} * aq2_{(u)} / 10$$

Η ομοιότητα δημογραφικών χαρακτηριστικών λοιπόν υπολογίζεται ως εξής: αθροίζονται οι απαντήσεις $d_{(u,v)}$ και διαιρούμε με το σύνολο των 8 ερωτήσεων

$$(2) \quad d_{1(u,v)} = \frac{\text{Άθροισμα_απαντήσεων_δημογραφικών_ερωτήσεων}}{\text{Σύνολο_δημογραφικών_ερωτήσεων}}$$

όπου u ο χρήστης στόχος και v οι υπόλοιποι χρήστες που έχουν απαντήσει στις ίδιες ερωτήσεις με τον χρήστη u . Η τιμή της ομοιότητας κυμαίνεται από 0 έως 1, όπου 1 σημαίνει πλήρη ταύτιση των χρηστών και 0 πλήρη αναντιστοιχία.

Παράδειγμα:

Έστω ότι στο σύστημά μας έχουμε τέσσερις χρήστες, user1, user2, user3, user4 με τα παρακάτω δημογραφικά στοιχεία:

data_D	ΠΕΔΙΑ							
Χρήστες	gender	age	nationality	denomination	education	profession	income	disabilities
user1	Γ	41-50	ΕΛΛΗΝΙΚΗ	KANENA	MET	ΔΥ	1-1500	0
user2	Γ	41-50	ΕΛΛΗΝΙΚΗ	KANENA	MET	ΔΥ	1-1500	0
user3	A	41-50	ΓΕΡΜΑΝΙΚΗ	KANENA	MET	ΔΥ	1-1500	0
user4	A	>61	ΕΛΛΗΝΙΚΗ	KANENA	Δ	ΕΕ	1501-3000	0

Πίνακας 1: Απαντήσεις δημογραφικών ερωτήσεων

Ο πίνακας aq1 που προκύπτει για τον χρήστη user1 σε σύγκριση με τους υπόλοιπους, θα είναι:

aq1	Πεδία							
Χρήστες	gender	age	nationality	denomination	education	profession	income	disabilities
user2	1	1	1	1	1	1	1	1
user3	0	1	0	1	1	1	1	1
user4	0	0	1	1	0	0	0	1

Πίνακας 2: Σύγκριση απαντήσεων χρήστη user1 με τις απαντήσεις των χρηστών 2, 3, 4.

Οι τιμές στον πίνακα pref_rec_dem για τον χρήστη user1 έστω ότι είναι οι εξής :

pref_rec_de m	Πεδία						
ΧΡΗΣΤΕΣ	rec_gend er	rec_ag e	rec_nati on	rec_deno m	rec_ed u	rec_pr of	rec_inco m
user 1	1	10	9	5	8	6	7

Πίνακας 3: Απαντήσεις του χρήστη user1 στις ερωτήσεις για συστάσεις (δημογραφικά στοιχεία)

Οι απαντήσεις στον aq1 εφαρμόζοντας την εξίσωση (1) μετατρέπονται ως εξής:

Χρήστ ες	Πεδία							
	gen der	ag e	nationali ty	denominati on	educati on	professi on	inco me	disabiliti es
user 2	0,1	1	0,9	0,5	0,8	0,6	0,7	1
user 3	0	1	0	0,5	0,8	0,6	0,7	1
user 4	0	0	0,9	0,5	0	0	0	1

Πίνακας 4: Απαντήσεις του χρήστη user1 σε σχέση με τις απαντήσεις των άλλων χρηστών πολλαπλασιασμένες με το βάρος σημαντικότητας

Η τελική ομοιότητα του χρήστη user1 με τους υπόλοιπους τρεις χρήστες αφού εφαρμοστεί η εξίσωση (2) θα είναι:

ΧΡΗΣΤΕΣ	user1	user2	user3	user4
user1	1	0,62	0,45	0,26

Πίνακας 5: Ομοιότητα χρήστη 1 με τους υπόλοιπους

Ο χρήστης 1 με βάση τα δημογραφικά χαρακτηριστικά είναι πιο κοντά στον χρήστη 2.

4.4.2.2 Ομοιότητα παραγόντων που επηρεάζουν τον χρήστη για την επιλογή ταξιδιού

Ένα δημοφιλές μέτρο που συγκρίνει αξιολογήσεις όπου οι επιδράσεις του μέσου όρου και της διακύμανσης έχουν αφαιρεθεί είναι η ομοιότητα **Pearson Correlation**.

$$(3) \quad d2_{(u,v)} = \frac{\sum_{i \in I_{3_{uv}}} (q3_{ui} - \bar{q3}_u)(q3_{vi} - \bar{q3}_v)}{\sqrt{\sum_{i \in I_{3_{uv}}} (q3_{ui} - \bar{q3}_u)^2 \sum_{i \in I_{3_{uv}}} (q3_{vi} - \bar{q3}_v)^2}}$$

Στο σύστημά μας είναι ένα μέτρο συσχέτισης δύο χρηστών u και v ως προς τις απαντήσεις που δίνουν στη δεύτερη οθόνη (παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές), είναι οι τιμές που αποθηκεύονται στον πίνακα `data_pref`. Η μετρική $d2_{(u,v)}$ είναι η συσχέτιση των χρηστών u και v , $q3_{ui}$ είναι η απάντηση του χρήστη u στην ερώτηση i , $q3_{vi}$ είναι η απάντηση του χρήστη v στην ερώτηση i , $\overline{q3_u}$ η μέση τιμή των απαντήσεων του χρήστη u , ενώ $\overline{q3_v}$ η μέση τιμή των απαντήσεων του χρήστη v . Σημειώνεται ότι $i \in I3_{uv}$ όπου το $I3_{uv}$ είναι οι ερωτήσεις της δεύτερης οθόνης στις οποίες έχουν απαντήσει και οι δύο χρήστες u, v .

Οι τιμές κυμαίνονται από -1 έως 1 και αντιπροσωπεύουν τη δύναμη της συσχέτισης. Η τιμή -1 σημαίνει ότι οι χρήστες έχουν πλήρη αναντιστοιχία (έχουν απαντήσει πολύ διαφορετικά), 1 ότι οι χρήστες έχουν πλήρη ταύτιση – (έχουν απαντήσει ακριβώς ίδια).

Παράδειγμα

Οι τιμές του πίνακα `data_pref` για τους τέσσερις χρήστες:

data_p ref	Πεδία										
Χρήστ ες	na tu re	monu ments	cult ure	soci ety	co st	trans port	acc ess	saf ety	r_tr_ne twork	facilitie s_servic es	clim e
user1	6	9	8	3	5	6	5	7	5	5	2
user2	8	10	10	2	7	3	2	8	6	6	1
user3	8	8	8	8	8	7	6	10	6	7	6
user4	1	1	1	1	2	3	1	5	4	6	8

Πίνακας 6: Απαντήσεις των χρηστών 1, 2, 3, 4 στις ερωτήσεις της δεύτερης οθόνης

Η ομοιότητα του χρήστη 1 με τους υπόλοιπους τρεις χρήστες εφαρμόζοντας την εξίσωση (3) θα είναι:

ΧΡΗΣΤΕΣ	user1	user2	user3	user4
user1	1	0,85	0,50	-0,47

Πίνακας 7: Ομοιότητα χρήστη 1 με τους υπόλοιπους με βάση τους παράγοντες που τους επηρεάζουν στην επιλογή

4.4.2.3 Ομοιότητα προτιμήσεων των χρηστών

Για κάθε ένα από τα πέντε κριτήρια (σκοπός, μεταφορικό μέσο, παρέα, προορισμός,

διάρκεια) της τελευταίας οθόνης του συστήματος υπολογίζεται η ομοιότητα των χρηστών. Ο τρόπος υπολογισμού είναι ο εξής: η κάθε απάντηση που δηλώνει τη συχνότητα πολλαπλασιάζεται με τη σημαντικότητα που δίνει ο χρήστης (ερωτήσεις δεύτερης οθόνης) δια 10.

$$(4) \quad aq2 = \text{Συχνότητα} * \text{Σημαντικότητα} / 10$$

Παράδειγμα:

Θα πάρουμε τις προτιμήσεις για το μεταφορικό μέσο. Η σημαντικότητα που δίνει ο χρήστης για την επιλογή του μεταφορικού μέσου (τιμές 1 έως 10) μαζί με τη συχνότητα που επιλέγει μεταφορικό μέσο (τιμές 0 έως 4) χρησιμοποιούνται ως εξής:

Έστω ότι ο χρήστης 1 δίνει σημαντικότητα 8, και οι προτιμήσεις του φαίνονται στον παρακάτω πίνακα:

	Μεταφορικό μέσο			
user1	αεροπλάνο	αυτοκίνητο	Λεωφορείο-τρένο	πλοίο
ΜΕΤΑΦΟΡΙΚΟ ΜΕΣΟ	πολύ συχνά-4	μερικές φορές-2	σπάνια-1	ποτε-0
ΣΗΜΑΝΤΙΚΟΤΗΤΑ	8			
Υπολογισμός aq2	$4*8/10=3,2$	$2*8/10=1,6$	$1*8/10=0,8$	0

Πίνακας 8: Υπολογισμός απαντήσεων χρήστη user1 για μεταφορικό μέσο

Οι τιμές αυτές υπολογίζονται για όλους τους χρήστες όποτε έχουμε:

	Μεταφορικό μέσο			
	αεροπλάνο	αυτοκίνητο	Λεωφορείο-τρένο	πλοίο
user1	3,2	1,6	0,8	0
user2	2	4	2	1
user3	3,6	3,6	0,9	0,9

Πίνακας 9: Απαντήσεις χρηστών μετά την εφαρμογή της εξίσωσης 4

Χρησιμοποιούμε τον υπολογισμό ομοιότητας Pearson Correlation και προκύπτει η ομοιότητα του user1 με τους υπόλοιπους χρήστες σε ότι αφορά την επιλογή του μεταφορικού μέσου και τη σημαντικότητα που δίνει ο καθένας ως εξής:

sim_trans	user1	user2	user3
user 1	1	0,33	0,87

Πίνακας 10: Ομοιότητα χρήστη 1 με τους χρήστες 2 και 3

Η ίδια διαδικασία ακολουθείται και για τα πέντε κριτήρια, επομένως θα έχουμε πέντε τιμές ομοιότητας: sim_{prop} (ομοιότητα ως προς τον σκοπό), sim_{trans} (ομοιότητα ως προς το μεταφορικό μέσο), sim_{comp} (ομοιότητα ως προς την παρέα), sim_{dest} (ομοιότητα ως προς τον προορισμό), sim_{dur} (ομοιότητα ως προς την διάρκεια).

Η ομοιότητα των προτιμήσεων προκύπτει από τον μέσο όρο των παραπάνω τιμών:

$$(5) \quad d_{3uv} = (sim_{prop_{uv}} + sim_{trans_{uv}} + sim_{comp_{uv}} + sim_{dest_{uv}} + sim_{dur_{uv}}) / 5.$$

4.4.2.4 Τελική Ομοιότητα χρηστών

Τέλος η ομοιότητα του χρήστη u από τους χρήστες v θα είναι το άθροισμα των τριών μετρικών $d1$ (εξίσωση 2), $d2$ (εξίσωση 3) και $d3$ (εξίσωση 5), δηλ.

$$(6) \quad D_{uv} = d1_{uv} + d2_{uv} + d3_{uv}.$$

Το σύστημα θα υπολογίζει για κάθε χρήστη u την ομοιότητα με τους υπόλοιπους v χρήστες οι οποίοι έχουν απαντήσει στις ίδιες ερωτήσεις με τον χρήστη u .

Στο παράδειγμά μας η τελική ομοιότητα του χρήστη 1 με τους υπόλοιπους χρήστες θα είναι:

ΧΡΗΣΤΕΣ	user1	user2	user3	user4
user1	3	1,77	1,95	-0,21

Πίνακας 11: Τελική ομοιότητα του χρήστη 1 με τους χρήστες 2, 3 και 4.

Όσο πιο κοντά στην τιμή 3 τόσο πιο μεγάλη είναι η ομοιότητα των χρηστών.

4.5 Επιλογή γειτόνων - Neighborhood Selection

Φιλτράρισμα Top-N

Για κάθε χρήστη, διατηρείται μια λίστα των πλησιέστερων γειτόνων και η αντίστοιχη τιμή ομοιότητας. Στο σύστημά μας την πρώτη φορά που γίνονται συστάσεις στον χρήστη u , επιλέγονται οι τρεις πιο κοντινοί του γείτονες, αυτοί με τη μεγαλύτερη τιμή D_{uv} (εξίσωση 6).

Στη συνέχεια συστήνονται στον χρήστη οι προορισμοί των τριών πιο κοντινών του γειτόνων ($k_u=3$). Σε αυτή τη φάση το σύστημα ζητάει από τον χρήστη να αξιολογήσει αυτές τις προτάσεις σε μια κλίμακα από 1 έως 5 (1-Δεν με ενδιαφέρει, 5-Με ενδιαφέρει πολύ). Η κάθε απάντηση έχει και μια αριθμητική τιμή (1 έως 5) αντίστοιχη με την κλίμακα απαντήσεων. Ο αλγόριθμος υπολογίζει το μέσο όρο των απαντήσεων και αν ο αριθμός αυτός είναι μικρότερος του 3, σημαίνει ότι δεν υπάρχει μεγάλο ενδιαφέρον του χρήστη u για αυτούς τους προορισμούς. Τότε το σύστημα συνεχίζει και προτείνει προορισμούς από τους επόμενους δύο πιο κοντινούς γείτονες ($k_u=5$) που βρίσκονται στη λίστα. Επαναλαμβάνεται η διαδικασία αξιολόγησης αυτών των προτάσεων και ο υπολογισμός του μέσου όρου και από τους 5 γείτονες. Κάθε φορά το σύστημα αυξάνει κατά δύο το k_u και όταν βρεθεί μια τιμή πάνω από 3 μονάδες για τον μέσο όρο, τότε το σύστημα σταματάει και καταχωρεί τον αριθμό k_u πλησιέστερων γειτόνων για τον χρήστη u . Την επόμενη φορά που θα κάνει συστάσεις για αυτόν τον χρήστη θα προτείνει προορισμούς από τους k_u πιο κοντινούς γείτονες.

4.6 Προσαρμογή προφίλ

Οι προτιμήσεις και τα ενδιαφέροντα των χρηστών αλλάζουν με την πάροδο του χρόνου. Από το σύστημα ο χρήστης έχει τη δυνατότητα να αλλάξει χειροκίνητα τα δημογραφικά του στοιχεία, ή τις προτιμήσεις του στις υπόλοιπες ερωτήσεις του συστήματος. Αν πραγματοποιηθεί οποιαδήποτε αλλαγή στα στοιχεία του ή/και στις προτιμήσεις του τότε το σύστημα επαναυπολογίζει την ομοιότητά του με τους υπόλοιπους n χρήστες του συστήματος και συστήνει προορισμούς αρχικά από τους 3 πιο κοντινούς και αναλόγως τροποποιεί τον αριθμό k_u πλησιέστερων γειτόνων σύμφωνα με την διαδικασία που αναφέρθηκε στην προηγούμενη παράγραφο.

Εκτός από τον εκ νέου υπολογισμό των k πλησιέστερων γειτόνων οι προορισμοί που συστήνονται εμφανίζονται στον χρήστη με φθίνουσα ταξινόμηση, δηλ. από τον πιο πρόσφατο χρονικά έως τον πιο παλιό. Και αυτό διότι για έναν χρήστη ο οποίος επισκέφθηκε έναν προορισμό πριν από μια πενταετία είναι πολύ πιθανόν οι προτιμήσεις του να έχουν αλλάξει και τα ταξίδια που πραγματοποίησε αρχικά τώρα μπορεί να μην τον εκφράζουν.

4.7 Χρήστες που δεν απαντούν στις ερωτήσεις προτίμησης

Σε περίπτωση που υπάρχουν χρήστες οι οποίοι δεν απαντούν στις ερωτήσεις προτιμήσεων, το σύστημα βασίζεται στα δημογραφικά τους στοιχεία (αποθηκευμένα στον πίνακα data_D) και υλοποιεί δημογραφικό φιλτράρισμα. Ταξινομεί τους χρήστες σε κλάσεις σύμφωνα με τα δημογραφικά τους χαρακτηριστικά και προτείνει στον χρήστη προορισμούς που έχουν επισκεφθεί άλλοι χρήστες της ίδιας τάξης.

Σε περίπτωση που χρήστες δίνουν ελλιπείς απαντήσεις στις ερωτήσεις ταξιδιωτικών προτιμήσεων, επιλογών κλπ. το σύστημα τους αποδίδει τον μέσο όρο των απαντήσεων των υπόλοιπων χρηστών του συστήματος.

4.8 Δημιουργία προφίλ από αναζητήσεις των χρηστών

Επιπλέον το σύστημα διατηρεί μια λίστα με τις αναζητήσεις ταξιδιωτικών προορισμών του χρήστη. Τις κατατάσσει σε φθίνουσα σειρά σύμφωνα με τον αριθμό των αναζητήσεων και πραγματοποιεί συστάσεις για τους πρώτους τρεις προορισμούς που αναζήτησε, εμφανίζοντας στοιχεία για τον κάθε προορισμό από τον πίνακα travel_history, όπως για ποιον σκοπό ταξίδεψαν οι άλλοι χρήστες, πόση ήταν η διάρκεια ταξιδιού τους σε αυτόν τον προορισμό κ.λ.π.

4.9 Συστάσεις

Τέλος οι συστάσεις που γίνονται στους χρήστες πέρα από τους προορισμούς συνοδεύονται και από την αιτιολόγησή τους. Παράδειγμα: σας συστήνονται οι α, β, γ, προορισμοί γιατί είναι προορισμοί που επισκέφθηκαν οι πιο κοντινοί προς εσάς χρήστες u1, u2, u3 ή συστήνονται οι α, β, γ, προορισμοί γιατί οι χρήστες που επισκέφθηκαν αυτούς τους προορισμούς η διάρκεια του ταξιδιού τους ήταν 1 με 2 διανυκτερεύσεις.

Ο χρήστης έχει τη δυνατότητα να δει τα προφίλ των χρηστών που είναι οι πιο κοντινοί γείτονές του και να επιλέξει αν θέλει να παραμείνουν στη λίστα των πιο κοντινών του ή να διαγραφούν. Το σύστημα δεν θα τον συμπεριλάβει στη λίστα των πλησιέστερων

γειτόνων του εκτός και αν συμβεί κάποια αλλαγή στα στοιχεία ενός εκ των δύο χρηστών (δημογραφικά ή προτιμήσεις).

