

**Πολυτεχνείο Κρήτης**  
**Σχολή Μηχανικών Παραγωγής και Διοίκησης**



**ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ**

**Ανάπτυξη Πολυκριτήριου Συστήματος Συστάσεων  
για προσωποποιημένο Μάρκετινγκ**  
(Development of Multicriteria Recommender System for  
personalized Marketing)

**Ιωάννης Ζήσος**

Χανιά - Ιούνιος 2018

## **Μέλη της επταμελούς εξεταστικής Επιτροπής**

**Νικόλαος Ματσατσίνης**, Καθηγητής (Επιβλέπων - Μέλος Τριμελούς Επιτροπής)

Σχολή Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης

**Ευάγγελος Γρηγορούδης**, Καθηγητής (Μέλος Τριμελούς Επιτροπής)

Σχολή Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης

**Αθανάσιος Σπυριδάκος**, Καθηγητής (Μέλος Τριμελούς Επιτροπής)

Σχολή Διοίκησης και Οικονομίας, Πανεπιστήμιο Δυτικής Αττικής

**Μάρω Βλαχοπούλου**, Καθηγήτρια

Σχολή Επιστημών Πληροφορίας, Πανεπιστήμιο Μακεδονίας

**Πάυλος Δελιάς**, Αναπληρωτής Καθηγητής

Τμήμα Λογιστικής και Χρηματοοικονομικής, ΤΕΙ Ανατολικής Μακεδονίας-Θράκης

**Στυλιανός Τσαφάρáκης**, Επίκουρος Καθηγητής

Σχολή Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης

**Νικόλαος Τσότσολας**, Επίκουρος Καθηγητής

Σχολή Διοίκησης και Οικονομίας, Πανεπιστήμιο Δυτικής Αττικής

Η παρούσα Διδακτορική διατριβή υπεβλήθη για τη μερική εκπλήρωση των υποχρεώσεων απόκτησης του Διδακτορικού Διπλώματος του Πολύτεχνείου Κρήτης.

Η έγκριση της Διδακτορικής Διατριβής, από τη Σχολή Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα. (Ν. 5343/1932)

Copyright © Ιωάννης Α. Ζήσος, 2018

All rights reserved. Με επιφύλαξη παντός δικαιώματος.

Στους γονείς μου,  
Θανάση και Μαρία  
με αγάπη.

--- “*ὅτι οὐδὲν οἶδα*”, Σωκράτης, 469- 399 π.Χ---

## Ευχαριστίες

Με την ολοκλήρωση της διατριβής μου, θα ήθελα να απευθύνω τα σέβη και τις ευχαριστίες μου σε όσους στάθηκαν δίπλα μου σ' αυτή την επίπονη και μακρά προσπάθεια και με βοήθησαν να την φέρω σε πέρας.

Θα ήθελα να εκφράσω τη βαθιά και απεριόριστη εκτίμηση μου για τον επιβλέποντα καθηγητή Δρ. Νικόλαο Ματσατσίνη καθώς και τους καθηγητές Δρ. Ευάγγελο Γρηγορούδη και Δρ. Αθανάσιο Σπυριδάκο, μέλη της τριμελούς επιτροπής της Διδακτορικής Διατριβής μου. Καθόλη τη διάρκεια της πορείας μου προς την ολοκλήρωση της έρευνας, ήταν δίπλα μου, προσφέροντας την αμέριστη συμπαράσταση και γνώση τους. Οι πόρτες τους ήταν πάντα ανοιχτές και οι συμβουλές τους σε ακαδημαϊκό και προσωπικό επίπεδο υπήρξαν καταλυτικές. Τους ευχαριστώ εκ βάθέων και τους ευγνωμονώ.

Τις θερμές μου ευχαριστίες θα ήθελα να εκφράσω, στα μέλη της εξεταστικής επιτροπής, καθ. Μ. Βλαχοπούλου και αν. καθ. Π. Δελιά για το χρόνο που αφιέρωσαν καθώς και για την αποδοχή της πρόσκλησης να συμμετέχουν στην επιτροπή. Ιδιαίτερα θα ήθελα να απευθυνθώ στον επ. Καθ. Στέλιο Τσαφάρακη, και να τον ευχαριστήσω για τις ατελείωτες επιστημονικές και προσωπικές συζητήσεις που κάναμε κατά την περίοδο της διαμονής μου στα Χανιά. Επίσης, οφείλω να ευχαριστήσω τον επ. Καθ. Νίκο Τσότσολα, που μου προσέφερε σημαντικές λύσεις στο τελευταίο στάδιο της εκπόνησης της ΔΔ.

Δεν θα μπορούσα να μην αναφερθώ στην αμέριστη βοήθεια που έλαβα από τον καθ. Λάζαρο Βρυζίδα και να τον ευχαριστήσω από καρδιάς.

Οι αδελφικοί μου φίλοι Άγγελος, Αναστασία, Παναγιώτης, Τάσος και Θοδωρής ήταν συνοδοιπόροι μου σε αυτό το ταξίδι και αυτοί με τους οποίους μοιράστηκα τον ψυχισμό μου. Για αυτό τους ευγνωμονώ.

Θα ήθελα τέλος να εκφράσω το πιο μεγάλο ευχαριστώ στους γονείς μου, Θανάση και Μαρία και τον αδελφό μου Στέλιο, που από την αρχή της διαδρομής αποτέλεσαν τους πυλώνες έμπνευσης, δύναμης και κινήτρου μου.

## Περίληψη

Ένα από τα πιο καίρια ερωτήματα για τους χρήστες του διαδικτύου, είναι πώς θα καταφέρουν να διαχειριστούν την τεράστια ποσότητα διαθέσιμης πληροφορίας, ώστε να καταλήξουν σε επιλογή προϊόντων που ανταποκρίνονται όσο το δυνατό καλύτερα στις προτιμήσεις και ανάγκες τους. Αντίστοιχα, οι εταιρείες που παρέχουν προϊόντα ή υπηρεσίες μέσω του διαδικτύου, προσπαθούν συστηματικά να εντοπίσουν μεθόδους ώστε να αποκωδικοποιήσουν με ακρίβεια τα προφίλ προτίμησης των χρηστών, με στόχο να καταφέρουν να προσαρμόσουν κατάλληλα τα προϊόντα τους και να αυξήσουν τις πωλήσεις τους.

Για τους παραπάνω λόγους, η επιστημονική και ερευνητική κοινότητα που δραστηριοποιείται στο τομέα της ανάλυσης δεδομένων και το μάρκετινγκ έχει επικεντρώσει την προσπάθεια της, στην δημιουργία μεθοδολογιών που θα απαντήσουν όσο το δυνατόν πιο αποτελεσματικά τα παραπάνω ερωτήματα. Οι περισσότερες από αυτές τις μεθοδολογίες καταλήγουν στην ανάπτυξη προσαρμοστικών συστημάτων που αντλούν δεδομένα από το διαδίκτυο και εξάγουν προτάσεις για τους χρήστες. Η πιο γνωστή κατηγορία τέτοιου είδους συστημάτων είναι τα συστήματα συστάσεων (Recommender Systems).

Στην παρούσα ερευνητική εργασία παρουσιάζεται η μεθοδολογία και τα αποτελέσματα πιλοτικής λειτουργίας ενός νέου υβριδικού συστήματος συστάσεων που βασίζεται στη χρήση μεθόδων ανάλυσης συναισθήματος, πολυκριτήριας ανάλυσης καθώς και μεθόδων φιλτραρίσματος. Η μεθοδολογία καταλήγει σε τέσσερα διαφορετικά είδη σύστασης, με άκρως ενδιαφέροντα αποτελέσματα.

Μέσω του μεθοδολογικού πλαισίου γίνεται εφικτός ο προσδιορισμός των προτιμησιακών προφίλ των χρηστών του συστήματος, τα οποία εν συνεχεία αντιστοιχίζονται σε «προφίλ πελατών» που επιλέγουν συγκεκριμένα προϊόντα/υπηρεσίες που τους «ταιριάζουν».

Έτσι, καταλήγουμε σε προσωποποιημένες συστάσεις προϊόντων στον χρήστη του συστήματος, που είναι ανάλογες των προτιμήσεων του. Επιπλέον δίνεται στο χρήστη η δυνατότητα να φιλτράρει τις διαθέσιμες εναλλακτικές με σχετική επιλογή από ένα σύνολο κατ' αποκοπή κριτηρίων. Η χρήση του κατωφλιού

ελάχιστης ικανοποίησης, που προσδιορίζεται από τα αποτελέσματα της ανάλυσης συναισθήματος στα σχόλια των πελατών, εγγυάται την ποιότητα των συστάσεων.

Τα δεδομένα του συστήματος είναι πραγματικές απόψεις και βαθμολογίες χρηστών για καταλύματα, καθώς και χαρακτηριστικά καταλυμάτων που αντλήθηκαν από γνωστή διαδικτυακή πλατφόρμα κρατήσεων.