Κεφάλαιο 5

Πόσο πρόθυμοι είναι οι χρήστες να χρησιμοποιήσουν το προτεινόμενο σύστημα συστάσεων

Στο Κεφάλαιο 4 έγινε η περιγραφή σχεδιασμού ενός ταξιδιωτικού συστήματος συστάσεων. Στο παρόν κεφάλαιο παρουσιάζεται η έρευνα η οποία έχει ως σκοπό την διερεύνηση των απόψεων του κοινού σχετικά με το ταξιδιωτικό σύστημα συστάσεων που περιγράφηκε ώστε να δούμε κατά πόσο είναι θετικοί οι χρήστες να χρησιμοποιήσουν ένα τέτοιο σύστημα. Ένα θετικό αποτέλεσμα θα μπορούσε να σημαίνει την υλοποίηση ενός τέτοιου συστήματος. Πιο συγκεκριμένα, το σύστημα για να παράγει ακριβείς συστάσεις χρειάζεται αρκετά προσωπικά στοιχεία (δημογραφικά, ταξιδιωτικές προτιμήσεις και επιλογές) των χρηστών του. Διερευνάται λοιπόν κατά πόσο είναι διατεθειμένοι οι χρήστες να αφιερώσουν χρόνο για τη δημιουργία του προφίλ τους και να δώσουν τα προσωπικά τους στοιχεία. Είναι ένα κρίσιμο ζήτημα γιατί, αν οι χρήστες είναι αρνητικοί τότε το σύστημα δεν θα έχει κάποια χρησιμότητα.

Αρχικά παρουσιάζεται η έρευνα, έπειτα αναλύονται τα αποτελέσματα με τη βοήθεια διαγραμμάτων και τέλος ερμηνεύονται τα αποτελέσματα.

5.1 Περιγραφή έρευνας - Ερευνητικά ερωτήματα

Το θέμα της έρευνας ήταν κατά πόσο οι χρήστες είναι διατεθειμένοι να συμπληρώσουν το προφίλ τους και να απαντήσουν σε ερωτήσεις σχετικά με τις ταξιδιωτικές τους επιλογές και προτιμήσεις στο σύστημα ώστε να λάβουν συστάσεις για έναν ταξιδιωτικό προορισμό. Τα βασικά ερευνητικά ερωτήματα είναι τα εξής:

- Θέλουνε οι χρήστες να λαμβάνουν συστάσεις από ένα σύστημα;

- Είναι διατεθειμένοι να συμπληρώσουν το προφίλ τους στο σύστημα (δημογραφικά στοιχεία, ταξιδιωτικές προτιμήσεις και επιλογές);
- Για την παραγωγή εξατομικευμένων συστάσεων θα ήθελαν να δηλώνουν από ποιες κατηγορίες χρηστών θα προέρχονται αυτές οι συστάσεις και να τους συγκρίνει το σύστημα με τους υπόλοιπους χρήστες;

Η έρευνα διεξήχθη ηλεκτρονικά μέσω ερωτηματολογίου (Παράρτημα Δ) με το δωρεάν εργαλείο Φόρμες Google. Το ερωτηματολόγιο στάλθηκε με email, κοινοποιήθηκε μέσω facebook και messenger σε συναδέλφους, φίλους, γνωστούς και σε ταξιδιωτικά γραφεία. Καλέσαμε το κοινό να συμμετάσχει στην έρευνα με το εξής μήνυμα:

Ονομάζομαι Μόκα Ιωάννα, είμαι εκπ/κός ΠΕ86 (Πληροφορικής), και στα πλαίσια της μεταπτυχιακής μου διατριβής διερευνώ τις απόψεις του κοινού σχετικά με τη χρήση ενός ταξιδιωτικού συστήματος συστάσεων (travel recommendation system).

Είναι σημαντικό να συλλέξω δεδομένα και θα ήταν πολύτιμη βοήθεια να αφιερώσετε 3 λεπτά για τη συμπλήρωση του ερωτηματολογίου στον ακόλουθο υπερσύνδεσμο: <https://goo.gl/forms/j5npgZPwJ4oVp8jt1>.

Σας ευχαριστώ!

Στην αρχή του ερωτηματολογίου υπήρχε το εξής ενημερωτικό σημείωμα για τους συμμετέχοντες:

Ενημερωτικό Σημείωμα προς Συμμετέχοντες/ουσες

Στα πλαίσια της μεταπτυχιακής μου διατριβής, περιγράφεται ένα ταξιδιωτικό σύστημα συστάσεων το οποίο προτείνει στους χρήστες του ταξιδιωτικούς προορισμούς. Το σύστημα αυτό χρειάζεται τα δημογραφικά στοιχεία των χρηστών του (π.χ. ηλικία, φύλο, εκπαίδευση, εισόδημα κ.α.) καθώς και τις προτιμήσεις τους σε ό,τι αφορά τις ταξιδιωτικές τους επιλογές (φύση, μνημεία, κλίμα, κόστος κ.α.) αλλά και από ποιους χρήστες θα θέλανε να προέρχονται οι συστάσεις (π.χ. από χρήστες με ίδιο επίπεδο εκπαίδευσης, ίδιο θρήσκευμα, με ίδιες προτιμήσεις ως προς τον σκοπό του ταξιδιού κ.α.). Διατηρεί ιστορικό των ταξιδιών που έχει πραγματοποιήσει ο χρήστης καθώς και ιστορικό αναζητήσεων του χρήστη.

Διασφάλιση της εμπιστευτικότητας

Το ερωτηματολόγιο αποτελεί μία ανώνυμη διαδικασία κατά την οποία όλες οι πληροφορίες που θα συλλεχθούν θα παραμείνουν αυστηρά εμπιστευτικές. Τα μοναδικά προσωπικά στοιχεία που ζητούνται αφορούν κυρίως δημογραφικά στοιχεία όπως το φύλο, η ηλικία, η εκπαίδευση, το επάγγελμα.

Σας ευχαριστώ για τον χρόνο σας.

Μόκα Ιωάννα

Εκπ/κός ΠΕ 86

Μεταπτυχιακή φοιτήτρια στο Ανοικτό Πανεπιστήμιο Κύπρου στο πρόγραμμα σπουδών "Κοινωνικά Πληροφοριακά Συστήματα".

Τίτλος μεταπτυχιακής διατριβής: "Σύστημα συστάσεων ταξιδιωτικών προορισμών: προσέγγιση συνεργατικού φιλτραρίσματος με βάση τα προφίλ και τις προτιμήσεις των χρηστών".

email: miekbg@gmail.com

Είμαι στη διάθεση σας, οπότε μη διστάσετε να κάνετε ερωτήσεις ή να εκφράσετε τον προβληματισμό σας σχετικά με το ερωτηματολόγιο.

Αρχικά συλλέγονται δημογραφικά στοιχεία των συμμετεχόντων όπως φύλο, ηλικία, επίπεδο εκπαίδευσης και επάγγελμα. Θέλουμε να δούμε ποιοι είναι διατεθειμένοι να χρησιμοποιήσουν ένα τέτοιο σύστημα και να διαθέσουν πληροφορίες στο σύστημα για αυτό.

Έπειτα ρωτάμε τη συχνότητα με την οποία ταξιδεύουν στο εσωτερικό και στο εξωτερικό. Υποθέτουμε ότι τα άτομα τα οποία ταξιδεύουν πιο συχνά θα θέλανε μια τέτοια υπηρεσία και για αυτό θα ήταν διατεθειμένοι να δώσουν τα στοιχεία τους στο σύστημα.

Οι επόμενες ερωτήσεις είναι από ποιον θα θέλανε να προέρχονται οι συστάσεις, από ένα σύστημα, από έναν άνθρωπο ή και από τα δύο.

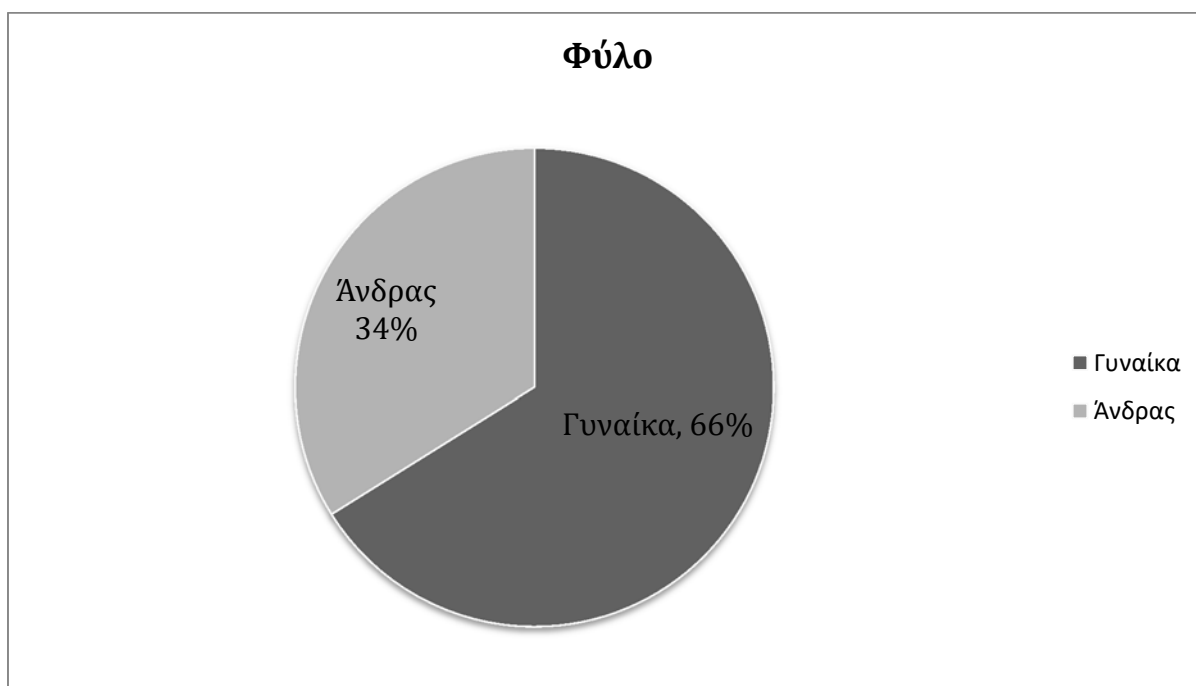
Τέλος, οι επόμενες ερωτήσεις, διερευνούν κατά πόσο είναι διατεθειμένοι να δώσουν τα δημογραφικά τους στοιχεία, τις ταξιδιωτικές τους προτιμήσεις και επιλογές ώστε να παράγει το σύστημα συστάσεις, καθώς και αν θα θέλανε να μαθαίνει το σύστημα για αυτούς μέσω των αναζητήσεων και των προηγούμενων ταξιδιών που έχουν πραγματοποιήσει. Και τέλος, αν θα θέλανε να τους συγκρίνει με άλλους χρήστες ώστε να βρει άλλους όμοιους με αυτούς χρήστες του συστήματος. Με αυτές τις ερωτήσεις θέλουμε να δούμε ποιοι χρήστες είναι διατεθειμένοι να δώσουν πληροφορίες για αυτούς και τι είδους πληροφορίες είναι αυτές που θα δώσουν στο σύστημα.

5.2 Παρουσίαση και ερμηνεία των αποτελεσμάτων

Στην έρευνα συμμετείχαν 180 άτομα. Παρακάτω γίνεται διαγραμματική παρουσίαση των αποτελεσμάτων.

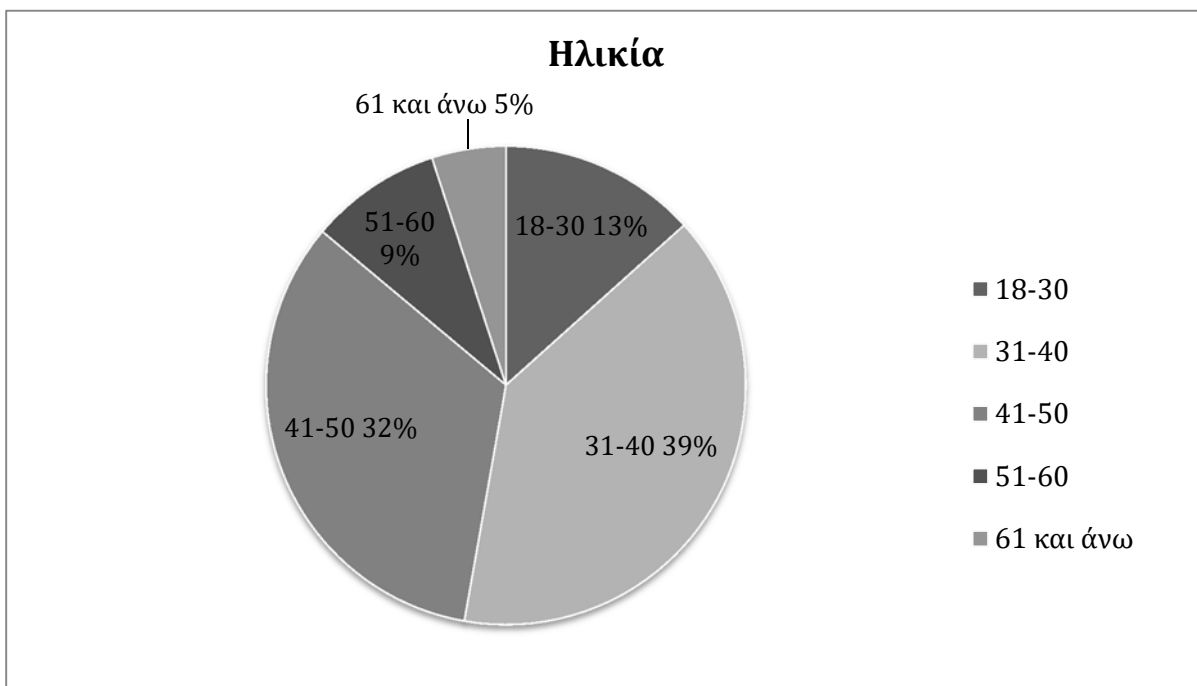
5.2.1 Δημογραφικά

Από τα 180 άτομα το μεγαλύτερο ποσοστό, 66%, είναι γυναίκες, ενώ το 34% είναι άντρες όπως φαίνεται και στο (Γράφημα 1: Φύλο



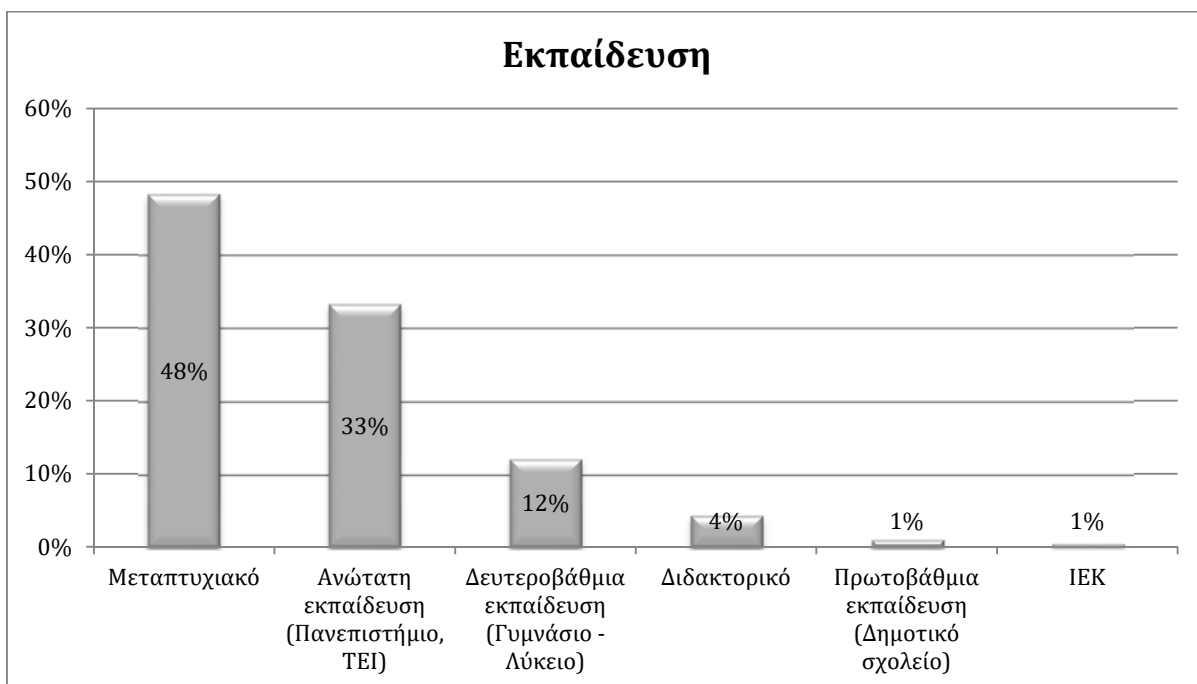
Γράφημα 1: Φύλο

Το μεγαλύτερο ποσοστό (39%) είναι μεταξύ 31 έως 40 ετών, το 32% από 41 έως 50 ετών, το 13% από 18 έως 30ετών, το 9% από 51 έως 60 ετών και τέλος το 5% από 61 ετών και πάνω (Γράφημα 2: Ηλικία).



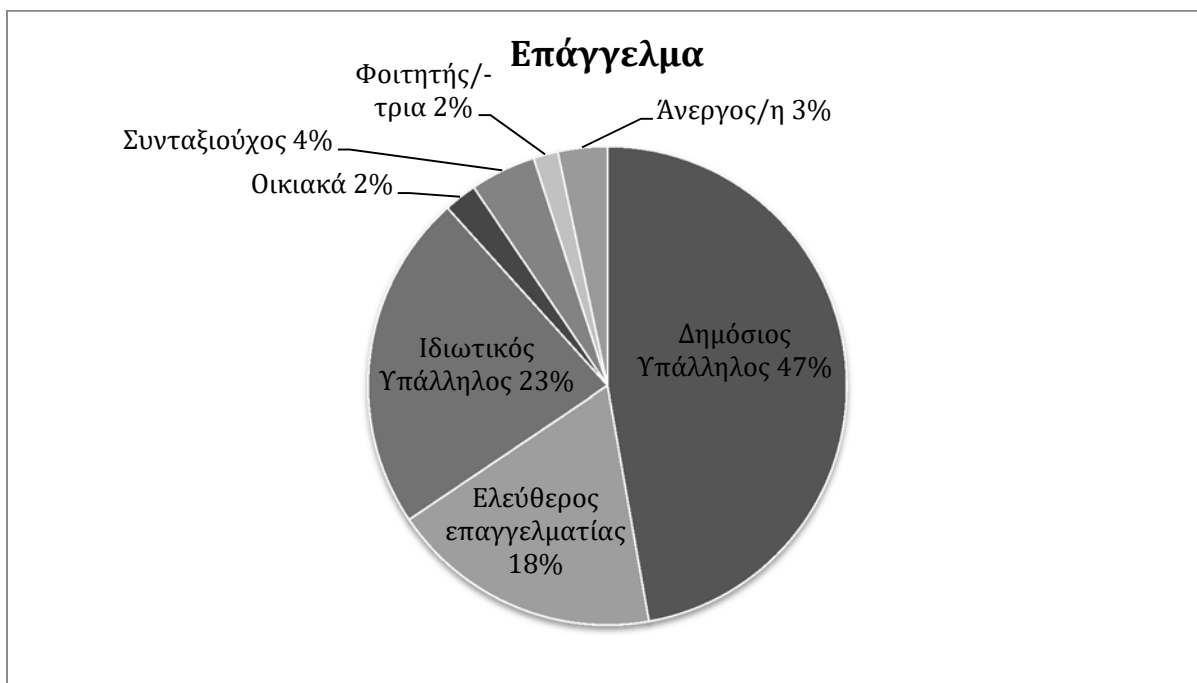
Γράφημα 2: Ηλικία

Ότι αφορά στο επίπεδο εκπαίδευσης το 48% των συμμετεχόντων κατέχουν μεταπτυχιακό δίπλωμα, το 33% έχουν τελειώσει την Ανώτατη εκπαίδευση, το 12% την Δευτεροβάθμια εκπαίδευση, το 4% κατέχει Διδακτορικό δίπλωμα και 1% τελείωσε την Πρωτοβάθμια εκπαίδευση και το 1% ΙΕΚ (Γράφημα 3: Εκπαίδευση).



Γράφημα 3: Εκπαίδευση

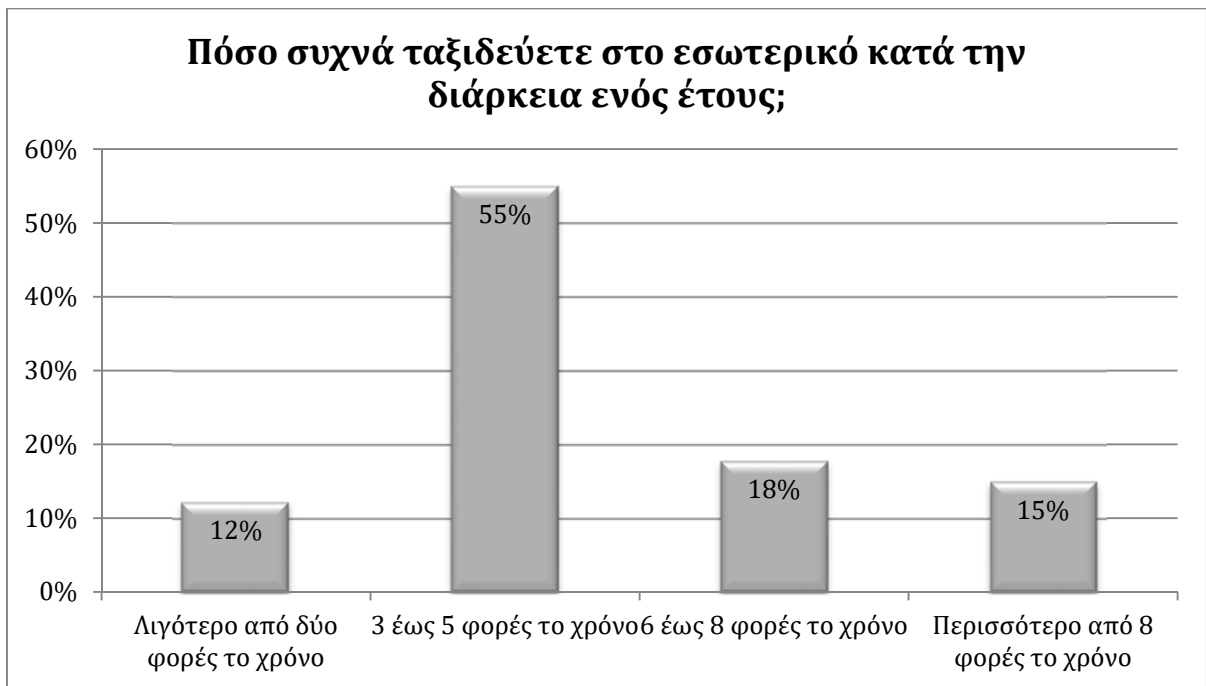
Τέλος, ως προς το επάγγελμα, το μεγαλύτερο ποσοστό, 48%, είναι Δημόσιοι υπάλληλοι, το 23% Ιδιωτικοί υπάλληλοι, το 18% Ελεύθεροι επαγγελματίες, το 5% συνταξιούχοι, το 3% άνεργοι, το 2% δηλώνουν οικιακά και το 2% είναι φοιτητές (Γράφημα 4: Επάγγελμα).



Γράφημα 4: Επάγγελμα

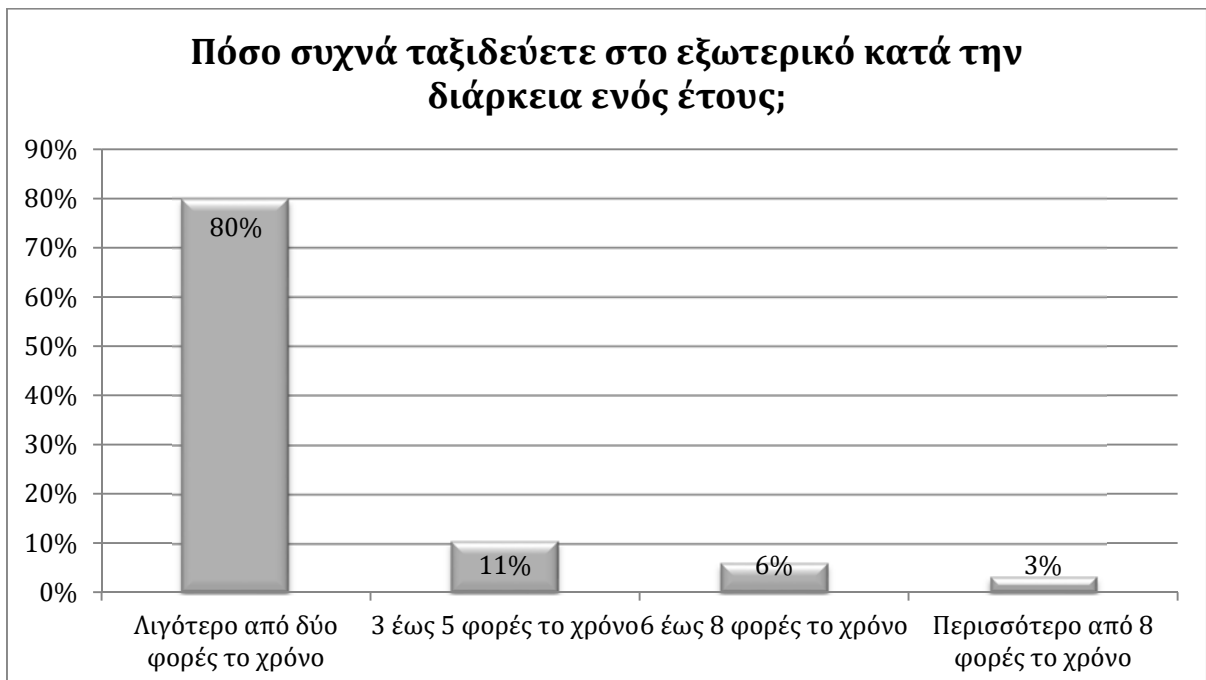
5.2.2 Συχνότητα ταξιδιών στο εσωτερικό/εξωτερικό

Στο εσωτερικό, πάνω από τους μισούς ερωτηθέντες (55%), ταξιδεύουν από 3 έως 5 φορές το χρόνο, το 18% 6 έως 8 φορές, το 15% περισσότερο από 8 φορές και το 12% λιγότερο από δύο φορές το χρόνο (Γράφημα 5).



Γράφημα 5: Συχνότητα ταξιδιών στο εσωτερικό

Στο εξωτερικό, το μεγαλύτερο ποσοστό 80%, ταξιδεύει λιγότερο από 2 φορές το χρόνο, το 11% 3 έως 5 φορές, το 6% 6 έως 8 φορές και το 3% περισσότερο από 8 φορές το χρόνο (Γράφημα 6).



Γράφημα 6: Συχνότητα ταξιδιών στο εξωτερικό

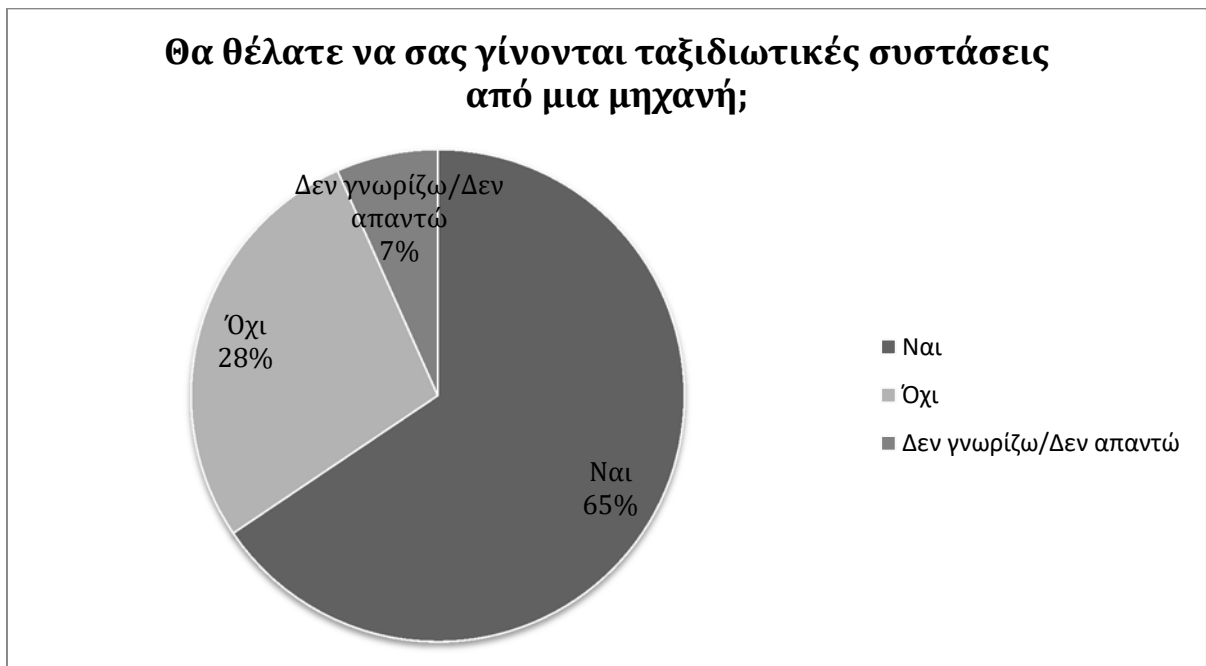
5.2.3 Προέλευση συστάσεων

Στην ερώτηση «Θα θέλατε να σας γίνονται ταξιδιωτικές συστάσεις από κάποιον άνθρωπο;» το 83% δηλώνει ότι ναι θα ήθελε, το 14% δηλώνει όχι και το 3% απάντησε ότι Δεν γνωρίζει/ Δεν απαντά (Γράφημα 7).



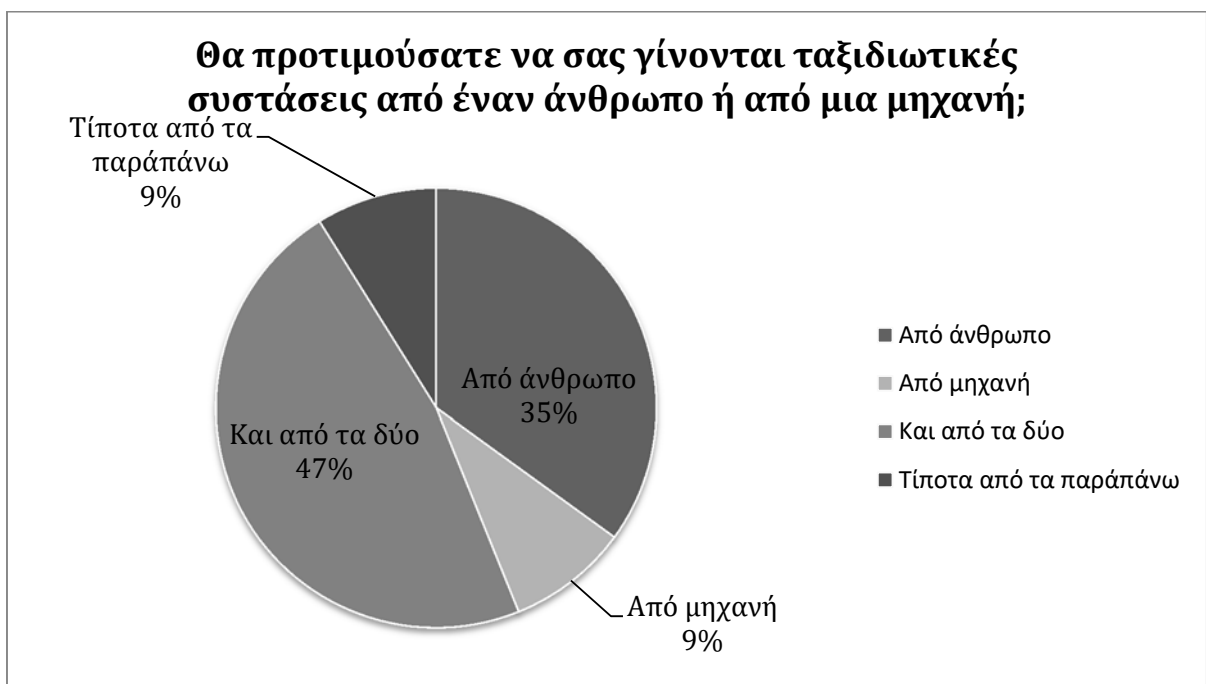
Γράφημα 7: Συστάσεις από άνθρωπο

Στην επόμενη ερώτηση, «Θα θέλατε να σας γίνονται ταξιδιωτικές συστάσεις από μια μηχανή;» το μεγαλύτερο ποσοστό, 65%, δηλώνει πως ναι, το 28% δηλώνει όχι ενώ το 7% Δεν γνωρίζει/Δεν απαντά (Γράφημα 8).



Γράφημα 8: Συστάσεις από μηχανή

Τέλος στην ερώτηση «Θα προτιμούσατε να σας γίνονται ταξιδιωτικές συστάσεις από έναν άνθρωπο ή από μια μηχανή;» το 47%, δηλώνει πως θα ήθελε και από τα δύο, το 35% από άνθρωπο, το 9% από μηχανή, και το 9% δηλώνει τίποτα από τα παραπάνω (Γράφημα 9).



Γράφημα 9: Συστάσεις από άνθρωπο ή από μηχανή

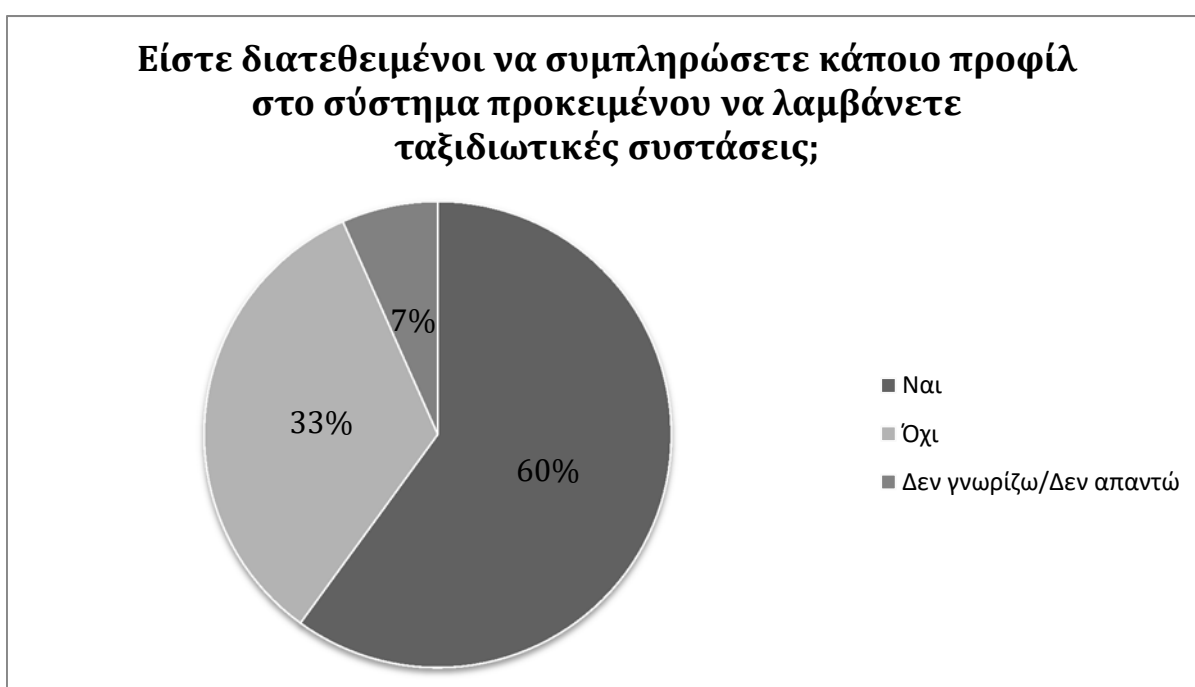
Παρατηρούμε ότι το 83% των ερωτηθέντων δηλώνει ότι θα ήθελε οι συστάσεις να γίνονται από άνθρωπο, το ποσοστό αυτό μειώνεται στην ερώτηση αν θα ήθελε οι συστάσεις να γίνονται από μηχανή (65%) αν και παραμένει υψηλό, ενώ σχεδόν οι μισοί, το 47% θα ήθελε οι συστάσεις να του γίνονται και από τα δύο. Να σημειωθεί ότι το 35% θα ήθελε μόνο από άνθρωπο ενώ το 9% μόνο από μηχανή.

5.2.4 Συμπλήρωση προφίλ, ταξιδιωτικών προτιμήσεων και επιλογών

Το επόμενο σκέλος της έρευνας αφορά στο κατά πόσο οι χρήστες είναι διατεθειμένοι να καταχωρούν τα προσωπικά τους στοιχεία, τις επιλογές και τις προτιμήσεις τους στο σύστημα ώστε να τους παρέχονται συστάσεις.