Η ανάπτυξη του συστήματος βασίστηκε στην μεθοδολογία CRISP-DM (Shearer, 2000a).

Η αξιολόγηση του συστήματος συστάσεων γίνεται με μέτρηση της ακρίβειας προβλέψεων αξιολογήσεων σε πείραμα με πραγματικούς χρήστες.

Για τη μελέτη περίπτωσης χρησιμοποιήθηκαν δεδομένα για τα τουριστικά καταλύματα του Νομού Χανίων. Τα πιο σημαντικά χαρακτηριστικά της συγκεκριμένης ερευνητικής προσπάθειας είναι:

α. Η χρήση μεγάλου όγκου πραγματικών δεδομένων σε αντίθεση με τις περισσότερες από τις υπάρχουσες έρευνες που χρησιμοποιούν έτοιμα τεστ σετ δεδομένων.

β. Η χρησιμοποίηση όλης της διαθέσιμης αντλημένης πληροφορίας προκειμένου να καταλήξουμε σε σύσταση. Πιο συγκεκριμένα στην παρούσα μεθοδολογία χρησιμοποιούνται δεδομένα βαθμολογιών προϊόντων για την ανάλυση ικανοποίησης πελατών, δεδομένα απόψεων για τη συναισθηματική ανάλυση, στατικά δεδομένα των προϊόντων σαν κατ' αποκοπή κριτήρια σε αντίθεση με τις περισσότερες μελέτες όπου χρησιμοποιούνται είτε βαθμολογίες, είτε ανάλυση σχολίων για την τελική σύσταση.

γ. Η χρήση των αποτελεσμάτων της ανάλυσης συναισθήματος ως κατώφλια ποιότητας σύστασης.

δ. Η επιτυχής εφαρμογή παραλλαγής της πρόσφατα παρουσιασθείσας μεθόδου WAP, για τη δημιουργία προτιμησιακού προφίλ χρήστη.

ε. Η απαίτηση για εισαγωγή ελάχιστων δεδομένων από το χρήστη.

στ. Ο ελάχιστος χρόνος που απαιτείται για την παραγωγή της σύστασης κατά τη λειτουργία του συστήματος.

η. Η αποφυγή άντλησης προσωπικών δεδομένων του χρήστη για χρήση τους στη διαδικασία της σύστασης.

θ. Η αποφυγή του προβλήματος της καθυστερημένης εκκίνησης (cold start).

ι. Τα άκρως ικανοποιητικά αποτελέσματα με βάση τις μετρικές που χρησιμοποιήθηκαν για τη δοκιμή του συστήματος.

κ. Η χρησιμότητα του συστήματος είναι χαρακτηριστική για όλες τις ομάδες στόχου, καθώς δίνεται η δυνατότητα παροχής χρήσιμης πληροφορίας τόσο στον πελάτη με σύσταση ανάλογη των αναγκών του, όσο και στον πάροχο υπηρεσίας/προϊόντος, προσδιορίζοντας του τις τάσεις όσον αφορά την ικανοποίηση των πελατών.

Οι κύριες μετρικές που χρησιμοποιούνται για την μέτρηση της ποιότητας των συστάσεων που παράγονται είναι οι: Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE).

Η διατριβή ολοκληρώνεται, με προτάσεις, για μελλοντική έρευνα και επέκταση του παρόντος συστήματος.

## Δημοσιεύσεις σε επιστημονικά περιοδικά και συνέδρια

I. Zisos, E. Grigoroudis, N. Tsotsolas, N. Matsatsinis, **“Recommending Hotels, a new approach: The case of Chania region.” “Marketing in the era of Social Media and Big Data: New Trends and Applications”**, Springer book. (Submitted)

I. Zisos, A. Spyridakos, E. Grigoroudis, N. Matsatsinis, **“Combining Multicriteria Methods with Sentiment Analysis to personalize recommendations”**, Operational Research: An International Journal. (Submitted)

I. Zisos, N. Matsatsinis, N. Tsotsolas, **“A Hybrid Multicriteria Hotel Recommender System using online ratings and reviews”**, 7th International Symposium and 29th National Conference on Operational Research “The contribution of OR, new technologies and innovation in agriculture and tourism”, p88-93, ISBN: 978-618-80361-7-8, June 14-16 2018, Chania, Crete, Greece.

I. Zisos, Grigoroudis E., Matsatsinis N., Spyridakos A., **“Diving into online ratings to determine hotels’ improvement priorities”**, International Journal of Decision Support Systems, Special Issue on: “Decision Support Systems in Commerce, Logistics and Transportation, May 2018. (Received Acceptance)

I. Zisos, A. Spyridakos, E. Grigoroudis, N. Matsatsinis **“Υβριδικό Σύστημα Συστάσεων για επιλογή τουριστικού καταλύματος”**, 16ο Ειδικό Συνέδριο της Ε.Ε.Ε.Ε. και 12η Συνάντηση Πολυκριτήριας Ανάλυσης Αποφάσεων “Εφαρμογές της Πολυκριτήριας Ανάλυσης στη Ναυτιλία και τις Μεταφορές”, Εμπορικό και Βιομηχανικό Επιμελητήριο Πειραιά, 15-17 Φεβρουαρίου 2018.

Zisos I., Grigoroudis E., Matsatsinis N., Spiridakos A. , **“Using online ratings to define Hotels Improvement Priorities”**, 14th Special Conference of the Hellenic Operational Research Society- 11th Meeting of Multicriteria Decision Analysis, Agrinio, March 2015



# ΠΕΡΙΕΧΟΜΕΝΑ

Ευχαριστίες .....	iv
Περίληψη .....	v
Δημοσιεύσεις σε επιστημονικά περιοδικά και συνέδρια .....	viii
Σχήματα .....	xi
Πίνακες .....	xiii
Γραφικές Παραστάσεις .....	xiv
<b>1. Εισαγωγή .....</b>	<b>1</b>
1.1. Περιγραφή του ερευνητικού προβλήματος .....	1
1.2. Ερευνητικά ζητήματα που διερευνώνται και ομάδες στόχου .....	4
1.3. Πρωτοτυπία και χρησιμότητα της έρευνας.....	5
<b>2. Υφιστάμενη Κατάσταση &amp; Θεωρητικό Υπόβαθρο .....</b>	<b>18</b>
2.1. Απόψεις/κριτικές και Αξιολογήσεις .....	18
2.2. Συστήματα Συστάσεων .....	24
2.2.1. Κατηγορίες Συστημάτων Συστάσεων .....	27
2.2.2. Πολυκριτήρια Συστήματα αποφάσεων .....	34
2.2.3. Η Σύσταση ως ένα πολυκριτήριο πρόβλημα λήψης απόφασης.....	42
2.2.3.4.1 Χαρακτηριστικά και μετρικές αξιολόγησης Συστήματος Συστάσεων .....	54
2.3. Μοντελοποίηση Προτίμησης.....	59
2.3.1. Καθορισμός βαρών μέσω ιεράρχησης προτεραιότητας (WAP).....	59
2.4. Ικανοποίηση Πελατών .....	67
2.4.1. Πολυκριτήρια Ανάλυση Ικανοποίησης (MUSA).....	69
2.5. Ανάλυση περιεχομένου και συναισθήματος (Content and Sentiment analysis).....	83
2.5.1. Ταξινόμηση συναισθήματος σε επίπεδο εγγράφου .....	89
2.5.2. Ταξινόμηση συναισθήματος με εποπτευόμενη μάθηση .....	90
2.5.3. Ταξινόμηση συναισθήματος με χρήση μη-εποπτευόμενης μάθησης .....	91
2.5.4. Ταξινόμηση Υποκειμενικότητας .....	91
<b>3. Μεθοδολογία Συστήματος .....</b>	<b>93</b>
3.1. Σχηματικό Διάγραμμα.....	95
3.2. Ανάλυση και Επεξεργασία Δεδομένων .....	99

3.3	Εφαρμογή MUSA .....	107
3.4	Εφαρμογή Ανάλυσης Συναισθήματος .....	109
3.5	Εφαρμογή Μοντελοποίησης Προτίμησης.....	112
3.6	Τελική Σύσταση .....	115
3.7	Μηχανισμός Ανατροφοδότησης.....	123
3.8	Συνοπτική Μαθηματική παρουσίαση της μεθοδολογίας.....	126
<b>4.</b>	<b>Σχεδίαση και Ανάπτυξη συστήματος .....</b>	<b>134</b>
4.1	Δεδομένα .....	134
4.2	Εφαρμογή MUSA .....	139
4.3	Εφαρμογή Sentiment analysis.....	142
4.4	Εφαρμογή WAP .....	143
4.5	Τελική Σύσταση .....	146
<b>5.</b>	<b>Παρουσίαση του συστήματος για μελέτη περίπτωσης: Τουριστικά Καταλύματα .....</b>	<b>150</b>
<b>6.</b>	<b>Αξιολόγηση αποτελεσμάτων.....</b>	<b>165</b>
<b>7.</b>	<b>Συμπεράσματα .....</b>	<b>172</b>
7.1	Περιορισμοί.....	177
7.2	Προτάσεις για μελλοντική έρευνα.....	178
<b>8.</b>	<b>Βιβλιογραφία .....</b>	<b>180</b>
<b>9.</b>	<b>Παράρτημα.....</b>	<b>189</b>
	<b>Σύντομο Βιογραφικό Σημείωμα .....</b>	<b>202</b>

## Σχήματα

Σχ. 2-1 Φάσεις Ανάπτυξης Πληροφοριακού Συστήματος Υποστήριξης Αποφάσεων.....	39
Σχ. 2-2 Σχηματικό Διάγραμμα της λειτουργίας της Μεθόδου WAP .....	66
Σχ. 2-3 Ράβδοι κύλισης για προσδιορισμό της διαφοράς σημαντικότητας κριτηρίων με $Z_{min}$ , $Z_{max}$ .....	66
Σχ. 2-4 Ολική ικανοποίηση .....	70
Σχ. 2-5 Ομάδες πελατών με διαφορετικό βαθμό απαιτητικότητας.....	75
Σχ. 2-6 Διάγραμμα Δράσης .....	76
Σχ. 2-7 Διάγραμμα Βελτίωσης.....	79
Σχ. 3-1 Διάγραμμα Φάσεων Crisp DM.....	94
Σχ. 3-2 Σχηματικό Διάγραμμα λειτουργιών νέου Συστήματος Συστάσεων.....	96
Σχ. 3-3 Spreadsheet με αποτελέσματα άντλησης.....	101
Σχ. 3-4 Συναρτήσεις για επεξεργασία των αρχικών δεδομένων.....	102
Σχ. 3-5 Μορφή δεδομένων κατόπιν της επεξεργασίας.....	102
Σχ. 3-6 Χαρακτηριστικά ξενοδοχείου στην αρχική μορφή άντλησης .....	103
Σχ. 3-7 Κατηγοριοποιημένα Χαρακτηριστικά ανά ξενοδοχείο .....	103
Σχ. 3-8 Παράδειγμα κριτικής και βαθμολογιών πελάτη από διαδικτυακή πλατφόρμα .....	105
Σχ. 3-9 Παράδειγμα On/Off κριτηρίων από διαδικτυακή πλατφόρμα .....	105
Σχ. 3-10 Σχηματικό διάγραμμα εισόδων-εξόδων MUSA στην μεθοδολογία.....	108
Σχ. 3-11 Σχηματικό διάγραμμα εισόδων-εξόδων ανάλυσης συναισθήματος στην μεθοδολογία.....	112
Σχ. 3-12 Σχηματικό διάγραμμα εισόδων-εξόδων WAP στη μεθοδολογία.....	115
Σχ. 3-13 Σύνολα εναλλακτικών P, Pfilt μετά την επιλογή On/Off από το χρήστη .....	118
Σχ. 3-14 Διαδικασία και αποτελέσματα WAP στην μεθοδολογία .....	119
Σχ. 3-15 Σύνολα εναλλακτικών P, PWESU μετά την εφαρμογή WAP.....	120
Σχ. 3-16 Σύνολα και υποσύνολα εναλλακτικών P, Pwesu, Pse .....	121
Σχ. 3-17 Σύνολα και υποσύνολα εναλλακτικών P, Pwesu, Pse, Pfilt.....	122
Σχ. 3-18 Υποσύνολο συστάσεων Prec και σύνολα εναλλακτικών P, Pwesu, Pwesuse, Pfilt.....	122
Σχ. 3-19 Εγγραφές σχετικές με τις επιλογές του χρήστη συστήματος συστάσεων .....	124
Σχ. 3-20 Αποτελέσματα συστάσεων και ανατροφοδότηση.....	125
Σχ. 3-21 Στιγμιότυπο εγγραφής στη βάση δεδομένων .....	125
Σχ. 4-1 Ράβδοι κύλισης για υπολογισμό $Z_{min}$ , $Z_{max}$ στη WAP .....	144
Σχ. 5-1 Σελίδα Εισόδου του συστήματος συστάσεων .....	151
Σχ. 5-2 Διαδικασία login στο σύστημα συστάσεων .....	152
Σχ. 5-3 Σελίδα ON/OFF filtering .....	153
Σχ. 5-4 Επιλογές τιμών για τα κατ' αποκοπή κριτήρια.....	154
Σχ. 5-5 Σελίδα παρουσίασης αποτελεσμάτων εναλλακτικών μετά το On/ Off Filtering .....	155

Σχ. 5-6 Σελίδα εισαγωγής επιλογών κατάταξης και διαφοράς σημαντικότητας κριτηρίων .....	156
Σχ. 5-7 Κατάταξη κριτηρίων και μετάβαση στην επιλογή διαφοράς σημαντικότητας .....	157
Σχ. 5-8 Επιλογή διαφορών σημαντικότητας μεταξύ συνεχόμενων κριτηρίων ...	158
Σχ. 5-9 Σελίδα παρουσίασης αποτελεσμάτων των 4 τύπων συστάσεων για χρήστη admin .....	159
Σχ. 5-10 Παρουσίαση συστάσεων .....	160
Σχ. 5-11 Σελίδα παρουσίασης αποτελεσμάτων συστάσεων Prec και πληροφοριών εναλλακτικής .....	161
Σχ. 5-12 Σελίδα με charts .....	162
Σχ. 5-13 Σελίδα με charts με αποτελέσματα .....	162
Σχ. 5-14 Σελίδα με αναλυτικά δεδομένα για κάθε εναλλακτική.....	163
Σχ. 5-15 Σελίδα για ανατροφοδότηση από το χρήστη .....	164

## Πίνακες

Πίνακας 1-1 Καταγραφή των σημαντικότερων μεθοδολογιών και συστημάτων Συστάσεων .....	6
Πίνακας 2-1:Υπολογισμός βαρών 12 κριτηρίων με τη μέθοδο Simos.....	61
Πίνακας 3-1 Δυναμικά Δεδομένα .....	104
Πίνακας 3-2 Στατικά δεδομένα .....	105
Πίνακας 3-3 Δεδομένα εισόδου MUSA .....	107
Πίνακας 3-4 Αποτελέσματα MUSA .....	108
Πίνακας 3-5 Είσοδοι για την εφαρμογή ανάλυσης συναισθήματος.....	111
Πίνακας 3-6 Έξοδοι από την ανάλυση συναισθήματος.....	111
Πίνακας 3-7 Έξοδοι εφαρμογής WAP .....	114
Πίνακας 3-8 Αποτελέσματα φάσης άντλησης και επεξεργασίας δεδομένων .....	115
Πίνακας 3-9 Αποτελέσματα MUSA .....	116
Πίνακας 3-10 Αποτελέσματα ανάλυσης συναισθήματος.....	116
Πίνακας 3-11 Αποτελέσματα μοντελοποίησης προτίμησης.....	117
Πίνακας 3-12 Επιλογές On/Off από το χρήστη .....	118
Πίνακας 4-1 Αριθμός κριτικών και ξενοδοχείων ανά κατηγορία ξενοδοχείων που αντλήθηκαν.....	136
Πίνακας 4-2 Στατικά δεδομένα που αντλήθηκαν για τα ξενοδοχεία.....	136
Πίνακας 4-3 Δεδομένα βαθμολογιών και κριτικών πελατών που αντλήθηκαν .	138
Πίνακας 4-4 Μορφή της κριτικής κατόπιν επεξεργασίας .....	139
Πίνακας 4-5 Αριθμός Ξενοδοχείων και κριτικών ανά ομάδα .....	140
Πίνακας 4-6 Αποτελέσματα MUSA για ένα ξενοδοχείο του δείγματος.....	141
Πίνακας 4-7 Αποτελέσματα ανάλυσης συναισθήματος για 2 ξενοδοχεία του δείγματος .....	142
Πίνακας 4-8 Μέσοι όροι των $Z_{min}$ , $Z_{max}$ με βάση τα αποτελέσματα του ερωτηματολογίου .....	145
Πίνακας 4-9 Αποτελέσματα WAP για ένα τυχαίο χρήστη.....	145
Πίνακας 6-1 Αποτελέσματα Μετρικών σφάλματος για συνολική βαθμολογία .	166
Πίνακας 6-2 Πλήθος βαθμολογιών ανά εύρος ποσοστιαίων σφαλμάτων .....	168
Πίνακας 9-1 Αποτελέσματα ερωτηματολογίου για τις τιμές $Z_{min}$ , $Z_{max}$ .....	190
Πίνακας 9-2: Βαθμολογίες των 73 χρηστών του πειράματος για 5 κριτήρια και συνολική βαθμολογία .....	191
Πίνακας 9-3 Μέσοι όροι Βαθμολογιών κριτηρίων και συνολικών βαθμολογιών για ξενοδοχεία .....	194