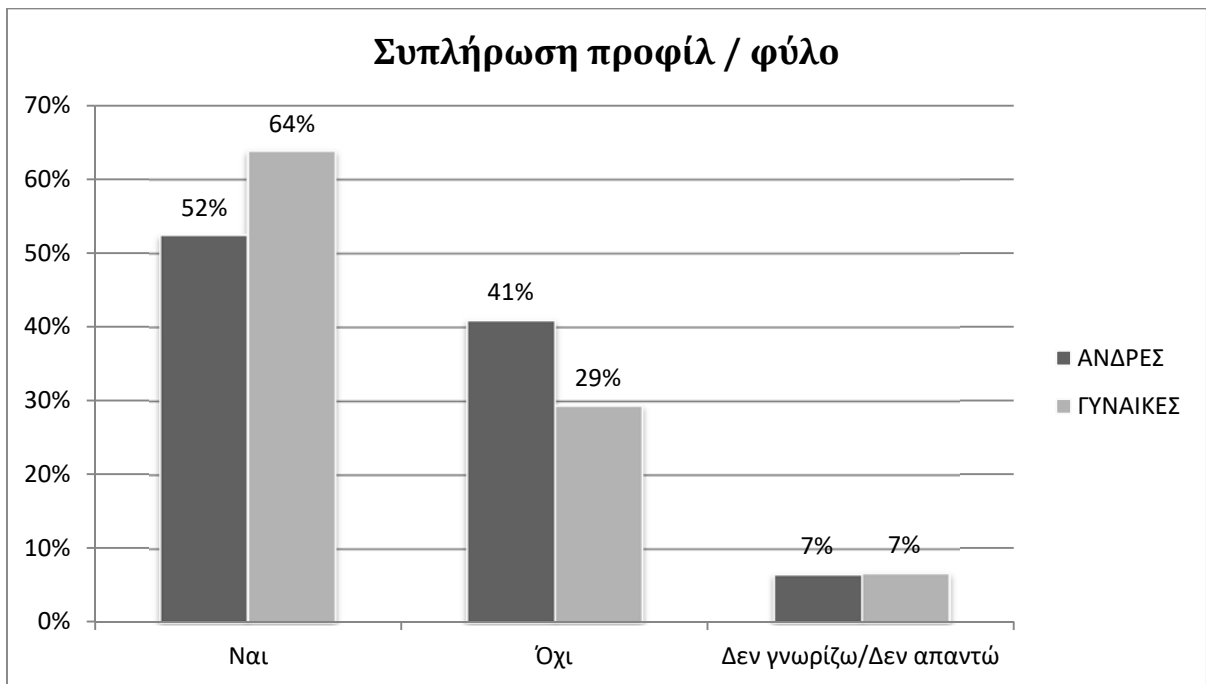
5.2.4.1 Συμπλήρωση προφίλ

Το 60% των ερωτηθέντων θα συμπλήρωνε το προφίλ του στο σύστημα, το 33% δηλώνει πως δεν θα το έκανε, ενώ το 7% Δεν γνωρίζει/ Δεν απαντά (Γράφημα 10).



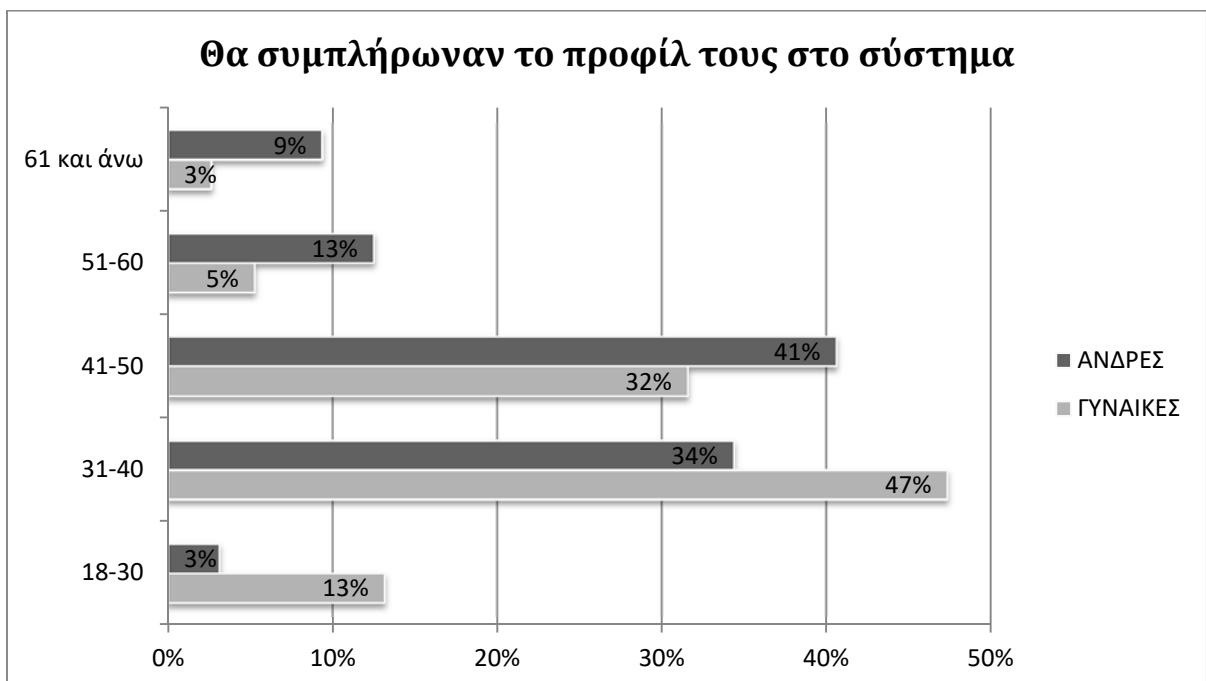
Γράφημα 10: Συμπλήρωση προφίλ

Από τις γυναίκες το 64% δηλώνει ναι, το 29% δηλώνει όχι και το 7% Δεν γνωρίζει / Δεν απαντά. Ενώ από τους άνδρες το 52% δηλώνει ναι, το 41% δηλώνει όχι και το 7% Δεν γνωρίζει / Δεν απαντά (Γράφημα 11).



Γράφημα 11: Συμπλήρωση προφίλ / φύλο

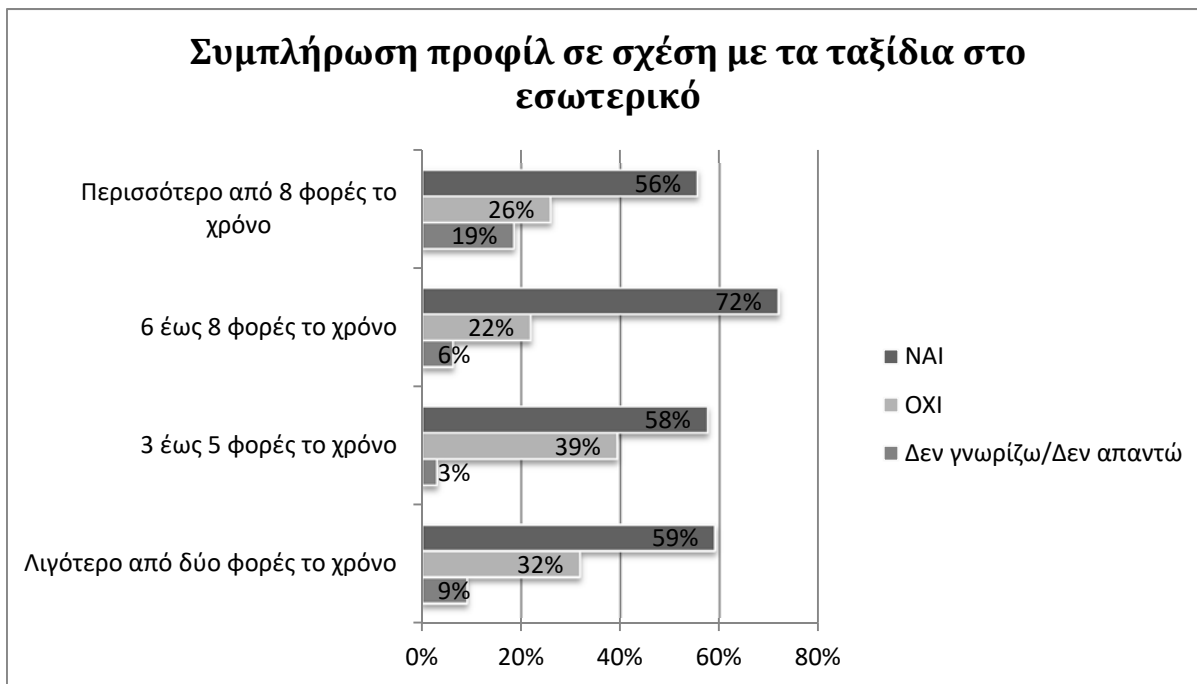
Από τις γυναίκες που θα συμπλήρωναν το προφίλ τους, το μεγαλύτερο ποσοστό, 47%, είναι ηλικίας 31 έως 40 ετών, ενώ από τους άνδρες το μεγαλύτερο ποσοστό 41% είναι ηλικίας 41 έως 50 ετών (Γράφημα 12).



Γράφημα 12: Συμπλήρωση προφίλ ανά ηλικία

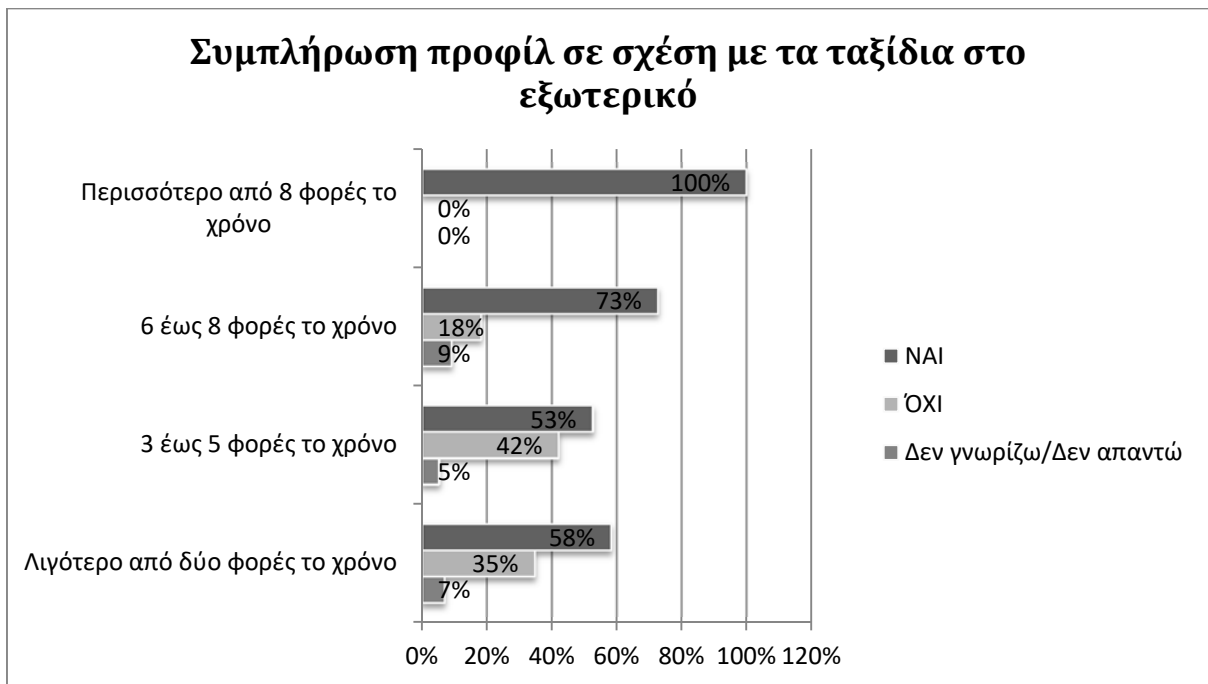
Σύμφωνα με τα προηγούμενα, οι γυναίκες ηλικίας 31 έως 40 ετών είναι περισσότερο θετικές στο να συμπληρώσουν το προφίλ τους.

Το μεγαλύτερο ποσοστό (72%), αυτών που είναι διατεθειμένοι να συμπληρώσουν ένα προφίλ στο σύστημα είναι αυτοί που ταξιδεύουν στο εσωτερικό 6 έως 8 φορές το χρόνο. Ενώ το μεγαλύτερο ποσοστό, 39%, αυτών που δεν είναι διατεθειμένοι να συμπληρώσουν ένα προφίλ στο σύστημα είναι αυτοί που ταξιδεύουν στο εσωτερικό 3 έως 5 φορές το χρόνο. (Γράφημα 13)



Γράφημα 13: Συμπλήρωση προφίλ / ταξίδια στο εσωτερικό

Όπως φαίνεται στο Γράφημα 14, το 100% αυτών που ταξιδεύουν στο εξωτερικό περισσότερες από 8 φορές θα συμπλήρωνε το προφίλ του στο σύστημα. Το ίδιο θα έκανε το 73%, αυτών που ταξιδεύουν από 6 έως 8 φορές το χρόνο. Ενώ, το μεγαλύτερο ποσοστό, 42%, αυτών που δεν είναι διατεθειμένοι να συμπληρώσουν ένα προφίλ στο σύστημα είναι αυτοί που ταξιδεύουν στο εσωτερικό 3 έως 5 φορές το χρόνο. Και αμέσως μετά, το 35% αυτοί που ταξιδεύουν λιγότερο από δυο φορές το χρόνο.

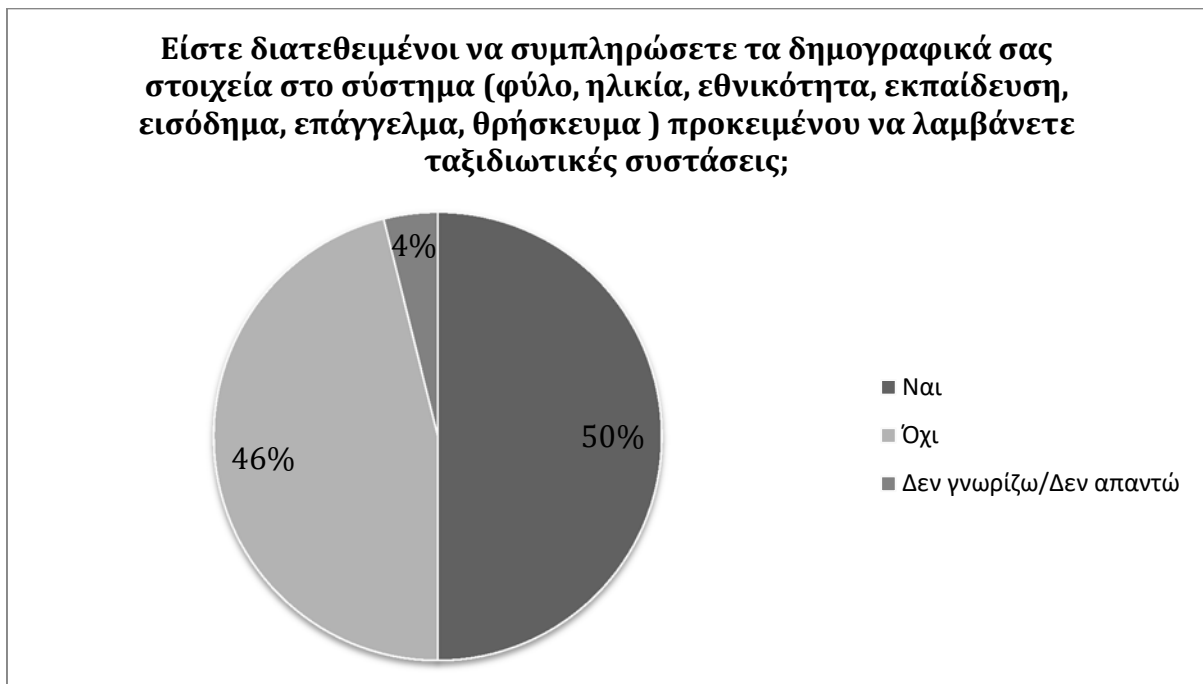


Γράφημα 14: Συμπλήρωση προφίλ / ταξίδια στο εξωτερικό

Σύμφωνα με τα αποτελέσματα περισσότερο διατεθειμένοι να συμπληρώσουν το προφίλ τους είναι αυτοί που ταξιδεύουν στο εξωτερικό πάνω από 6 φορές το χρόνο. Ενώ για το εσωτερικό είναι αυτοί που ταξιδεύουν 6-8 φορές.

5.2.4.2 Συμπλήρωση δημογραφικών στοιχείων

Το 50% των ερωτηθέντων θα συμπλήρωνε τα δημογραφικά του στοιχεία στο σύστημα, το 46% δηλώνει πως δεν θα το έκανε, ενώ το 4% Δεν γνωρίζει/ Δεν απαντά (Γράφημα 15).



Γράφημα 15: Συμπλήρωση δημογραφικών στοιχείων

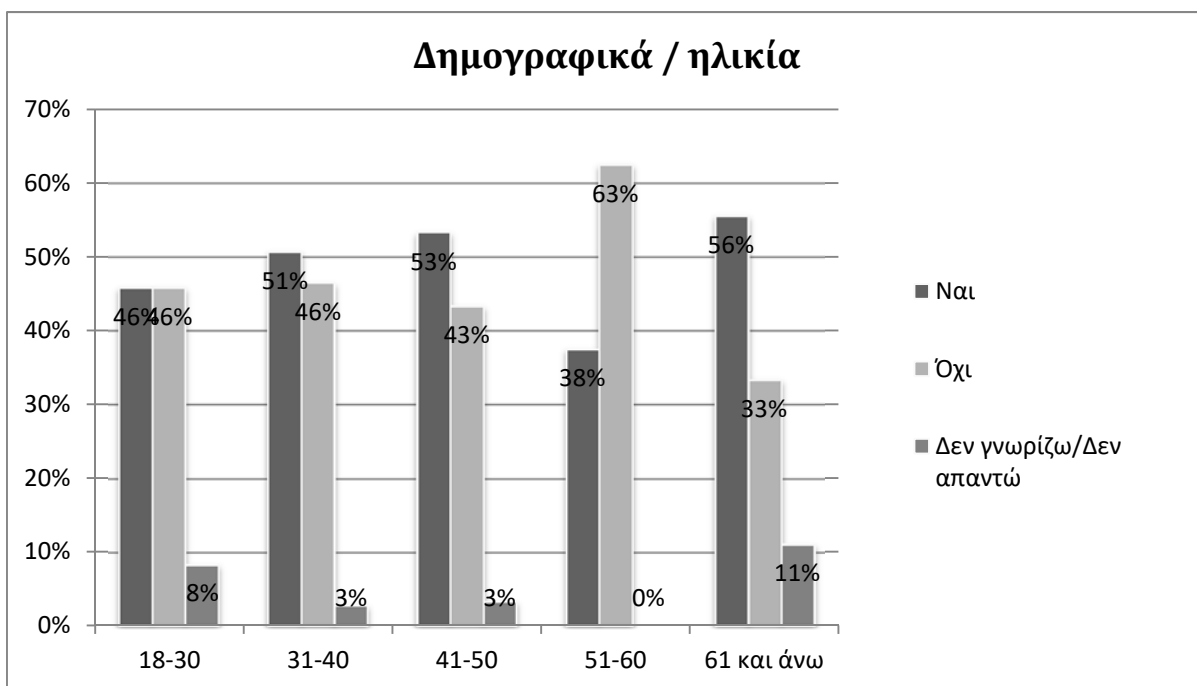
Από αυτούς που θα συμπλήρωναν τα δημογραφικά τους στοιχεία, περισσότερες είναι οι γυναίκες (52%) και λιγότεροι οι άνδρες (46%), περισσότερο αρνητικοί είναι οι άντρες (49%), λιγότερο οι γυναίκες (45%) (Γράφημα 16).