## Γραφικές Παραστάσεις

Γραφική 2-1 Διάγραμμα διακύμανσης ολικής συνάρτησης ικανοποίησης .....	82
Γραφική 6-1 Ποσοστιαία σφάλματα μεταξύ βαθμολογιών χρήστη και ξενοδοχείου .....	168
Γραφική 6-2 Σύγκριση βαθμολογιών από ανατροφοδότηση χρηστών με βαθμολογίες ξενοδοχείων .....	169
Γραφική 9-1 Ποσοστιαίο Σφάλμα για overall Rating .....	196
Γραφική 9-2 Ποσοστιαίο Σφάλμα για Καθαριότητα .....	197
Γραφική 9-3 Ποσοστιαίο Σφάλμα για την Εξυπηρέτηση .....	197
Γραφική 9-4 Ποσοστιαίο Σφάλμα για Άνεση .....	198
Γραφική 9-5 Ποσοστιαίο Σφάλμα για Κατάσταση.....	198
Γραφική 9-6 Ποσοστιαίο Σφάλμα για Γειτονιά.....	199
Γραφική 9-7 Μέτρηση τιμών PE για τιμές -20% ως 20% για overall Rating.....	199
Γραφική 9-8 Μέτρηση τιμών PE για τιμές -20% ως 20% για Καθαριότητα .....	200
Γραφική 9-9 Μέτρηση τιμών PE για τιμές -20% ως 20% για Εξυπηρέτηση.....	200
Γραφική 9-10 Μέτρηση τιμών PE για τιμές -20% ως 20% για Άνεση.....	201
Γραφική 9-11 Μέτρηση τιμών PE για τιμές -20% ως 20% για Κατάσταση .....	201
Γραφική 9-12 Μέτρηση τιμών PE για τιμές -20% ως 20% για Γειτονιά.. .....	202

<b>1.</b>	<b>Εισαγωγή</b>
1.1.	Περιγραφή του ερευνητικού προβλήματος
1.2	Ερωτήματα που διερευνώνται και ομάδες στόχου
1.3	Πρωτοτυπία και χρησιμότητα της έρευνας

## **1. Εισαγωγή**

### **1.1. Περιγραφή του ερευνητικού προβλήματος**

Ένας από τους βασικούς στόχους της επιστημονικής και ερευνητικής κοινότητας που απασχολείται με τον κλάδο της ροής πληροφορίας στο διαδίκτυο (Jackson,n.d., Benkler, 2006, Freeman, 2000), καθώς και το διαδικτυακό marketing, είναι η βελτιστοποίηση της υποστήριξης των χρηστών μέσα από τη διάχυση της κατάλληλης πληροφορίας που ταυτίζεται με το κοινωνικό προφίλ τους, τις προτιμήσεις και τα ενδιαφέροντα τους. Η παροχή διαδικτυακού περιβάλλοντος και επιλογών στους χρήστες, προσαρμοζόμενου στις ανάγκες τους αποτελεί κρίσιμο παράγοντα για την αποτελεσματική αξιοποίηση του διαδικτύου για αγορά προϊόντων και υπηρεσιών. Η δυσκολία του συγκεκριμένου εγχειρήματος αυξάνει αν λάβουμε υπ' όψιν την τεράστια ποσότητα πληροφορίας που είναι διαθέσιμη στον παγκόσμιο ιστό, καθώς και την εκθετική αύξηση των χρηστών και των δεδομένων που είναι διαθέσιμα και προσπελάσιμα στο διαδίκτυο (Mirilas, 2010).

Η αποτελεσματική ομαδοποίηση των χρηστών με βάση την ανάλυση της συμπεριφοράς τους, τα ενδιαφέροντά τους και τις προτιμήσεις τους, αποτελεί καθοριστικής σημασίας αντικείμενο έρευνας στον τομέα του Μάρκετινγκ. Εμπορικές και παραγωγικές επιχειρήσεις σχεδιάζουν τις στρατηγικές και επιτελικές επιλογές τους βασιζόμενες στην καταγραφή και αποτύπωση της άποψης του «κοινού» στο πλαίσιο στοχευμένων εμπορικών δραστηριοτήτων τους. Ένα μεγάλο ποσοστό του «κοινού» σήμερα χρησιμοποιεί και αξιοποιεί τα συστήματα Κοινωνικής Δικτύωσης. Το ποσοστό αυτό συνεχώς αυξάνεται και η ανάγκη της παροχής αποτελεσματικής προσωποποιημένης πληροφορίας που

ικανοποιεί τις προτιμήσεις του «κοινού» (Raad et al., 2011, European Commission and Directorate-General for the Information Society and Media, 2010) γίνεται όλο και πιο καθοριστική τόσο για τις επιχειρήσεις όσο και για τους ίδιους τους χρήστες.

Είναι επίσης ευρέως γνωστό, ότι ακόμα και εταιρείες κολοσσοί του διαδικτύου, βρίσκονται σε μια συνεχή διαδικασία διερεύνησης και άντλησης όσο το δυνατόν περισσότερων λεπτομερειών για τα προφίλ των χρηστών, προκειμένου να σκιαγραφήσουν το προφίλ των αναγκών τους αλλά και της προσωπικότητας τους. Δυστυχώς σε αρκετές περιπτώσεις η διαδικασία αυτή καταλήγει στην παραβίαση κανόνων προσωπικών δεδομένων, ενώ η προσπάθεια διείσδυσης γίνεται όλο και εντονότερη.

Στόχος της παρούσας διδακτορικής διατριβής είναι ο αξιόπιστος και άμεσος προσδιορισμός των χαρακτηριστικών και προτιμήσεων χρηστών και διαδικτυακών κοινοτήτων ούτως ώστε, με τη χρήση μεθόδων συναισθηματικής και πολυκριτήριας ανάλυσης, να προκύπτουν ακριβείς και προσωποποιημένες συστάσεις προϊόντων/υπηρεσιών σε χρήστες ή κοινότητες πελατών.

Στα πλαίσια της παρούσας έρευνας, δε γίνεται καμία προσπάθεια άντλησης προσωπικών δεδομένων ή διείσδυσης στην προσωπική ζωή των χρηστών. Τα προτιμησιακά μοντέλα που δομούνται, στηρίζονται σε επιλογές των χρηστών που γίνονται με τη θέληση τους, και όλα τα δεδομένα που χρησιμοποιούνται είναι δημόσια.

Μέσω του μεθοδολογικού πλαισίου γίνεται εφικτός ο προσδιορισμός των δυναμικών προφίλ των χρηστών του διαδικτύου και των αντικειμένων ενδιαφέροντος τους όπως είναι και η περίπτωση της αγοράς συγκεκριμένων προϊόντων και υπηρεσιών έτσι ώστε να υποστηριχθούν οι λειτουργίες του προσωποποιημένου μάρκετινγκ (Lakiotaki et al., 2011, Lakiotaki and Matsatsinis, 2012, Manouselis' and Matsatsinis, 2001, Manouselis and Costopoulou, 2007a, Grigoroudis and Siskos, 2010, Antunes and Costa, 2012) για στοχευμένες ομάδες δυνητικών πελατών. Τα προφίλ προτιμήσεων των χρηστών αντιστοιχίζονται σε «προφίλ» προϊόντων/υπηρεσιών που εμπεριέχουν τα χαρακτηριστικά που αποτελούν προτεραιότητα για τους συγκεκριμένους χρήστες και εν συνεχεία



προτείνονται σε αυτούς. Η συγκεκριμένη ερευνητική προσπάθεια επιχειρεί να βοηθήσει στην επίλυση των εξής προβλημάτων:

Για τους πελάτες:

1. Την δυσκολία επιλογής προϊόντων/υπηρεσιών, λόγω της τεράστιας ποσότητας πληροφορίας που είναι διαθέσιμη στο διαδίκτυο, με αποτέλεσμα την πλειονότητα των φορών να κατευθύνονται σε επιλογές διαφορετικές από το προτιμησιακό τους μοντέλο, που τελικά δεν τους ικανοποιούν.
2. Την παροχή συστάσεων για αγορά προϊόντων/υπηρεσιών στους πελάτες, που προκύπτουν από επεξεργασία δεδομένων που αντλούνται εν αγνοία τους, με βάση τις σελίδες που επισκέπτονται στο διαδίκτυο και αρκετές φορές παραβιάζουν το προσωπικό απόρρητο. Με βάση την ερευνητική μεθοδολογία που αναπτύσσεται στη διατριβή, ο χρήστης καλείται να παρέχει με τη θέληση του ελάχιστες πληροφορίες που αποτελούν επιλογές/προτιμήσεις του σε ποιοτικά κριτήρια.
3. Την παροχή συστάσεων με βάση μεμονωμένες ή περιορισμένου εύρους έρευνες αγοράς, καθώς τα προτιμησιακά μοντέλα προκύπτουν από μεγάλο όγκο δεδομένων που αποτελούν τόσο ποσοτικές όσο και ποιοτικές μεταβλητές (ratings και reviews).
4. Τον αυξημένο χρόνο που καταναλώνουν προκειμένου να ερευνήσουν περισσότερες επιλογές που ίσως να τους ικανοποιούν.