Γράφημα 16: Συμπλήρωση δημογραφικών ανά φύλο

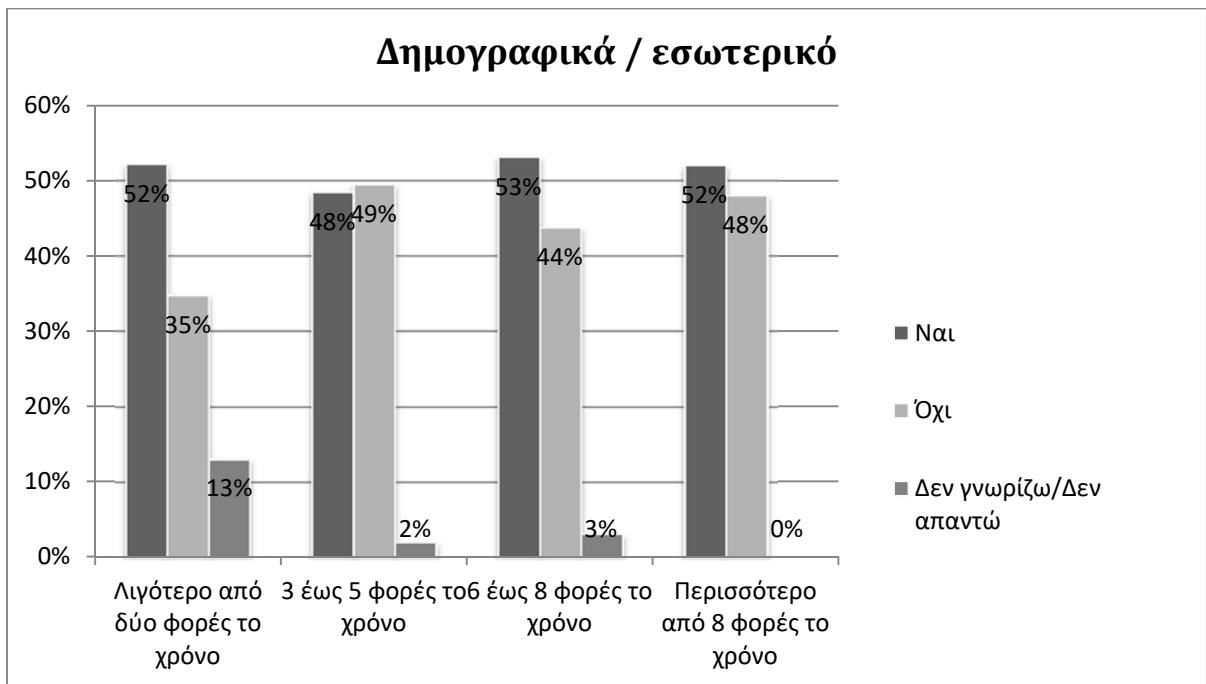
Να σημειώσουμε ότι ενώ το 60% είναι διατεθειμένο να συμπληρώσει το προφίλ του στο σύστημα, το ποσοστό μειώνεται στο 50% όταν αυτές οι ερωτήσεις αφορούν σε δημογραφικά στοιχεία. Επίσης περισσότερο διατεθειμένες να συμπληρώσουν τα δημογραφικά τους στοιχεία είναι οι γυναίκες, ενώ περισσότερο αρνητικοί είναι οι άντρες.

Οι περισσότερο διατεθειμένοι να συμπληρώσουν τα δημογραφικά τους στοιχεία είναι ηλικίας 61 και άνω (56%), 41 έως 50 (53%) και 31 έως 40 (51%). Ενώ οι περισσότερο αρνητικοί είναι ηλικίας 51 έως 60 ετών, το 63% απάντησε πως δεν θα συμπλήρωνε τα δημογραφικά του στοιχεία. Τα άτομα ηλικίας 18 έως 30 είναι μοιρασμένοι, 46% δηλώνει ότι θα συμπλήρωνε και 46% ότι δεν θα συμπλήρωνε, ενώ ένα 8% Δεν γνωρίζει/Δεν απαντά (Γράφημα 17).



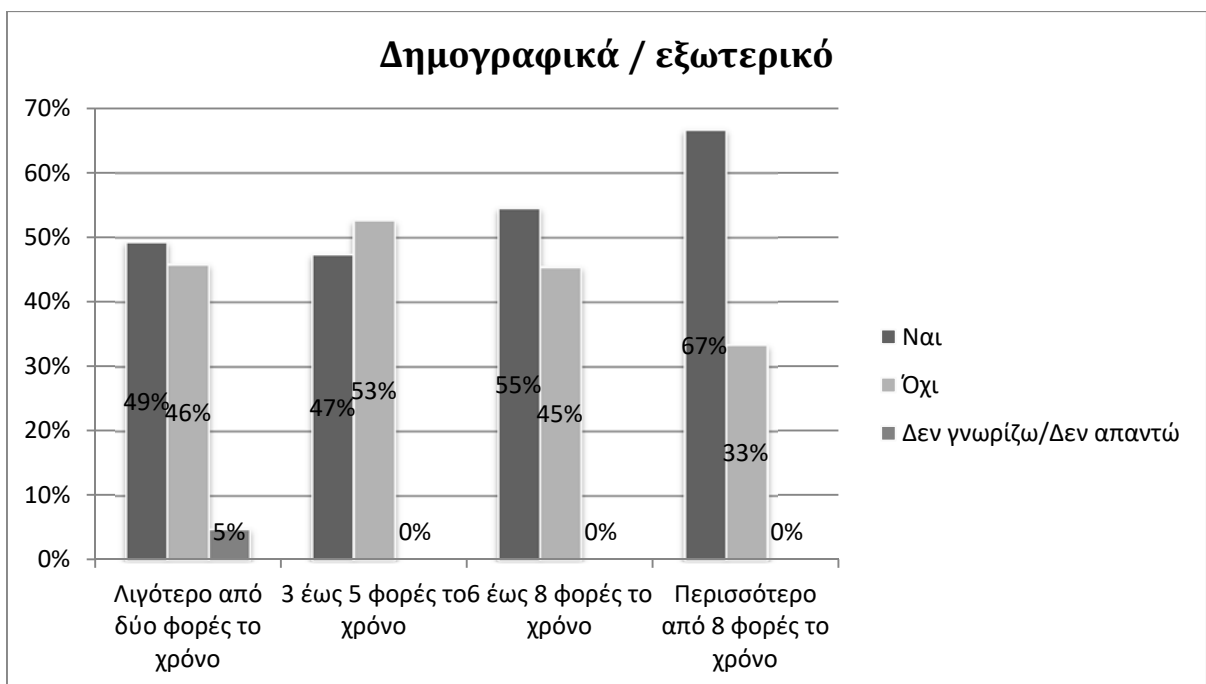
Γράφημα 17: Συμπλήρωση δημογραφικών ανά ηλικία

Περισσότερο διατεθειμένοι να συμπληρώσουν τα δημογραφικά τους στοιχεία είναι αυτοί που ταξιδεύουν 6 έως 8 φορές το χρόνο στο εσωτερικό (53%), ακολουθούν αυτοί που ταξιδεύουν λιγότερο από δύο φορές (52%) και αυτοί που ταξιδεύουν περισσότερο από 8 φορές το χρόνο (52%). Ενώ από τα άτομα που ταξιδεύουν 3 έως 5 φορές το χρόνο το 48% θα συμπλήρωνε τα δημογραφικά του στοιχεία ενώ το 49% δεν θα το έκανε (Γράφημα 18).



Γράφημα 18: Συμπλήρωση δημογραφικών ανά συχνότητα ταξιδιών στο εσωτερικό

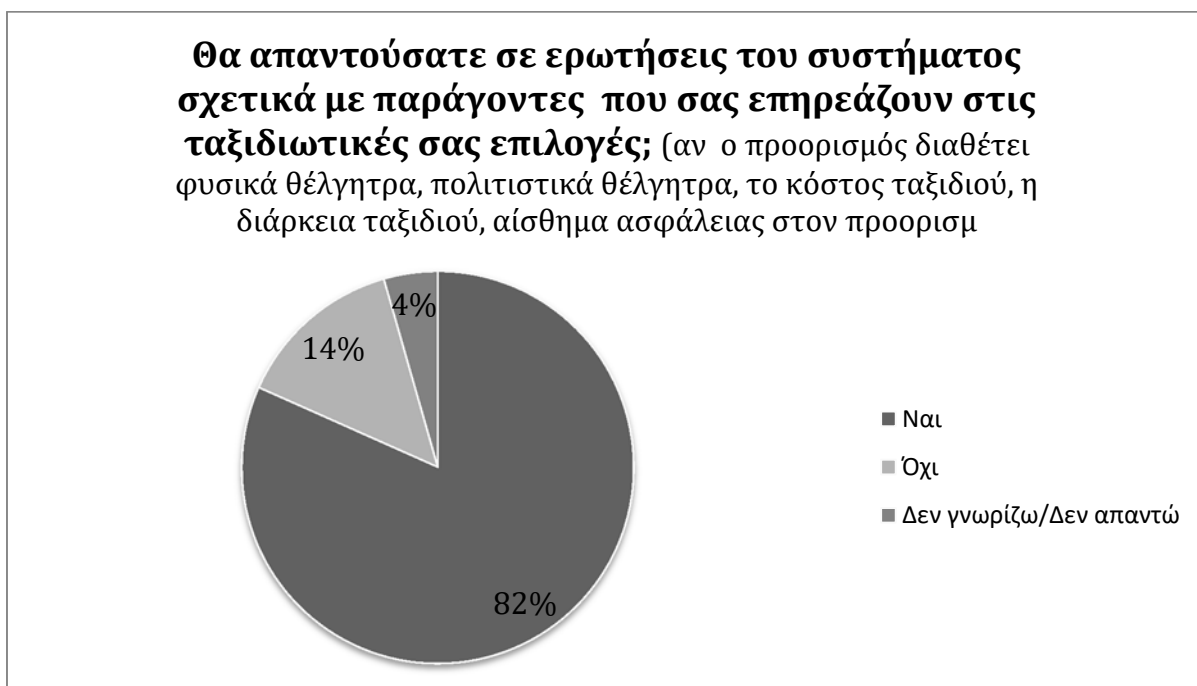
Όσον αφορά στη συχνότητα ταξιδιών στο εξωτερικό, οι περισσότερο διατεθειμένοι να συμπληρώσουν τα δημογραφικά τους στοιχεία είναι αυτοί που ταξιδεύουν περισσότερο από 8 φορές το χρόνο (67%) και λιγότερο διατεθειμένοι είναι αυτοί που ταξιδεύουν 3 έως 5 φορές το χρόνο (53%) (Γράφημα 19).



Γράφημα 19: Συμπλήρωση δημογραφικών ανά συχνότητα ταξιδιών στο εξωτερικό

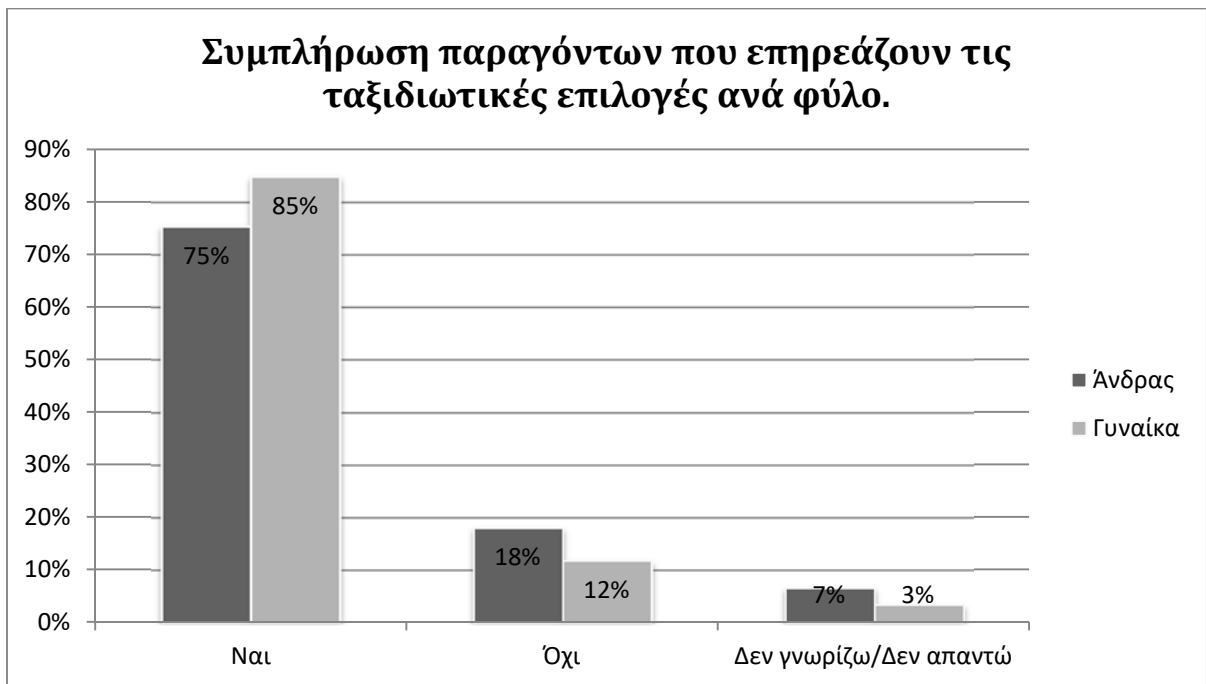
5.2.4.3 Συμπλήρωση ερωτήσεων σχετικά με τους παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές.

Στην επόμενη ερώτηση ένα μεγάλο ποσοστό, το 82%, θα απαντούσε σε ερωτήσεις του συστήματος σχετικά με τους παράγοντες που το επηρεάζουν στις επιλογές των ταξιδιών του, το 14% δηλώνει πως δεν θα το έκανε, ενώ το 4% Δεν γνωρίζει/ Δεν απαντά (Γράφημα 20).



Γράφημα 20: Συμπλήρωση ταξιδιωτικών επιλογών

Και πάλι οι γυναίκες είναι περισσότερο διατεθειμένες να συμπληρώσουν ερωτήσεις σχετικές με τις ταξιδιωτικές τους επιλογές στο σύστημα (85%) και λιγότερο οι άντρες (75%) (Γράφημα 21).



Γράφημα 21: Συμπλήρωση ταξιδιωτικών επιλογών ανά φύλο

5.2.4.4 Συμπλήρωση ερωτήσεων σχετικά με τις ταξιδιωτικές προτιμήσεις.

Παρόμοια αποτελέσματα έχουμε και στην επόμενη ερώτηση, το 76% θα απαντούσε σε ερωτήσεις του συστήματος σχετικά με τις ταξιδιωτικές του προτιμήσεις, το 21% δηλώνει πως δεν θα το έκανε, ενώ το 3% Δεν γνωρίζει/ Δεν απαντά (Γράφημα 22).



Γράφημα 22: Συμπλήρωση ταξιδιωτικών προτιμήσεων

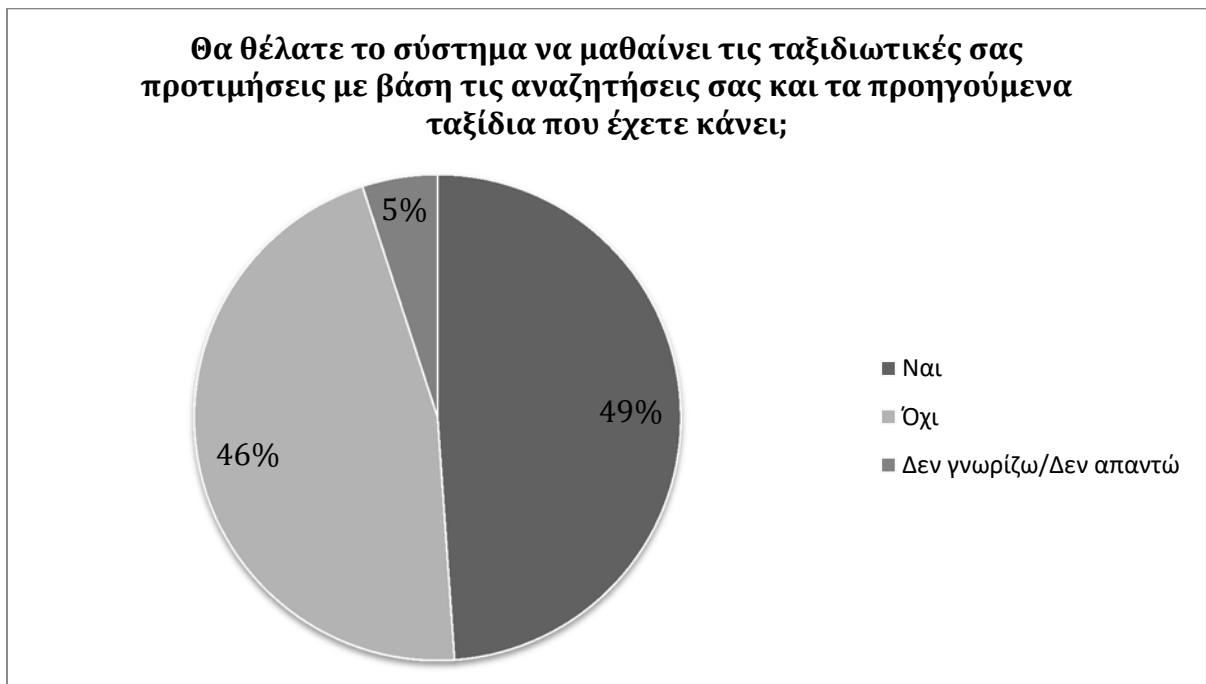
Και πάλι βλέπουμε ότι οι γυναίκες είναι περισσότερο θετικές στο να συμπληρώσουν τις ταξιδιωτικές τους προτιμήσεις (79%) και λιγότερο θετικοί οι άνδρες (70%) (Γράφημα 23).