Για τους παρόχους υπηρεσιών/προϊόντων:

1. Την ελλιπή ή μεμονωμένη πληροφορία που έχουν στη διάθεση τους σχετικά με τις προτιμήσεις των πελατών ή των κατηγοριών πελατών. Αυτό έχει σαν αποτέλεσμα την μη σωστή κατανομή των πόρων της εταιρείας για την παραγωγή ή/και την προώθηση προϊόντων/υπηρεσιών τους.
2. Την επένδυση σε τομείς των υπηρεσιών που δεν έχουν αντίκτυπο στους πελάτες ή δεν έχουν την αναμενόμενη επιστροφή κεφαλαίου.

3. Το χρόνο που καταναλώνουν προκειμένου να διερευνήσουν τις πιο ελκυστικές για τους πελάτες υπηρεσίες ή παραμέτρους προϊόντων.

## **1.2 Ερευνητικά ζητήματα που διερευνώνται και ομάδες στόχου**

Τα βασικά ερευνητικά ζητήματα που διερευνήθηκαν στα πλαίσια της παρούσας διατριβής είναι τα εξής:

1. Ανάλυση και συσχέτιση των αλληλεπιδράσεων και των χαρακτηριστικών των πελατών-χρηστών διαδικτύου για δημιουργία προτύπων συμπεριφοράς χρηστών (Lakiotaki et al., 2009).
2. Ομαδοποίηση και τμηματοποίηση του τεράστιου πλήθους πελατών που είναι διαδικτυακοί χρήστες καθώς και των παρεχόμενων προϊόντων και υπηρεσιών, με βάση τα κοινωνικά χαρακτηριστικά και τις προτιμήσεις τους.
3. Η βελτίωση της διαδραστικότητας μεταξύ των τμημάτων marketing παρόχων υπηρεσιών/προϊόντων με τους πελάτες και η επίτευξη πιο στοχοθετημένου και προσωποποιημένου marketing.
4. Η εφαρμογή πολυκριτηριακών συστημάτων συστάσεων σε πραγματικά δεδομένα προτιμήσεων χρηστών του διαδικτύου. Τα συστήματα συστάσεων έχουν αναπτυχθεί σποραδικά με ένα ad hoc τρόπο με χρήση κυρίως στατικών δεδομένων και υπάρχει ανάγκη για έρευνα στο συγκεκριμένο ερευνητικό πεδίο με χρήση πραγματικών δεδομένων σε πραγματικό χρόνο λειτουργίας των δικτύων (Tsafarakis et al., 2010).
5. Η απλούστευση της διαδικασίας επιλογής αγοράς για τους πελάτες, με παροχή όσο το δυνατόν λιγότερων πληροφοριών.
6. Η μείωση του χρόνου που καταναλώνουν οι πελάτες για έρευνα αγοράς στο διαδίκτυο.

7. Η μείωση του χρόνου που καταναλώνουν οι πάροχοι για να διερευνήσουν τις πτυχές βελτίωσης των υπηρεσιών/προϊόντων, με βάση τις ανάγκες και τις προτιμήσεις των πελατών.

#### Ομάδες Στόχου

1. Οι πελάτες χρήστες διαδικτύου. Μέσα από των εντοπισμό και την περιγραφή των προτύπων συμπεριφοράς, την τμηματοποίηση και κατηγοριοποίηση των χρηστών και των κοινοτήτων τους και από τις ατομικές τους αναδράσεις (Lakiotaki et al., 2009) θα παρέχεται το πλαίσιο περιγραφής των πραγματικών προτιμήσεων των ομάδων των χρηστών. Το αποτελεσματικό profiling των χρηστών θα συντελέσει στην εξατομικευμένη διάχυση πληροφορίας, με ένα προσαρμοσμένο και στοχοθετημένο στο χρήστη τρόπο. Τελικός στόχος είναι, μέσα από τη έρευνα στο συγκεκριμένο πεδίο, να παρέχονται στο χρήστη οι κατάλληλες πληροφορίες και δεδομένα που θα συντελούν στην λήψη των καλύτερων δυνατών αποφάσεων.
2. Οι πάροχοι υπηρεσιών/προϊόντων. Οι επιχειρήσεις βρίσκονται σε μια αέναη αναζήτηση για καλύτερη και πιο προσαρμοσμένη στους καταναλωτές προσφορά προϊόντων και υπηρεσιών. Η τεράστια εξάπλωση των κοινωνικών δικτύων και η εκθετική καθημερινά αύξηση των συνδεδεμένων σε αυτά χρηστών (Benkler, 2006) αποτελούν τεράστιο κίνητρο για τα τμήματα Marketing των εταιρειών να αξιοποιήσουν αυτό το φτηνό και μαζικής χρήσης τρόπο επικοινωνίας με τους δυνητικούς πελάτες τους. Το στοχευμένο marketing απαιτεί σε βάθος γνώση των προτιμήσεων του πελάτη κάτι το οποίο μπορεί να επιτευχθεί μέσα από έναν δομημένο και δυναμικό μεθοδολογικό πλαίσιο.

### **1.3 Πρωτοτυπία και χρησιμότητα της έρευνας**

Προκειμένου να περιγραφεί η διαφορετικότητα και πρωτοτυπία της παρούσας ερευνητικής προσπάθειας, παρακάτω αρχικά προβαίνουμε σε καταγραφή

αντίστοιχων μεθοδολογιών και συστημάτων συστάσεων που αναπτύχθηκαν και περιγράφηκαν ερευνητικά τα τελευταία χρόνια (Πίνακας 1-1). Εν συνεχεία περιγράφουμε συνοπτικά τα στοιχεία της μεθοδολογίας που αναπτύξαμε καθώς και τα πρωτότυπα σημεία τα οποία της προσδίδουν προστιθέμενη αξία και καινοτομία.

**Πίνακας 1-1 Καταγραφή των σημαντικότερων μεθοδολογιών και συστημάτων Συστάσεων**

	Citation	Review element	User profile	Recommending method	Tested products	Evaluation metric
1	(Leal et al., 2017)	Ratings	Ratings	crowd-sourced multi-criteria rating profiling - collaborative filtering	Expediaand TripAdvisor	MAE, RMSE Precision, Recall
2	(Musto et al., 2017)	Featured based Opinions	Opinion Analysis sentiment analysis	collaborative filtering (CF)	Restaurants	NA
3	(Colomo-Palacios et al., 2017)	Ratings, Featured based Opinion	Ratings, aspect opinions	Context based recommendation	Travel	Kendall's coefficient, Precision, Recall
4	(Hinduja et al., 2017)	Demographic preferences modeling	Preference	Multi Criteria Recommender System Fuzzy Logic Utility theory	Life insurance	Customer survey
5	(Ranjbar Kermany and Alizadeh, 2017)	Ratings	Ratings	item-based fuzzy multi-criteria collaborative filtering	Movies	NA
6	(Nilashi et al., Oct2016)	Ratings	Classification and Regression Tree	Collaborative Filtering	Yahoo! Movies and TripAdvisor datasets	precision
7	(Traub et al., 2015)	User Interactions( Views, Likes,Booking s)	No Personalization of User, popularity based.	Hybrid Recommendation	ExpediaAffiliate Network(EAN), Hotels	NA
8	(Chen and Chen, 2014)	Aspect-level opinions under contexts	Context-independent context-dependent weight preference for aspects	Preference-based ranking	TripAdvisor and Yelp, Restaurants	Hit Ratio and MRR
9	(Zhang et al., 2013)	Sentiment words and emoticons	Inferred ratings by aggregating words' and emoticons' sentiments	User-based and item-based Collaborative Filtering	Youku, Videos Amazon, Books	Precision
10	(Musat et al., n.d.)	Frequent topics based on opinion count	Topic profile	Weighted average of product's ratings for ranking	TripAdvisor, Hotels	MAE, Kendall's tau rank coeffecient
11	(Pero and Horváth, 2013)	Ratings, Sentiment words	Real ratings and inferred opinion ratings	Biased MF	Amazon, Movies,Music, Books	RMSE
12	(Ganu et al., 2013)	Aspect-level opinions	Ratings and aspect opinions	Soft clustering and user-based Collaborative Filtering	New York Citysearch, Restaurants	RMSE
13	(Chen and Wang, 2013)	Feature opinions	Weight preference for features	Latent class regression model (LCRM) based clustering and user-based CF	Amazon, Cameras,Laptops	Hit Ratio and MRR
14	(Wang et al., 2013)	Feature opinions	Value preference for attributes	Preference-based ranking	Amazon, Cameras	Hit Ratio
15	(Santosa et al., 2012)	User Data Ratings	Preference	Multi-Criteria Decision Making Weighted Sum model (WSM)	Tourism	NA

16	(Dong et al., 2013)	Feature opinions and popularity	Product case with feature sentiment and popularity	Product ranking based on sentiment improvement and Popularity similarity	Amazon,GPS,devices, laptops, tablets	Average Better score, ratings benefit
17	(Levi et al., 2012)	Features - frequent nouns and noun phrases-	Weight preference for features in contexts (nationality and trip intent)	Preference-based ranking	Venere and TripAdvisor,Hotels	User satisfaction
18	(Raghavan et al., 2012)	Review helpfulness (votes by readers)	Rating weighted by quality score	Quality-aware probabilistic MF	Amazon-Books,audio CDs	RMSE
19	(Moshfeghi et al., 2011)	Review emotions via emotion classifier	Ratings	Extended LDA and gradient boosted tree	MovieLens ,Movies)	MSE
20	(Esparza et al., 2011)	Frequent words weighted by TF-IDF	Term-based user index	Content-based method	Blippr , Movies, Applications, books, games Flixter, Movies	Novelty, diversity, Coverage ,precision, recall, F1
21	(Poirier et al., 2010)	Frequent words	Inferred ratings via opinion classification	Item-based Collaborative Filtering	Flixster , Movies	RMSE
22	(Chapphannarungsri and Maneeroj, 2009a)	Frequency of preference	User preference	Multicriteria recommender	Movies	NA
23	(Jakob et al., 2009)	Aspect-level opinions	Ratings, aspect opinions, and number of opinions	Multi-relational MF	IMDb ,Movies	RMSE
24	(Lakiotaki et al., 2008)	Ratings	Ratings	Multicriteria recommender UTARec	Movies	Kendall's tau and ROC curve analysis
25	(Adomavicius and Kwon, 2007)	Ratings	similarity-based approach, aggregation-function-based approach	recommender systems, collaborative filtering, multi-criteria ratings,	Yahoo! Movies website (movies.yahoo.com)	Cosine-based similarities, Chebyshev multidimensional distance metric
26	(Leung, n.d.)	Feature opinions	Inferred ratings by aggregating words' opinion strengths	User-based CF Normal	MovieLens, Movies	NA
27	(Plantié et al., 2005a)	opinions or critics	decision tree classification method	Decision Making Support System	movies	NA
28	(Karacapilidis and Hatzieleftheriou, 2003a)	Ratings, critique based preference modeling	Clustering	Multicriteria knowledge-based and collaborative filtering techniques	Tourism, CityGuide	Robustness

Οι Leal et al., (2017) πρότειναν κατασκευή του προφίλ πελατών με χρήση πολυκριτήριων βαθμολογιών ξενοδοχείων και για την πρόβλεψη των βαθμολογιών χρησιμοποίησαν αλγόριθμους ελαχιστοποίησης τετραγώνων. Τα δεδομένα που χρησιμοποίησαν προέρχονταν από την Expedia και το Tripadvisor. Οι βασικές μετρικές που χρησιμοποίησαν για να αξιολογήσουν το σύστημα συστάσεων που δημιούργησαν είναι το MAE, RMSE, Precision, Recall.

Οι Musto et al., (2017) παρουσίασαν ένα πολυκριτήριο σύστημα συστάσεων που χρησιμοποιεί σχόλια χρηστών και υπολογίζει αυτόματα σκορ συναισθήματος και ενδιαφέροντα των χρηστών με χρήση collaborative filtering στο τομέα των

εστιατορίων. Οι Colomo-Palacios et al., (2017) παρουσίασαν ένα σύστημα που υποστηρίζει τον κύκλο ζωής της τουριστικής αφοσίωσης (loyalty) μετά την πρώτη επίσκεψη, χρησιμοποιώντας τόσο βαθμολογίες όσο και απόψεις χρηστών. Το σύστημα σύστασης αναπτύχθηκε με βάση συστήματα τεχνητής νοημοσύνης και ανάλυσης απόψεων. Οι Hinduja et al., (2017) πρότειναν ένα σύστημα συστάσεων για ασφάλειες ζωής. Χρησιμοποίησαν θεωρία χρησιμότητας (utility theory) για να προτείνουν κατάλληλες πολιτικές επιλογής ασφάλειας στην Ινδία, με δεδομένο τον τεράστιο ανταγωνισμό. Το σύστημα είναι πολυκριτήριο και περιελάμβανε και πτυχές από τη θεωρία ασαφούς λογικής. Η αξιολόγηση του στηρίχθηκε σε έρευνα αγοράς.

Οι Ranjbar Kermany and Alizadeh, (2017) παρουσίασαν μια υβριδική πολυκριτήρια μέθοδο συστάσεων που ενσωματώνει collaborative filtering, ασαφή λογική, δημογραφική πληροφορία και βασιζόμενη στα προϊόντα μέθοδο φιλτραρίσματος για συστάσεις ταινιών. Στα πλαίσια της συγκεκριμένης προσπάθειας χρησιμοποιείται προσαρμοστικό νευρο-ασαφές σύστημα για να ανακαλυφθεί η σχέση μεταξύ κάθε κριτηρίου και της ολικής βαθμολογίας.

Οι Nilashi et al., (2016), παρουσίασαν μια αρκετά ενδιαφέρουσα μέθοδο συστάσεων όπου χρησιμοποίησαν δέντρο ταξινόμησης και παλινδρόμησης (Classification and Regression Tree) και μεγιστοποίηση προσδοκίας (Expectation Maximization) για βελτίωση της ακρίβειας των πολυκριτηρίων συστημάτων συστάσεων. Τα σετ δεδομένων προέρχονται από Yahoo!Movies και το Tripadvisor και με μετρική το precision, έδειξαν ότι βελτίωσαν την ακρίβεια της σύστασης.

Οι Traub et al., (2015) παρουσίασαν υβριδικό σύστημα συστάσεων ξενοδοχείων με βάση δεδομένα για τις αλληλεπιδράσεις και τις κρατήσεις ξενοδοχείων από δεδομένα που άντλησαν από το Expedia Affiliate Network(EAN). Δεν προχώρησαν σε καμία εξατομίκευση των προσωπικών χαρακτηριστικών των χρηστών ενώ οι συστάσεις στηρίζονται στην δημοφιλία των προϊόντων. Στόχος της μελλοντικής τους έρευνας ήταν να δώσουν τη δυνατότητα στο χρήστη να προσδίδει αυτός παραμέτρους και βάρη για τις προτιμήσεις του σε επίπεδο ξενοδοχείου.

Οι Chen and Chen, (2014) πρότειναν έναν αλγόριθμο συστάσεων που βασίζεται σε σχόλια χρηστών, με βάση τον οποίο οι εξαρτημένες προτιμήσεις των χρηστών προκύπτουν από διαφορετικές στρατηγικές εφαρμογής βαρυτήτων. Οι εξαρτημένες προτιμήσεις συνδυάζονται με ανεξάρτητες προτιμήσεις των χρηστών προκειμένου να γίνει η σύσταση. Τα σετ δεδομένων που χρησιμοποίησαν ήταν για εστιατόρια από το TripAdvisor και το Yelp.

Οι Zhang et al., (2013) κατάφεραν να εξαγάγουν «εικονικές» βαθμολογίες μέσα από την επεξεργασία απόψεων και συναισθημάτων για videos από το Youku και βιβλία από Amazon. Οι «εικονικές» βαθμολογίες κατά τους Leung, n.d., Poirier et al., (2010), Zhang et al., (2013) είναι πολύ πιθανό να είναι συγκρίσιμες με τις πραγματικές βαθμολογίες, για χρήση τους σε collaborative filtering τεχνικές.

Οι Musat et al., n.d. παρουσίασαν μια μη εξατομικευμένη μέθοδο κατάταξης που ονομάστηκε κατάταξη βασισμένη στη δημοφιλία. Σε αυτή τη μέθοδο δεν απαιτείται καμιά πληροφορία προτίμησης. Τα σκορ για τα υποψήφια προϊόντα προκύπτουν βρίσκοντας το μέσο όρο των βαθμολογιών όλων των χρηστών. Οι βασικές μετρικές που χρησιμοποιήθηκαν ήταν οι MAE, Kendall's tau rank coefficient.

Οι Pero and Horváth, (2013) προσπάθησαν να ενσωματώσουν τόσο τις πραγματικές βαθμολογίες όσο και τις «εικονικές» βαθμολογίες από τις απόψεις, σε σύστημα συστάσεων, όπου χρησιμοποίησαν τις πραγματικές και τις «εικονικές» βαθμολογίες, τόσο ξεχωριστά σαν εισόδους στο σύστημα και συνδύασαν τα αποτελέσματα, όσο και συνδυαστικά στην κατασκευή του μοντέλου. Τα δεδομένα αφορούσαν ταινίες, μουσική και βιβλία από Amazon και βασική μετρική που χρησιμοποίησαν ήταν το RMSE.