Γράφημα 23: Συμπλήρωση ταξιδιωτικών προτιμήσεων ανά φύλο

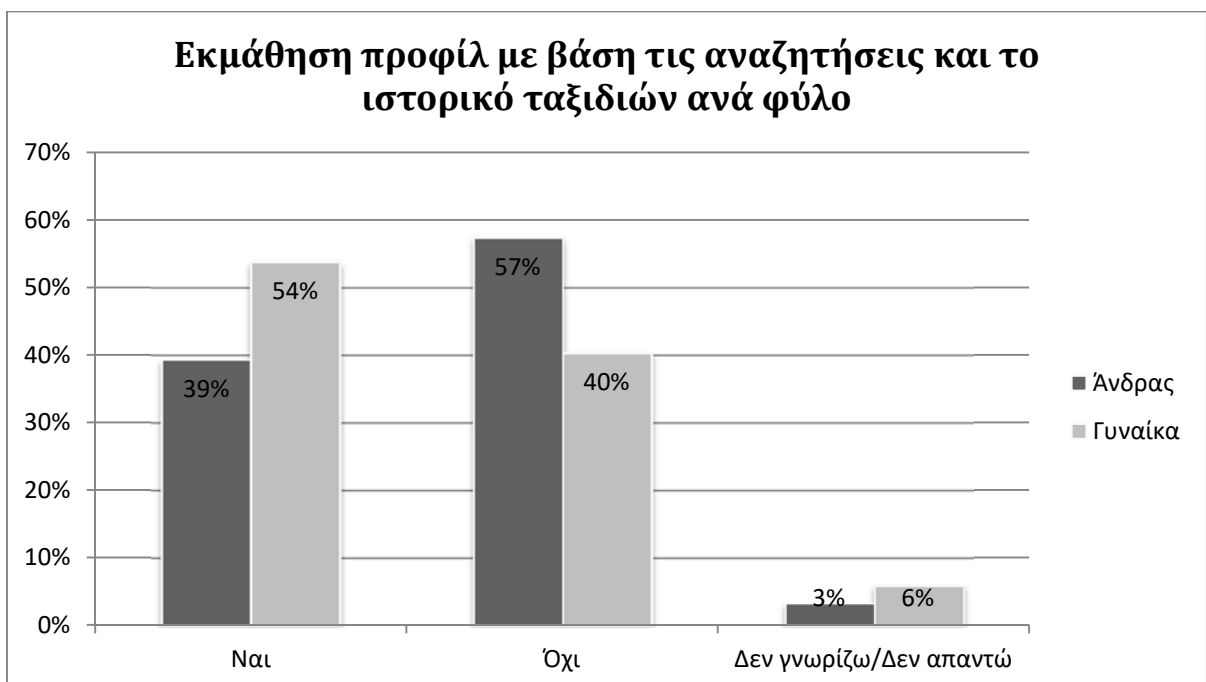
5.2.4.5 Εκμάθηση προφίλ μέσω των αναζητήσεων των χρηστών και το ιστορικό των ταξιδιών τους.

Το 49% θα ήθελε το σύστημα να μαθαίνει τις ταξιδιωτικές του προτιμήσεις με βάση τις αναζητήσεις του και τα προηγούμενα ταξίδια που έχει πραγματοποιήσει, ενώ το 46% δεν θα ήθελε και το 5% Δεν γνωρίζει/ Δεν απαντά (Γράφημα 24).



Γράφημα 24: Ταξιδιωτικές προτιμήσεις με βάση τις αναζητήσεις και το ιστορικό ταξιδιών τους

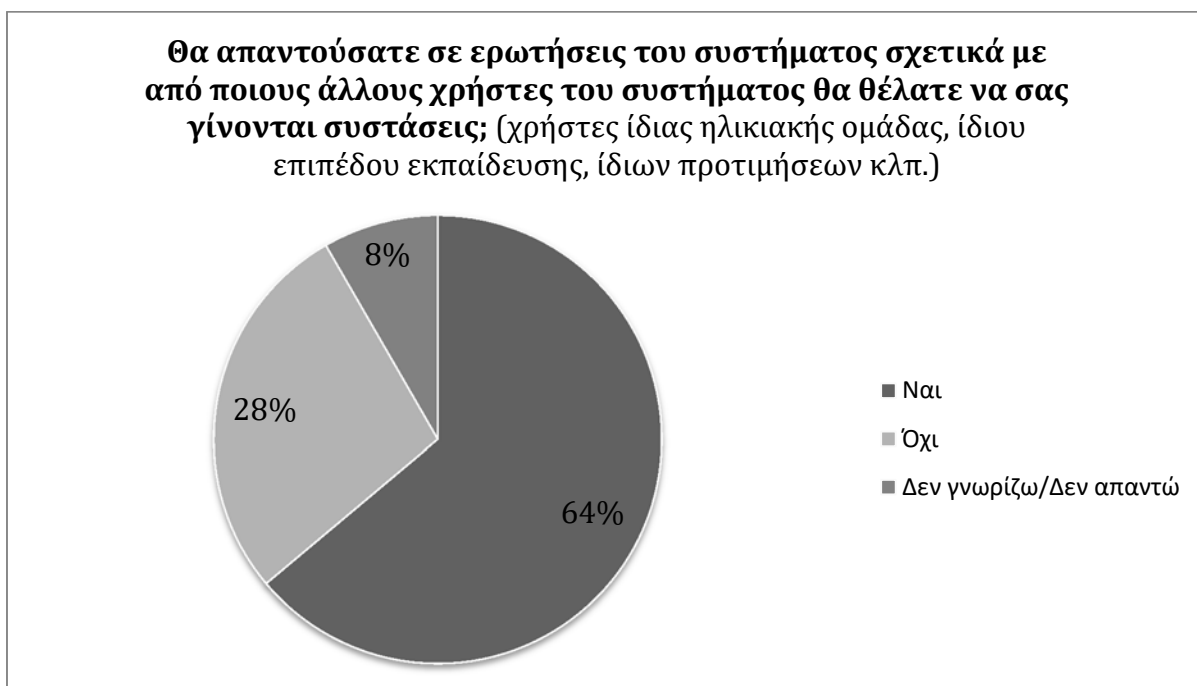
Οι περισσότερες γυναίκες (54%) και σε αυτή την ερώτηση δηλώνουν θετικές στο να τους μαθαίνει το σύστημα με βάση τις αναζητήσεις τους και το ιστορικό των ταξιδιών τους, ενώ το 40% είναι αρνητικό. Από την άλλη οι περισσότεροι άντρες (57%) είναι αρνητικοί ενώ λιγότεροι (39%) είναι θετικοί (Γράφημα 25).



Γράφημα 25: Εκμάθηση προφίλ μέσω αναζητήσεων και ιστορικού ταξιδιών ανά φύλο

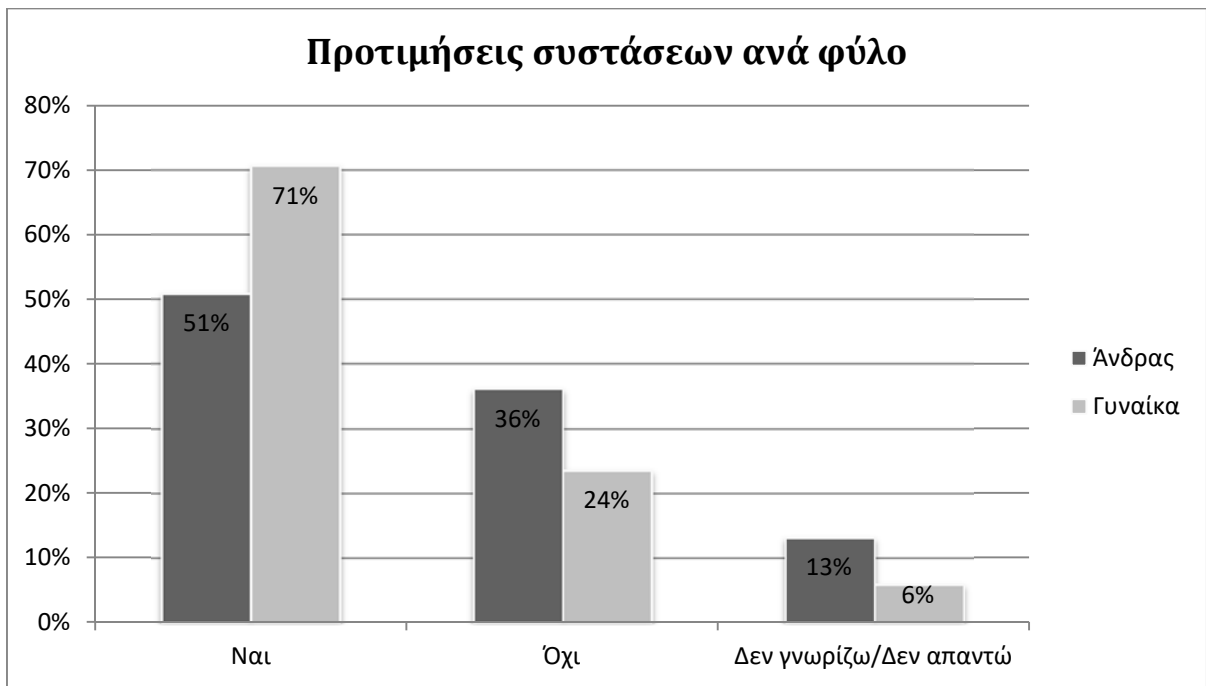
5.2.4.6 Ερωτήσεις συστήματος από ποιες ομάδες χρηστών θα ήθελαν να τους γίνονται οι συστάσεις

Το 64% θα ήθελε να δηλώνει στο σύστημα από ποια κατηγορία χρηστών θα ήθελε να του γίνονται συστάσεις, το 28% δεν θα ήθελε ενώ το 8% Δεν γνωρίζει/ Δεν απαντά (Γράφημα 26).



Γράφημα 26: Προτιμήσεις συστάσεων

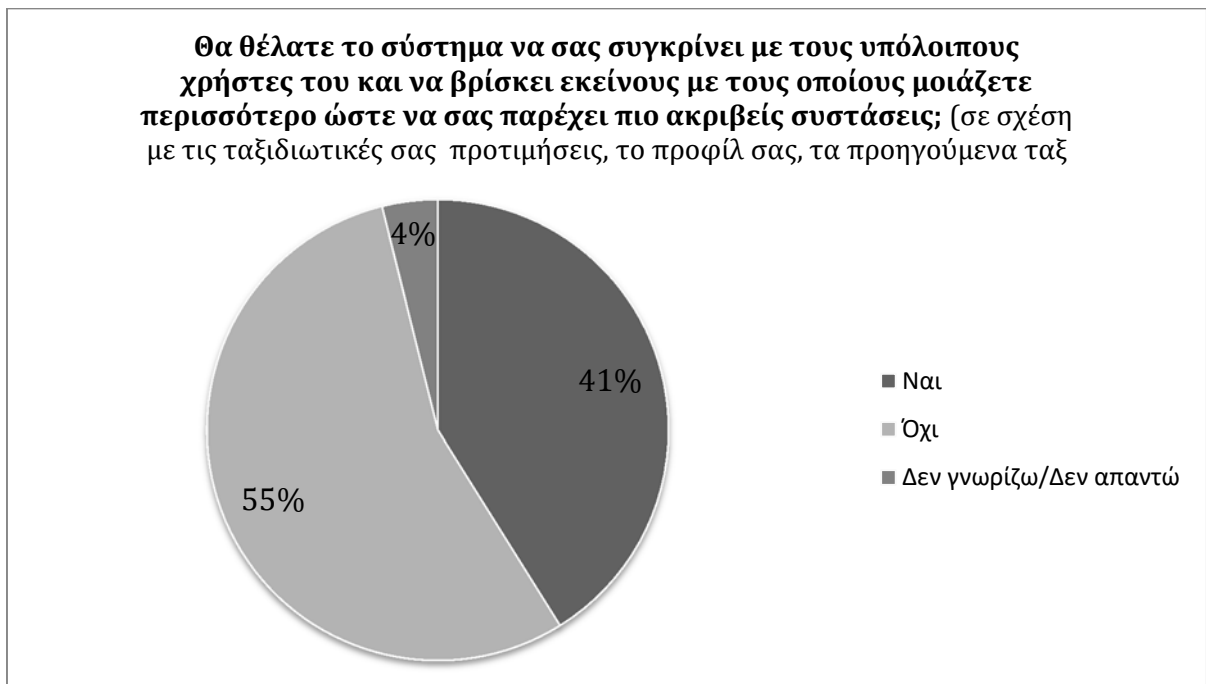
Από τις γυναίκες το 71% θα απαντούσε στις ερωτήσεις ενώ το 24% όχι. Από τους άντρες το 51% θα απαντούσε στις ερωτήσεις και το 36% δεν θα το έκανε (Γράφημα 27).



Γράφημα 27: Συστάσεις

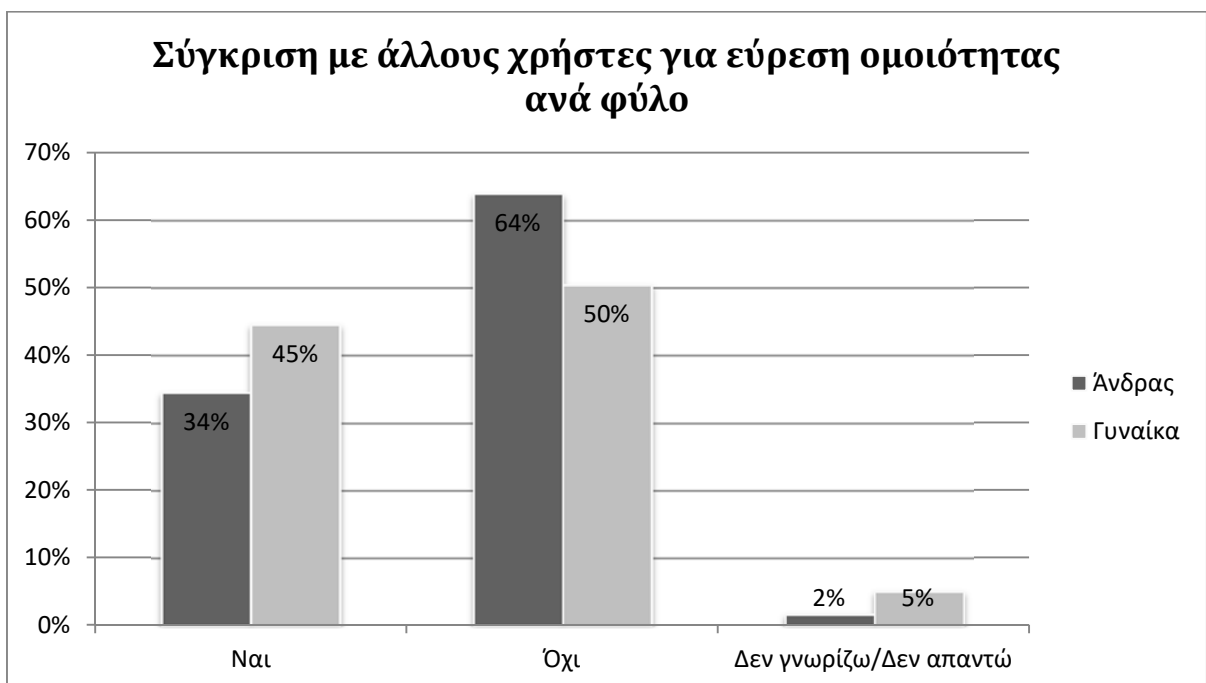
5.2.4.7 Ερωτήσεις συστήματος για εύρεση ομοιότητας βάση της σύγκρισης με άλλους χρήστες

Τέλος, στην τελευταία ερώτηση το 55% δηλώνει ότι δεν θα ήθελε το σύστημα να τους συγκρίνει με άλλους χρήστες του συστήματος ώστε να βρίσκει τους όμοιους προς αυτούς, το 41% δηλώνει πως θα ήθελε, ενώ το 4% Δεν γνωρίζει/Δεν απαντά (Γράφημα 28).



Γράφημα 28: Σύγκριση με άλλους χρήστες

Οι περισσότεροι άνδρες, το 64%, δεν θα απαντούσαν στις ερωτήσεις, ενώ το 34% θα το έκανε. Όσον αφορά τις γυναίκες, το 50% δεν θα απαντούσε, ενώ το 45% θα το έκανε (Γράφημα 29).



Γράφημα 29: Ερωτήσεις για εύρεση ομοιότητας με άλλους χρήστες

Κεφάλαιο 6

Επίλογος – Συμπεράσματα

6.1 Πλεονεκτήματα – Μειονεκτήματα συστήματος

Το σύστημα που παρουσιάζουμε είναι ένα χειροκίνητο σύστημα, ζητά από τους χρήστες να καταχωρήσουν τις ταξιδιωτικές επιλογές και προτιμήσεις τους καθώς και από ποιους χρήστες θα θέλανε συστάσεις. Το πλεονέκτημα είναι η διαφάνεια της συμπεριφοράς του συστήματος. Ένα από τα κύρια πλεονεκτήματα αυτής της μεθόδου είναι η Αιτιολόγηση (Justifiability) (Ricci, Rokach, Shapira, & Kantor, 2011). Μπορεί, δηλ. το σύστημα να αιτιολογήσει γιατί έγιναν οι συγκεκριμένες συστάσεις. Έτσι ο χρήστης γνωρίζοντας για ποιο λόγο του έγινε η σύσταση, κατανοεί τη σύσταση και δημιουργείται σχέση εμπιστοσύνης ανάμεσα σε αυτόν και το σύστημα, συνεχίζει να το χρησιμοποιεί, το οποίο αποτελεί απαραίτητη προϋπόθεση για την λειτουργία του συστήματος.

Ένα από τα ισχυρά σημεία των συστημάτων που βασίζονται στη γειτονιά είναι η αποτελεσματικότητά τους. Σε αντίθεση με τα περισσότερα συστήματα που βασίζονται σε μοντέλα, δεν απαιτούν δαπανηρές φάσεις εκπαίδευσης, οι οποίες πρέπει να διεξάγονται σε συχνά διαστήματα σε μεγάλες εμπορικές εφαρμογές. Ενώ η φάση συστάσεων είναι συνήθως πιο ακριβή από ό, τι οι μέθοδοι που βασίζονται σε μοντέλα, οι πλησιέστεροι γείτονες μπορούν να υπολογιστούν εκ των προτέρων σε ένα βήμα εκτός σύνδεσης, παρέχοντας σχεδόν άμεσες συστάσεις. Επιπλέον, η αποθήκευση αυτών των πλησιέστερων γειτόνων απαιτεί πολύ λίγη μνήμη, καθιστώντας τέτοιες προσεγγίσεις κλιμακωτές σε εφαρμογές που έχουν εκατομμύρια χρήστες και αντικείμενα (Ricci, Rokach, Shapira, & Kantor, 2011).

Το μειονέκτημα της χειροκίνητης εισαγωγής και ενημέρωσης στοιχείων είναι ότι απαιτεί προσπάθεια από τον χρήστη. Συνήθως οι χρήστες δεν είναι πρόθυμοι να αφιερώσουν χρόνο και να απαντήσουν τις ερωτήσεις του συστήματος (Montaner, López, & de la Rosa, 2003). Επιπλέον μπορεί οι χρήστες να μην μπορούν να

προσδιορίσουν και να απαντήσουν με ακρίβεια για το ποιες είναι οι προτιμήσεις τους και οι ταξιδιωτικές επιλογές τους. Τέλος, οι προτιμήσεις του χρήστη είναι πολύ πιθανό να αλλάξουν με την πάροδο του χρόνου, οπότε θα χρειαστεί να ενημερώσουν τα στοιχεία τους αν θέλουν πιο ακριβείς συστάσεις.

Το κύριο ζήτημα της συνεργατικής προσέγγισης είναι ότι απλουστεύουμε ένα ταξίδι στο σημείο όπου τα ταξίδια δύο χρηστών είναι ίδια, όμως ο κάθε χρήστης/ταξιδιώτης μπορεί να το βιώσει διαφορετικά (Ricci, Travel Recommender Systems, 2002).

6.2 Συμπεράσματα έρευνας

Όσον αφορά τα αποτελέσματα της έρευνας καταλήγουμε στα εξής:

Από πού θέλουν να προέρχονται οι συστάσεις:

- Ένα πολύ μεγάλο ποσοστό (83%) απαντάει ναι στην ερώτηση αν θα ήθελε οι ταξιδιωτικές συστάσεις να προέρχονται από κάποιο άτομο.
- Μικρότερο ποσοστό, 69%, απαντάει ναι στην ερώτηση αν θα ήθελε οι ταξιδιωτικές συστάσεις να προέρχονται από μηχανή,
- Ενώ στην ερώτηση αν θα ήθελε συστάσεις από άνθρωπο ή από μηχανή το μεγαλύτερο ποσοστό 47% δηλώνει και από τα δύο, το 35% μόνο από άνθρωπο, ενώ μόλις το 9% μόνο από μηχανή.

Ερωτήσεις του συστήματος:

- Οι περισσότεροι είναι διατεθειμένοι να συμπληρώσουν το προφίλ τους στο σύστημα και να απαντήσουν σε ερωτήσεις σχετικά με τις ταξιδιωτικές προτιμήσεις τους και τις επιλογές τους. Να σημειωθεί ότι ενώ σε όλες τις ερωτήσεις οι θετικές απαντήσεις έχουν αρκετή διαφορά από τις αρνητικές δεν ισχύει το ίδιο στις ερωτήσεις που αφορούν τα δημογραφικά τους στοιχεία (50%-ναι, 46%-όχι). Παρόμοια, μικρή διαφορά παρατηρείται και στην ερώτηση για το αν θα ήθελε να τους μαθαίνει το σύστημα μέσω των αναζητήσεων που κάνουν και από το ιστορικό των ταξιδιών τους (49%-ναι, 46%-όχι). Αυτό ίσως να σημαίνει ότι οι χρήστες νιώθουν ανασφάλεια στο να παρέχουν ιδιωτικά δεδομένα. Θα μπορούσε το σύστημα κατά την διαδικασία της εγγραφής των χρηστών να τους ενημερώνει για τον τρόπο διαφύλαξης και διαχείρισης των

προσωπικών τους δεδομένων, ώστε να πεισθούν ότι τα δεδομένα τους είναι ασφαλή.

- Οι γυναίκες είναι περισσότερο διατεθειμένες από τους άνδρες να συμπληρώσουν τα στοιχεία τους στο σύστημα.

Συστάσεις:

- Ένα πολύ μεγάλο ποσοστό θα απαντούσε στις ερωτήσεις του συστήματος σχετικά με από ποιες κατηγορίες χρηστών θα προέρχονται οι συστάσεις (64%-Ναι, 28%-Όχι).
- Ενώ αντίθετα, αλλά με μικρότερη διαφορά των απαντήσεων ναι – όχι, οι περισσότεροι δεν θα θέλανε το σύστημα να τους συγκρίνει με άλλους χρήστες ώστε να τους παρέχει πιο ακριβείς συστάσεις (41%-Ναι, 55%-Όχι).

Κατά την διεξαγωγή της έρευνάς μας, το ερωτηματολόγιο στάλθηκε και σε ταξιδιωτικά γραφεία ζητώντας τους την άδεια για την ανάρτησή του στη σελίδα που διατηρούν σε κοινωνικά μέσα δικτύωσης. Σε αρκετά από αυτά οι συνεργάτες τους ή οι διαχειριστές της σελίδας απαντήσανε στο ερωτηματολόγιο, κάποια μας επιτρέψανε να την αναρτήσουμε, ενώ τρία από αυτά εκδήλωσαν ενδιαφέρον για την εργασία.

Από όλα τα παραπάνω το αποτέλεσμα που προκύπτει είναι θετικό. Θα μπορούσε λοιπόν, με περισσότερη βιβλιογραφική μελέτη και βελτιώσεις να υλοποιηθεί ένα τέτοιο σύστημα.

6.3 Επεκτάσεις

Ως μελλοντική επέκταση του συστήματος που σχεδιάσαμε, θα μπορούσαμε να αναφέρουμε την σύνδεσή του με κοινωνικά μέσα δικτύωσης και με ταξιδιωτικούς ιστοτόπους και εφαρμογές αξιολόγησης τουριστικών υπηρεσιών. Η σύνδεση με κοινωνικά μέσα δικτύωσης κυρίως θα προσανατολιζόταν στην εξαγωγή (α) του προφίλ χρήστη ώστε να μην είναι αναγκασμένος να συμπληρώνει το προφίλ του στο σύστημα, (β) του δικτύου των φίλων του ώστε να βελτιωθεί η λειτουργία εύρεσης όμοιων χρηστών. Επιπλέον, με τη σύνδεση σε ιστότοπους και εφαρμογές αξιολόγησης τουριστικών υπηρεσιών θα μπορούσαμε να εξάγουμε τις προτιμήσεις και τις ταξιδιωτικές επιλογές των χρηστών.


Τέλος, θα μπορούσαμε να επεκτείνουμε το σχεδιασμό του συστήματός μας ώστε να παράγει συστάσεις σε ομάδες χρηστών. Από τη στιγμή που έχει συγκεντρώσει τις προτιμήσεις τους θα μπορεί να ομαδοποιεί τα άτομα σύμφωνα με αυτές και να τους προτείνει προορισμούς.

Παράρτημα Α

Ερωτήσεις Συστήματος

Α Ερωτήσεις συστήματος

Α.1 Δημογραφικά στοιχεία



Ερωτήσεις συστήματος

Οθόνη 1η - Δημογραφικά στοιχεία

Φύλλο

Άνδρας

Γυναίκα

Σε ποιά ηλικιακή κατηγορία ανήκετε;

18-30

31-40

41-50

51-60

61 και άνω

Ποιά είναι η εθνικότητά σας;

Η απάντησή σας _____

Σε ποιο θρήσκευμα ανήκετε;

Χριστιανικό Ορθόδοξο

Χριστιανικό Καθολικό

Μουσουλμανικό

Εβραϊκό

Κανένα

Άλλο: _____

Ποιό είναι το επίπεδο εκπαίδευσής σας;

Καμία τυπική εκπαίδευση

Βασική εκπαίδευση (Δημοτικό σχολείο)

Δευτεροβάθμια εκπαίδευση (Γυμνάσιο - Λύκειο)

Ανώτατη εκπαίδευση (Πανεπιστήμιο, ΤΕΙ)

Μεταπτυχιακό

Διδακτορικό

Άλλο: _____

Σε ποιά κατηγορία ανήκει το επάγγελμά σας;

- Δημόσιος Υπάλληλος
- Ιδιωτικός Υπάλληλος
- Ελεύθερος Επαγγελματίας
- Συνταξιούχος
- Οικιακά
- Φοιτητής/-τρια
- Ανεργος

Σε ποιά κατηγορία ανήκει το μηνιαίο καθαρό εισόδημά σας;

- 1€-500€
- 501€-1.000€
- 1.001€-1.500€
- 1.501€-2.000€
- 2.001€-3.000€
- 3.001€ και άνω

Εάν έχετε κάποιες ιδιαίτερες δυσκολίες παρακαλώ προσδιορίστε ποιες είναι αυτές

- Κινητικές δυσκολίες
- Νοητικές δυσκολίες
- Προβλήματα ακοής
- Προβλήματα όρασης
- Δυσκολίες στην επικοινωνία
- Άλλο: _____

A.2 Παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές του χρήστη

Οθόνη 2η - Παράγοντες που επηρεάζουν τις ταξιδιωτικές επιλογές

Για την επιλογή ενός ταξιδιωτικού προορισμού εκτός από τον κύριο σκοπό του ταξιδιού σας, πόσο σημαντικό είναι για εσάς τα παρακάτω:

Ο προορισμός να διαθέτει φυσικά θέλγητρα (φυσικά τοπία, οικοσυστήματα, δρυμοί, ακρογιαλιές, σπήλαια, βουνά, λίμνες κ.λ.π.)

1 2 3 4 5 6 7 8 9 10
Καθόλου σημαντικό ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Πάρα πολύ σημαντικό

Ο προορισμός να διαθέτει κατασκευασμένα θέλγητρα (μνημεία, ιστορικοί-αρχαιολογικοί χώροι, κλασική- σύγχρονη αρχιτεκτονική, πάρκα, αθλητικά κέντρα, ενυδρεία κ.λ.π.)

1 2 3 4 5 6 7 8 9 10
Καθόλου σημαντικό ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Πάρα πολύ σημαντικό

Ο προορισμός να διαθέτει πολιτιστικά θέλγητρα (μουσεία, πινακοθήκες, ναούς, θέατρα, χώρους τέχνης κ.λ.π.)

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πάρα πολύ σημαντικό

Ο προορισμός να διαθέτει κοινωνικά θέλγητρα (ήθη, έθιμα, τρόπος ζωής ντόπιου πληθυσμού κ.λ.π.)

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πάρα πολύ σημαντικό

Για την επιλογή ταξιδιωτικού προορισμού, πόσο λαμβάνεται υπ' όψιν το διαθέσιμο εισόδημά σας (ή το κόστος ταξιδιού)

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πάρα πολύ σημαντικό

Πόσο σημαντικό είναι το μεταφορικό μέσο για την επιλογή ενός ταξιδιωτικού προορισμού;

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πολύ σημαντικό

Πόσο σημαντική είναι η ευκολία προσέγγισης ενός ταξιδιωτικού προορισμού;

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πολύ σημαντικό

Πόσο σημαντική είναι η διάρκεια ταξιδιού στην επιλογή ενός προορισμού;

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πολύ σημαντικό

Πόσο σημαντικό είναι το αίσθημα ασφάλειας στον προορισμό για την επιλογή του;

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πολύ σημαντικό

Πόσο λαμβάνεται υπ'όψιν την καλή κατάσταση οδικού δικτύου και μεταφορών του προορισμού για την επιλογή του;

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό

Πολύ σημαντικό

Για την επιλογή ενός προορισμού πόσο σημαντικό είναι ο προορισμός να διαθέτει εγκαταστάσεις και να προσφέρει υπηρεσίες (καταλύματα, εστιατόρια κ.λ.π.) διαφόρων κατηγοριών;

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό Πολύ σημαντικό

Πόσο σημαντικό είναι για εσάς το κλίμα ενός προορισμού (καιρικές συνθήκες) για την επιλογή του;

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό Πολύ σημαντικό

Παρακαλώ προσδιορίστε κάτι άλλο (εκτός από τον κύριο σκοπό του ταξιδιού) που είναι σημαντικό για εσάς στην επιλογή ενός ταξιδιωτικού προορισμού καθώς και το επίπεδο σημαντικότητας βαθμολογώντας από το 1 έως το 10.

Η απάντησή σας

A.3 Προτιμήσεις συστάσεων

Οθόνη 3η - Προτιμήσεις συστάσεων

Επιλέξτε πόσο σημαντικό είναι για εσάς, οι συστάσεις να προέρχονται από χρήστες με

το ίδιο φύλλο

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό Πάρα πολύ σημαντικό

την ίδια ηλικιακή ομάδα

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό Πάρα πολύ σημαντικό

την ίδια εθνικότητα

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό Πάρα πολύ σημαντικό

το ίδιο θρήσκευμα

1 2 3 4 5 6 7 8 9 10

Καθόλου σημαντικό Πάρα πολύ σημαντικό

το ίδιο επίπεδο εκπαίδευσης

1 2 3 4 5 6 7 8 9 10

Καθόλου
σημαντικό

Πάρα πολύ
σημαντικό

το ίδιο επάγγελμα

1 2 3 4 5 6 7 8 9 10

Καθόλου
σημαντικό

Πάρα πολύ
σημαντικό

το ίδιο επίπεδο εισοδήματος

1 2 3 4 5 6 7 8 9 10

Καθόλου
σημαντικό

Πάρα πολύ
σημαντικό

την ίδια οικογενειακή κατάσταση

1 2 3 4 5 6 7 8 9 10

Καθόλου
σημαντικό

Πάρα πολύ
σημαντικό

Επιλέξτε πόσο σημαντικό είναι για εσάς, οι συστάσεις να προέρχονται από χρήστες με ίδιες προτιμήσεις ως προς

τον σκοπό του ταξιδιού

1 2 3 4 5 6 7 8 9 10

Καθόλου
σημαντικό

Πάρα πολύ
σημαντικό

τον προορισμό (εσωτερικό-εξωτερικό)

1 2 3 4 5 6 7 8 9 10

Καθόλου
σημαντικό

Πάρα πολύ
σημαντικό

την διάρκεια του ταξιδιού

1 2 3 4 5 6 7 8 9 10

Καθόλου
σημαντικό

Πάρα πολύ
σημαντικό

Παρακαλώ προσδιορίστε κάποια άλλη κατηγορία χρηστών από τους οποίους θα θέλατε να σας προτείνει το σύστημα καθώς και το επίπεδο σημαντικότητας βαθμολογώντας από το 1 έως το 10.

Η απάντησή σας

A.4 Προτιμήσεις και συχνότητα επιλογής

Πόσο συχνά ταξιδεύετε για

	Ποτέ	Σπάνια	Μερικές φορές	Συχνά	Πολύ συχνά
Διακοπές/ Αναψυχή	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Επίσκεψη φίλων/ συγγενών	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Επαγγελματικούς λόγους	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Θρησκεία/ Προσκυνήματα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Λόγους υγείας	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Αθλητικούς λόγους (παρακολούθηση ή συμμετοχή)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Πολιτιστικούς-πνευματικούς λόγους (μουσεία, αρχαιολογικοί χώροι κ.λ.π.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Οικονομικοί λόγοι (ψώνια)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Λόγους γαστρονομικού ενδιαφέροντος	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Εκπαιδευτικούς/Επιστημονικούς λόγους(συνέδρια κλπ.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Γνωριμία - επαφή με τη φύση (οικολογικός τουρισμός)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Κοινωνικό τουρισμό	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Πόσο συχνά ταξιδεύετε

	Ποτέ	Σπάνια	Μερικές φορές	Συχνά	Πολύ συχνά
Μόνος	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Με σύντροφο	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Με παρέα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Με παιδιά	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Με ομάδες (group)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Πόσο συχνά ταξιδεύετε με:

	Ποτέ	Σπάνια	Μερικές φορές	Συχνά	Πολύ συχνά
Αεροπλάνο	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Αυτοκίνητο	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Λεωφορείο / τρένο	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Πλοίο	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Πόσο συχνά επιλέγετε προορισμό

	Ποτέ	Σπάνια	Μερικές φορές	Συχνά	Πολύ συχνά
εκτός της χώρας διαμονής σας (εξωτερικό)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
εντός της χώρας διαμονής σας (εσωτερικό)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Πόση είναι συνήθως η διάρκεια ταξιδιού;

	Ποτέ	Σπάνια	Μερικές φορές	Συχνά	Πολύ συχνά
Σύντομο (1 έως δύο διανυκτερεύσεις)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2-7 διανυκτερεύσεις	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8-15 διανυκτερεύσεις	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15-30 διανυκτερεύσεις	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Πάνω από ένα μήνα	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Παράρτημα Β

Πίνακες Συστήματος

Β.1 Προφίλ χρήστη

ΠΕΔΙΑ	ΠΕΡΙΓΡΑΦΗ	ΧΡΗΣΗ	ΥΠΟΧΡΕΩΤΙΚΟ	ΠΙΝΑΚΑΣ
user_id	αναγνωριστικό χρήστη	για δημογραφικό φιλτράρισμα	ΝΑΙ	data_D
gender	φύλλο		ΝΑΙ	
age	ηλικιακή ομάδα		ΝΑΙ	
nationality	εθνικότητα		ΝΑΙ	
denomination	θρήσκευμα		ΝΑΙ	
education	επίπεδο εκπ/σης		ΝΑΙ	
profession	επάγγελμα		ΝΑΙ	
income	εισοδηματική κλάση		ΝΑΙ	
disabilities	δυσκολίες		ΌΧΙ	

ΠΕΔΙΑ	ΠΕΡΙΓΡΑΦΗ	ΧΡΗΣΗ	ΥΠΟΧΡΕΩΤΙΚΟ	ΠΙΝΑΚΑΣ
user_id	αναγνωριστικό χρήστη	για προσδιορισμό ταξιδιωτικών προτιμήσεων	ΝΑΙ	data_pref
nature	φυσικά θέλγητρα		ΌΧΙ	
monuments	κατασκευασμένα		ΌΧΙ	
culture	πολιτιστικά θέλγητρα		ΌΧΙ	
society	κοινωνικά θέλγητρα		ΌΧΙ	
cost	κόστος ταξιδιού		ΌΧΙ	
access	πρόσβαση		ΌΧΙ	
safety	αίσθημα ασφάλειας		ΌΧΙ	
r_tr_network	οδικό δίκτυο και δίκτυο μεταφορών		ΌΧΙ	
facilities_services	εγκαταστάσεις και υπηρεσίες τουρισμού		ΌΧΙ	
clime	κλίμα-καιρικές συνθήκες		ΌΧΙ	
free_field	ελεύθερο πεδίο για να καταγράψει ο χρήστης τι άλλο είναι σημαντικό για αυτόν		ΌΧΙ	