Οι Ganu et al., (2013) ανέπτυξαν ταξινομητή κειμένου για να ταξινομούν προτάσεις σε διαφορετικές κατηγορίες συναισθήματος και αντικειμένου. Οι προτάσεις ταξινομούνται ανά αντικείμενο και συναίσθημα και στη συνέχεια χρησιμοποιούνται για να υπολογιστεί η βαθμολογία της άποψης για το συγκεκριμένο αντικείμενο. Εν συνεχεία, οι βαθμολογίες χρησιμοποιούνται για να παραχθούν συστάσεις μέσα από αλγόριθμους παλινδρόμησης ή ομαδοποίησης. Οι Chen and Wang, (2013) μοντελοποίησαν τις απόψεις με χρήση

παλινδρόμησης, ώστε να εντοπίσουν τις προτιμήσεις των χρηστών σε επίπεδο ομάδας, στα χαρακτηριστικά. Εν συνεχεία, χρησιμοποίησαν τις προτιμήσεις για υπολογίσουν την ομοιότητα μεταξύ χρηστών κατά τη διαδικασία σύστασης.

Σε αρκετές περιπτώσεις χρησιμοποιείται η συνολική άποψη ενός κειμένου για την ανάλυση συναισθήματος και την εξαγωγή της τελικής βαθμολογίας. Για να προκύψει αυτό συνδυάζονται τα συναισθήματα για όλες τις λέξεις άποψης που συμπεριλαμβάνονται σε μια άποψη αξιολόγησης (review). (Leung, n.d., Zhang et al., 2013). Εναλλακτικά χρησιμοποιείται machine learning αλγόριθμος για να «μάθει» την άποψη και να την ταξινομήσει στην κατάλληλη κατηγορία συναισθήματος. (Poirier et al., 2010). Εν συνεχεία οι συνολικές απόψεις μετατρέπονται σε «εικονικές» βαθμολογίες που είτε χρησιμοποιούνται αντί των πραγματικών βαθμολογιών ή βελτιώνουν τις πραγματικές βαθμολογίες (Pero and Horváth, 2013).

Σε μια διαφορετική προσέγγιση οι Zhang et al., (2013) χρησιμοποίησαν εικονίδια (emojicons), σαν συμβολικές αναπαραστάσεις συναισθημάτων, προκειμένου να υπονοήσουν βαθμολογίες που αντιστοιχούν στο συναίσθημα των πελατών.

Οι Dong et al., (2013) με βάση χαρακτηριστικά προϊόντων και αντίστοιχες απόψεις, έχτισαν προφίλ προϊόντων. Τα προφίλ χρησιμοποιούνται για να προτεραιοποιήσουν προϊόντα που είναι όμοια με τις επιλογές του χρήστη.

Οι Levi et al., (2012) πρότειναν ότι οι απόψεις που εκφράζονται από τους χρήστες ξενοδοχείων, μπορούν να συσχετιστούν με άλλα στοιχεία όπως ο σκοπός του ταξιδιού και η εθνικότητα, για να προκύψουν σχέσεις μεταξύ των πτυχής και γενικότερου πλαισίου.

Ο συνδυασμός των αξιολογήσεων-απόψεων με στατικά χαρακτηριστικά μπορεί να αποτελέσει πολύ χρήσιμο εργαλείο για να κατανοήσουμε τις προτιμήσεις του χρήστη σε συγκεκριμένα χαρακτηριστικά σύμφωνα με τους Wang et al., (2013).

Οι Raghavan et al., (2012) χρησιμοποίησαν σαν παράμετρο τις ψήφους στις κριτικές χρηστών που δείχνουν πόσο βοηθητικές ήταν για τους άλλους χρήστες, προκειμένου να βελτιώσουν την ποιότητα των προβλέψεων τους.



Οι Santosa et al., (2012) παρουσίασαν πολυκριτήριο σύστημα συστάσεων για τουριστικό προορισμό, και αναφέρθηκαν στη δυσκολία διαχείρισης της διαθέσιμης πληροφορίας λόγω πολυπλοκότητας.

Οι Moshfeghi et al., (2011) κατέληξαν, μετά από πειράματα στα σετ δεδομένων από το Movielens, ότι ενσωματώνοντας συναισθηματική και σημασιολογική πληροφορία, βελτιώνεται η ακρίβεια της πρόβλεψης για τα collaborative filtering μοντέλα. Η βασική μετρική που χρησιμοποίησαν για τα αποτελέσματα τους ήταν η MSE.

Οι Esparza et al., (2011) επεξεργάστηκαν απόψεις πραγματικού χρόνου που εκφράστηκαν στην υπηρεσία Blippr, αντίστοιχη του Twitter, και ανέπτυξαν τον τρόπο που οι χρήστες και τα προϊόντα μπορούν να αναπαρασταθούν από όρους που χρησιμοποιούνται στις απόψεις τους. Χρησιμοποίησαν μετρικές όπως το precision, recall, F1.

Οι Jakob et al., (2009) χρησιμοποίησαν Multi-Relational Matrix Factorization για να μοντελοποιήσουν αλληλεπιδράσεις μεταξύ χρηστών, δεδομένων και απόψεων χρηστών για ταινίες.

Οι Adomavicius and Kwon, (2007) ανέπτυξαν πολυκριτηριακό σύστημα που χρησιμοποιεί απόψεις χρήστη για πολλαπλές πτυχές ενός αντικειμένου προκειμένου να βελτιώσει τη σύσταση. Τα δεδομένα αφορούσαν ταινίες από το Yahoo!Movies και χρησιμοποίησαν συν-ημιτονοειδή ομοιότητα (cosine based similarities) και τη μετρική απόστασης Chebyshev προκειμένου να μετρήσουν την ποιότητα των συστάσεων.

Οι Chapphannarungsri and Maneeroj, (2009b) παρουσίασαν ένα εξελεγμένο Σύστημα Συστάσεων για Ταινίες που παρέχει υψηλής ποιότητας συστάσεις συνδυάζοντας την Πολυκριτηριακή Αξιολόγηση και Πολυδιάστατες μεθόδους. Για την προσέγγιση με πολλαπλά κριτήρια πρότειναν μια μέθοδο που αλλάζει το τρόπο που υπολογίζονται τα βάρη των κριτηρίων και δίνουν περισσότερη βαρύτητα στη συχνότητα με την οποία επιλέγονται χαρακτηριστικά των ταινιών. Σύμφωνα με την πειραματική αξιολόγηση ο συνδυασμός των πολυκριτηρίων αξιολογήσεων με τις Πολυδιάστατες προσεγγίσεις παρέχει πιο ακριβή

αποτελέσματα συστάσεων σε σχέση με τα σύγχρονα υβριδικά συστήματα Συστάσεων.

Οι Lakiotaki et al., (2008) παρουσίασαν το UTA-Rec, σύστημα συστάσεων που ενσωματώνει μεθοδολογίες πολυκριτήριας ανάλυσης και το εφάρμοσαν για τη μελέτη περίπτωσης σύστασης ταινιών με μετρικές ποιότητας συστάσεων Kendall's tau και ROC curve analysis.

Οι Plantié et al., (2005b) επίσης ανέπτυξαν μια εφαρμογή για σύσταση ταινιών. Πιο συγκεκριμένα ανέπτυξαν ένα πληροφοριακό σύστημα που συλλέγει, επεξεργάζεται και διαχειρίζεται όσο πιο αυτοματοποιημένα γίνεται τις απόψεις και κριτικές για ταινίες που είναι διαθέσιμες στο διαδίκτυο, σε σελίδες σχετικές, για να υποστηρίξει τη πολυκριτήρια αξιολόγηση για σύσταση. Το Σύστημα Συστάσεων αρχικά εφαρμόζει τεχνικές εξαγωγής πληροφορίας προκειμένου να κατηγοριοποιηθούν οι κριτικές των χρηστών που χρησιμοποιήθηκαν για την αξιολόγηση και να τους αποδοθούν αυτόματα αριθμητικές τιμές. Στη συνέχεια εφαρμόζονται πολυκριτήριες τεχνικές για να υπολογιστούν αριθμητικά, να συγκριθούν και να καταταχθούν οι ταινίες στις οποίες αναφέρονται οι κριτικές. Στο τέλος το Σύστημα Συστάσεων παρουσιάζει με διαδραστικό τρόπο τα αποτελέσματα των κατατάξεων.

Οι Karacapilidis and Hatzieleftheriou, (2003b) παρουσίασαν μια εφαρμογή που προσφέρει συστάσεις για επισκέψεις διαφορετικών πόλεων σε μια χώρα. Οι συγκεκριμένοι χρησιμοποίησαν μια πολυκριτήρια προσέγγιση προκειμένου να παρέχουν στους χρήστες μια κατάταξη των εναλλακτικών, ενώ ταυτόχρονα μέσω τους συστήματος επιτρέπουν στο χρήστη να υποβάλλει αξιολογήσεις σχετικά με υπάρχοντα αντικείμενα. Μεγάλη προσοχή έχει δοθεί στο βαθμό στον οποίο η αξιολόγηση του χρήστη επηρεάζει τις τιμές των αποθηκευμένων αντικειμένων. Στη συγκεκριμένη ανάπτυξη έχει γίνει χρήση των ασαφών μέτρων ομοιότητας.