ΠΕΔΙΑ	ΠΕΡΙΓΡΑΦΗ	ΧΡΗΣΗ	ΥΠΟΧΡΕΩΤΙΚΟ	ΠΙΝΑΚΑΣ
user_id	αναγνωριστικό χρήστη	για προσδιορισμό συστάσεων με βάση	ΌΧΙ	pref_rec_de m
rec_gende r	ίδιο φύλλο			
rec_age	ίδια ηλικιακή ομάδα			

rec_nation	ίδια εθνικότητα	δημογραφικά χαρακτηριστικά		
rec_deno m	ίδιο θρήσκευμα			
rec_edu	ίδιο επίπεδο εκπ/σης			
rec_prof	ίδιο επάγγελμα			
rec_incom	ίδια εισοδηματική κλάση			

ΠΕΔΙΑ	ΠΕΡΙΓΡΑΦΗ	ΧΡΗΣΗ	ΥΠΟΧΡΕΩΤΙΚΟ	ΠΙΝΑΚΑΣ
user_id	αναγνωριστικό χρήστη	για προσδιορισμό συστάσεων με βάση τις προτιμήσεις των χρηστών	ΟΧΙ	pref_rec
rec_purpose	ίδιος σκοπός			
transport	μεταφορικό μέσο			
company	με ποιον ταξιδεύει			
rec_destination	προορισμός(εσ.-εξ.)			
rec_durat	ίδια διάρκεια			
free_field	ελεύθερο πεδίο για να καταγράψει ο χρήστης τι άλλο είναι σημαντικό για αυτόν			

ΠΕΔΙΑ	ΠΕΡΙΓΡΑΦΗ	ΧΡΗΣΗ	ΥΠΟΧΡΕΩΤΙΚΟ	ΠΙΝΑΚΑΣ
user_id	αναγνωριστικό χρήστη	ιστορικό ταξιδιών του χρήστη		travel_history
country	χώρα			
city	πόλη			
purpose	σκοπός			
company	με ποιόν ταξίδεψε			
transport	μεταφορικό μέσο			
destination	προορισμός			
duration	διάρκεια			
date_tr	ημερομηνία			
rate	αξιολόγηση			

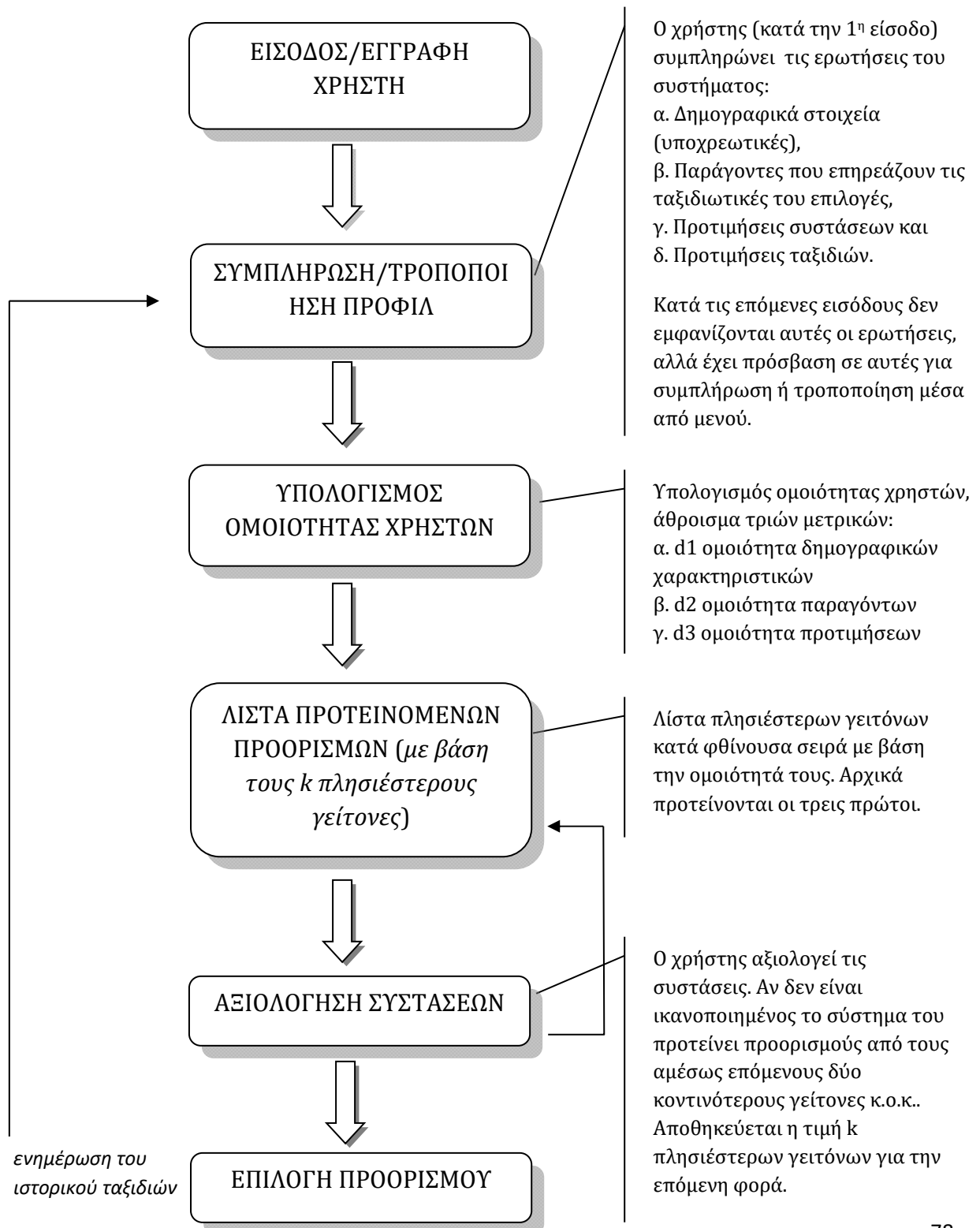
B.2 Προφίλ προορισμού

ΠΕΔΙΑ	ΠΕΡΙΓΡΑΦΗ	ΠΙΝΑΚΑΣ
destination_id	Αναγνωριστικό προορισμού	travel
country	ΧΩΡΑ	
city	ΠΟΛΗ	
purpose	ΣΚΟΠΟΣ	
company	ΜΕ ΠΟΙΟΝ ΤΑΞΙΔΕΥΣΕ	
transport	ΜΕΤΑΦΟΡΙΚΟ ΜΕΣΟ	
destination	ΠΡΟΟΡΙΣΜΟΣ	
duration	ΔΙΑΡΚΕΙΑ ΤΑΞΙΔΙΟΥ	

Παράρτημα Γ

Διάγραμμα

Γ.1 Διάγραμμα κύριων λειτουργιών του συστήματος



Παράρτημα Δ

Ερωτηματολόγιο

Δ.1 Ερωτηματολόγιο έρευνας

Ενημερωτικό Σημείωμα προς Συμμετέχοντες/ουσες

Στα πλαίσια της μεταπτυχιακής μου διατριβής, περιγράφεται ένα ταξιδιωτικό σύστημα συστάσεων το οποίο προτείνει στους χρήστες του ταξιδιωτικούς προορισμούς. Το σύστημα αυτό χρειάζεται τα δημογραφικά στοιχεία των χρηστών του (π.χ. ηλικία, φύλο, εκπαίδευση, εισόδημα κ.α.) καθώς και τις προτιμήσεις τους σε ό,τι αφορά τις ταξιδιωτικές τους επιλογές (φύση, μνημεία, κλίμα, κόστος κ.α.) αλλά και από ποιους χρήστες θα θέλανε να προέρχονται οι συστάσεις (π.χ. από χρήστες με ίδιο επίπεδο εκπαίδευσης, ίδιο θρήσκευμα, με ίδιες προτιμήσεις ως προς τον σκοπό του ταξιδιού κ.α.). Διατηρεί ιστορικό των ταξιδιών που έχει πραγματοποιήσει ο χρήστης καθώς και ιστορικό αναζητήσεων του χρήστη.

Διασφάλιση της εμπιστευτικότητας

Το ερωτηματολόγιο αποτελεί μία ανώνυμη διαδικασία κατά την οποία όλες οι πληροφορίες που θα συλλεχθούν θα παραμείνουν αυστηρά εμπιστευτικές. Τα μοναδικά προσωπικά στοιχεία που ζητούνται αφορούν κυρίως δημογραφικά στοιχεία όπως το φύλο, η ηλικία, η εκπαίδευση, το επάγγελμα.

Σας ευχαριστώ για τον χρόνο σας.

Μόκα Ιωάννα

Εκπ/κός ΠΕ 86

Μεταπτυχιακή φοιτήτρια στο Ανοικτό Πανεπιστήμιο Κύπρου στο πρόγραμμα σπουδών "Κοινωνικά Πληροφοριακά Συστήματα".

Τίτλος μεταπτυχιακής διατριβής: "Σύστημα συστάσεων ταξιδιωτικών προορισμών: προσέγγιση συνεργατικού φιλτραρίσματος με βάση τα προφίλ και τις προτιμήσεις των χρηστών".

email: miekbg@gmail.com

Είμαι στη διάθεση σας, οπότε μη διστάσετε να κάνετε ερωτήσεις ή να εκφράσετε τον προβληματισμό σας σχετικά με το ερωτηματολόγιο.

1. Φύλο

- Άνδρας
- Γυναίκα

2. Ηλικία

- 18-30
- 31-40
- 41-50
- 51-60
- 61 και άνω

3. Εκπαίδευση

- Πρωτοβάθμια εκπαίδευση (Δημοτικό σχολείο)
- Δευτεροβάθμια εκπαίδευση (Γυμνάσιο - Λύκειο)
- Ανώτατη εκπαίδευση (Πανεπιστήμιο, ΤΕΙ)
- Μεταπτυχιακό
- Διδακτορικό
- Τίποτα από τα παραπάνω
- Άλλο:

4. Επάγγελμα

- Δημόσιος Υπάλληλος
- Ιδιωτικός Υπάλληλος
- Ελεύθερος επαγγελματίας
- Συνταξιούχος
- Οικιακά
- Φοιτητής/τρια
- Άνεργος/η

5. Πόσο συχνά ταξιδεύετε στο εσωτερικό κατά την διάρκεια ενός έτους;

- Λιγότερο από δύο φορές το χρόνο
- 3 έως 5 φορές το χρόνο
- 6 έως 8 φορές το χρόνο
- Περισσότερο από 8 φορές το χρόνο

6. Πόσο συχνά ταξιδεύετε στο εξωτερικό κατά την διάρκεια ενός έτους;

- Λιγότερο από δύο φορές το χρόνο
- 3 έως 5 φορές το χρόνο
- 6 έως 8 φορές το χρόνο
- Περισσότερο από 8 φορές το χρόνο

7. Θα θέλατε να σας γίνουν ταξιδιωτικές συστάσεις από κάποιον άνθρωπο;

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

8. Θα θέλατε να σας γίνονται ταξιδιωτικές συστάσεις από μια μηχανή;

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

9. Θα προτιμούσατε να σας γίνονται ταξιδιωτικές συστάσεις από έναν άνθρωπο ή από μια μηχανή ;

- Από άνθρωπο
- Από μηχανή
- Και από τα δύο
- Τίποτα από τα παραπάνω

10. Είστε διατεθειμένοι να συμπληρώσετε κάποιο προφίλ στο σύστημα προκειμένου να λαμβάνετε ταξιδιωτικές συστάσεις;

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

11. Είστε διατεθειμένοι να συμπληρώσετε τα δημογραφικά σας στοιχεία στο σύστημα (φύλο, ηλικία, εθνικότητα, εκπαίδευση, εισόδημα, επάγγελμα, θρήσκευμα) προκειμένου να λαμβάνετε ταξιδιωτικές συστάσεις;

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

12. Θα απαντούσατε σε ερωτήσεις του συστήματος σχετικά με παράγοντες που σας επηρεάζουν στις ταξιδιωτικές σας επιλογές; (αν ο προορισμός διαθέτει φυσικά θέλγητρα, πολιτιστικά θέλγητρα, το κόστος ταξιδιού, η διάρκεια ταξιδιού, αίσθημα ασφάλειας στον προορισμό, το κλίμα κλπ.)

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

13. Θα απαντούσατε σε ερωτήσεις του συστήματος σχετικά με από ποιους άλλους χρήστες του συστήματος θα θέλατε να σας γίνονται συστάσεις; (χρήστες ίδιας ηλικιακής ομάδας, ίδιου επιπέδου εκπαίδευσης, ίδιων προτιμήσεων κλπ.)

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

14. Θα απαντούσατε σε ερωτήσεις του συστήματος σχετικά με τις ταξιδιωτικές σας προτιμήσεις; (σκοπός ταξιδιού, διάρκεια, μεταφορικό μέσο, κόστος ταξιδιού κλπ.)

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

15. Θα θέλατε το σύστημα να μαθαίνει τις ταξιδιωτικές σας προτιμήσεις με βάση τις αναζητήσεις σας και τα προηγούμενα ταξίδια που έχετε κάνει;

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

16. Θα θέλατε το σύστημα να σας συγκρίνει με τους υπόλοιπους χρήστες του και να βρίσκει εκείνους με τους οποίους μοιάζετε περισσότερο ώστε να σας παρέχει πιο ακριβείς συστάσεις; (σε σχέση με τις ταξιδιωτικές σας προτιμήσεις, το προφίλ σας, τα προηγούμενα ταξίδια που έχετε κάνει κλπ.)

- Ναι
- Όχι
- Δεν γνωρίζω/Δεν απαντώ

Βιβλιογραφία

- Ardissono, L., Goy, A., Petrone, G., Segnan, M., & Torasso, P. (2003, September). Intrigue: Personalized recommendation of tourist attractions for desktop and hand held devices. *Applied Artificial Intelligence*, 17, σσ. 687-714.
- Asnicar, F. A., & Tasso, C. (1997). ifWeb: a Prototype of User Model-Based Intelligent Agent for Document Filtering and Navigation in the World Wide Web. *Sixth International Conference on User Modeling*, (σσ. 2-5). Sardinia.
- Berka, T., & Plößnig, M. (2004). Designing recommender systems for tourism. ENTER 2004: 11th International Conference on Information Technology in Travel & Tourism.
- Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013, July 1). Recommender systems survey. *Knowledge-Based Systems*, σσ. 109-132.
- Buhalis, D. (2003). *E-tourism. Information Technology for Strategic Tourism Management*. London: Financial Times/Prentice Hall.
- Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism Management*, 29, σσ. 609 - 623.
- Burke, R. (2007). Hybrid Web Recommender Systems. Στο R. Burke, *The Adaptive Web: Methods and Strategies of Web Personalization* (σσ. 377-408). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-72079-9_12.
- Burke, R. (2007). Hybrid Web Recommender Systems. Στο P. Brusilovsky, A. Kobsa, & W. Nejdl, *The Adaptive Web: Methods and Strategies of Web Personalization* (σσ. 377-408). Berlin, Heidelberg: Springer Berlin Heidelberg.

- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: reasoning about a highly connected world* / David Easley, Jon Kleinberg. New York: Cambridge University Press, 2010. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edshlc&AN=edshlc.012657068-X&site=eds-live&authtype=ip,athens>
- García-Crespo, A., Chamizo, J., Rivera, I., Mencke, M., Colomo-Palacios, R., & Gómez-Berbís, J. M. (2009, August). SPETA: Social pervasive e-Tourism advisor. *Telematics and Informatics*, 26, σσ. 306-315.
- Gottapu, R. D., & Sriram Monangi, L. V. (2017, January 1). Point-Of-Interest Recommender System for Social Groups. *Procedia Computer Science*, σσ. 159-164.
- Havaragi, A. B., & Prathibha, S. B. (2016, Apr. - Jun). User-Based Collaborative Filtering System for Tourist Attraction Recommendations. *International Journal of Advanced Research in Computer Science & Technology (IJARCST 2016)*, 4(2), σσ. 194-198.
- Hui, T. K., Wan, D., & Ho, A. (2007). Tourists' satisfaction, recommendation and revisiting Singapore. *Tourism Management*, 28(4), σσ. 965-975.
- Jameson, A. (2004). More Than the Sum of Its Members: Challenges for Group Recommender Systems. *Proceedings of the International Working Conference on Advanced Visual Interface*, (σσ. 48-54). Gallipoli, Italy.
- Jameson, A., Baldes, S., & Kleinbauer, T. (2004). Two methods for enhancing mutual awareness in a group recommender system. (σσ. 447-449). AVI.
- Jeckmans, A., Beye, M., Erkin, Z., Hartel, P., Lagendijk, R., & Tang, Q. (2013). Privacy in Recommender Systems. Στο N. Ramzan, R. van Zwol, J.-S. Lee, K. Clüver, & X.-S. Hua, *Social Media Retrieval* (σσ. 263-281). London: Springer Verlag.

- Kabassi, K. (2010, February). Personalizing recommendations for tourists. *Telematics and Informatics*, 27(1), σσ. 51-66.
- Lam, S. K., Frankowski, D., Riedl, J., & Müller, G. (2006). Do You Trust Your Recommendations? An Exploration of Security and Privacy Issues in Recommender Systems. In *Proceedings of the International Conference on Emerging Trends in Information and Communication Security (ETRICS)*, (σσ. 14-29). Springer Berlin Heidelberg.
- McCarthy, J. F. (2002). Pocket Restaurant Finder: A situated recommender systems for groups. *Proceeding of Workshop on Mobile Ad-Hoc Communication at the 2002 ACM Conference on Human Factors in Computer Systems*.
- Montaner, M., López, B., & de la Rosa, J. L. (2003, 06). A Taxonomy of Recommender Agents on the Internet. *Artificial Intelligence Review*, 19(4), σσ. 285–330.
- Pazzani, M. J. (1999). A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review*, 13(5/6), σσ. 393-408.
- Ricci, F. (2002, November/December). Travel recommender systems. *IEEE Intelligent Systems*, σσ. 55-57.
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. B. (2011). *Recommender Systems Handbook*. New York: Springer.
- Wang, Y., Chan, S. C.-F., & Ngai, G. (2012). Applicability of Demographic Recommender System to Tourist Attractions: A Case. *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*. <https://doi.org/10.1109/WI-IAT.2012.133>.
- Xiaoyuan, S., & Taghi, K. M. (2009, January). A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence, 2009*, σσ. 1-20.

Ηγουμενάκης, Ν. Γ., Κραβαρίτης, Κ. Ν., & Λύτρας, Π. Ν. (1999). *Εισαγωγή στον Τουρισμό*. Αθήνα: Interbooks.

Σιγάλα, Μ. (2008). *Νέες Τεχνολογίες στον Τουρισμό*. Αθήνα: Εθνική Σχολή Δημόσιας Διοίκησης.