Όταν αναφερόμαστε σε πολυδιάστατα συστήματα συστάσεων θέλουμε να προβλέψουμε βαθμολογίες για περισσότερα του ενός κριτήρια. Στην περίπτωση των ταινιών είναι συνήθως αρκετό να χρησιμοποιηθεί η συνολική βαθμολογία προκειμένου να προκύψει μια καλή πρόβλεψη. Στην περίπτωση όμως των

ξενοδοχείων, μια συνολική βαθμολογία δεν είναι αρκετή, καθώς υπάρχουν άλλες παράμετροι που διαφορετικοί χρήστες θα θεωρήσουν πιο σημαντικές όπως το service, η εγγύτητα στην παραλία, καθαριότητα κτλ. Σε εφαρμογές σαν την περίπτωση των ξενοδοχείων, πολυδιάστατα συστήματα συστάσεων έχουν καλύτερη αποδοτικότητα (Pessemier et al., 2017).

Δραστηριότητες όπως τα ταξίδια, τα γεύματα και η βραδινή ζωή είναι συνήθως κοινωνικές δραστηριότητες όπου συμμετέχουν αρκετοί άνθρωποι. Είναι λοιπόν απαραίτητο να λαμβάνονται υπόψη οι προτιμήσεις για όλους τους αποφασίζοντας όταν παρέχουμε συστάσεις (Moreno et al., n.d.).

Στην παρούσα έρευνα, συνδυάζοντας τα αποτελέσματα της ανάλυσης ικανοποίησης πελατών, της ανάλυσης συναισθήματος, του φιλτραρίσματος επιλογών με κατ' αποκοπή κριτήρια σε μεγάλου όγκου δεδομένα, που έχουμε αντλήσει από το διαδίκτυο (Hotels.com), καθώς και της εφαρμογής του μοντέλου WAP (Spyridakos and Tsotsolas, 2015, Tsotsolas et al., 2016a) για δημιουργία προτιμησιακού προφίλ χρήστη, καταλήγουμε σε πολυκριτήρια σύσταση προϊόντων σε χρήστη.

Το νέο σύστημα συστάσεων χρησιμοποιεί τόσο πραγματικές απόψεις (review-opinions) όσο και βαθμολογίες (ratings) τουριστών για καταλύματα, καθώς και στατικά χαρακτηριστικά καταλυμάτων, τα οποία έχουν αντληθεί στα πλαίσια της διατριβής με χρήση μεθόδων εξόρυξης άποψης από διαδικτυακή πλατφόρμα κρατήσεων. Τις εισόδους του συστήματος αποτελούν επιλογές του χρήστη σε κατ' αποκοπή κριτήρια καθώς και κατάταξη συγκεκριμένων κριτηρίων προκειμένου να δημιουργηθεί το προτιμησιακό του μοντέλο.

Σε επίπεδο μεθοδολογίας χαρακτηρίζεται σαν ένα πολυκριτήριο υβριδικό σύστημα συστάσεων που επεξεργάζεται τόσο πραγματικές βαθμολογίες όσο και σχόλια χρηστών, προκειμένου να καταλήξει στη δημιουργία προτιμησιακών μοντέλων για τα προϊόντα και να τα «ταιριάζει» με προτιμησιακά μοντέλα χρηστών που δημιουργούνται από τις επιλογές τους στο σύστημα, ώστε να καταλήξει στη σύσταση. Οι κύριες μετρικές που χρησιμοποιούνται για την μέτρηση της ποιότητας των συστάσεων που παράγονται είναι οι: MSE, RMSE, MAE, MAPE.

Με βάση την παραπάνω βιβλιογραφική καταγραφή των ερευνητικών προσπαθειών στο αντίστοιχο πεδίο η πρωτοτυπία της παρούσας έρευνας έγκειται στα παρακάτω σημεία:

1. Χρήση μεγάλου όγκου πραγματικών δεδομένων (σαν είσοδος στο σύστημα) που αντλήσαμε με μεθόδους εξόρυξης δεδομένων από το διαδίκτυο σε αντίθεση με τις περισσότερες μελέτες που έχουν χρησιμοποιήσει έτοιμα τεστ σετ δεδομένων που υπήρχαν διαθέσιμα στο διαδίκτυο.
2. Χρησιμοποίηση όλης της διαθέσιμης αντλημένης πληροφορίας προκειμένου να καταλήξουμε σε σύσταση. Πιο συγκεκριμένα, οι περισσότερες από τις μελέτες, χρησιμοποιούν είτε πραγματικές βαθμολογίες, είτε βαθμολογίες που προκύπτουν από ανάλυση συναισθήματος ή κειμένου προκειμένου να καταλήξουν σε τελική σύσταση. Στην παρούσα μεθοδολογία χρησιμοποιούνται δεδομένα βαθμολογιών προϊόντων για την ανάλυση ικανοποίησης πελατών, δεδομένα απόψεων για τη συναισθηματική ανάλυση, στατικά δεδομένα των προϊόντων σαν κατ' αποκοπή κριτήρια, με ουσιαστική εκμετάλλευση όλης της διαθέσιμης πληροφορίας που προκύπτει από την άντληση.
3. Χρήση των αποτελεσμάτων της ανάλυσης συναισθήματος ως κατώφλια ποιότητας σύστασης, σε αντίθεση με τις υπάρχουσες μεθοδολογίες που χρησιμοποιούν τα αποτελέσματα της ανάλυσης, «εικονικές βαθμολογίες», ως είσοδο στο σύστημα σύστασης. Ο ρόλος της συγκεκριμένης ανάλυσης είναι άκρως σημαντικός στη μεθοδολογία καθώς λειτουργεί σαν εγγύηση ποιότητας τελικής σύστασης με βάση έξτρα πληροφορία. Η συγκεκριμένη χρήση της ανάλυσης συναισθήματος δεν παρατηρήθηκε σε καμία από τις μεθοδολογίες.
4. Συνδυασμός τριών διαφορετικών μεθόδων ανάλυσης (MUSA, Sentiment Analysis, Filtering) σε δεδομένα αξιολόγησης/κριτικής από πολλαπλές πηγές πληροφορίας, με εξαιρετική ακρίβεια στην δημιουργία προφίλ πελατών για τη χρήση του στην διαδικασία σύστασης.

5. Επιτυχής εφαρμογή παραλλαγής της πρόσφατα παρουσιασθείσας μεθόδου WAP για τη δημιουργία προτιμησιακού προφίλ χρήστη και συνδυασμού της με τα αποτελέσματα των προαναφερθεισών μεθόδων για τη δημιουργία προσωποποιημένης σύστασης σε χρήστη. Με βάση την συγκεκριμένη παραλλαγή της WAP, ο τελικός χρήστης του συστήματος σύστασης χρειάζεται απλά να επιλέξει μια ποιοτική μεταβλητή, προκειμένου να προσδώσει τη διαφορά σημαντικότητας μεταξύ των καταταγμένων κριτηρίων.
6. Δυνατότητα δημιουργίας προτιμησιακού προφίλ του χρήστη, με ελάχιστα δεδομένα, τα οποία παρέχονται με τη θέληση του και είναι ποιοτικά και συνεπώς ευκολότερα να τα κατανοήσει, σε αντίθεση με τις περισσότερες προσεγγίσεις που κάνουν χρήση προσωπικών δεδομένων και ιστορικού περιήγησης των χρηστών προκειμένου να δημιουργήσουν το προφίλ του.
7. Το σύστημα ξεπερνά ένα συνήθη περιορισμό που τίθεται στα συστήματα συστάσεων που είναι η κρύα εκκίνηση (cold start). Λόγω του τρόπου με τον οποίο καταλήγει σε σύσταση και του αλγοριθμικού πλαισίου, δεν απαιτεί προηγούμενα δεδομένα από το προφίλ του χρήστη ή τις επιλογές του.
8. Ελάχιστος χρόνος από την είσοδο του χρήστη στο σύστημα, μέχρι την παραγωγή της σύστασης. Αυτό οφείλεται στη δομή της μεθοδολογίας, καθώς πριν την είσοδο του χρήστη, έχει προηγηθεί ανάλυση ικανοποίησης και συναισθηματική ανάλυση, και κατά τη χρήση του συστήματος, πραγματοποιείται μόνο φιλτράρισμα των επιλογών με κατ' αποκοπή επιλογές του χρήστη καθώς και δημιουργία του προτιμησιακού του προφίλ με χρήση της μεθόδου WAP. Έτσι η απαίτηση της μεθοδολογίας σε επίπεδο πόρων είναι αρκετά μικρή κατά τη περίοδο που ο χρήστης χρησιμοποιεί το σύστημα.
9. Τελείως διαφορετική προσέγγιση σύστασης προϊόντος σε χρήστη σε σχέση με τις υπάρχουσες μεθοδολογίες στο διαδίκτυο. Πιο συγκεκριμένα, οι συστάσεις στις περισσότερες διαδικτυακές πλατφόρμες αγοράς προϊόντων, γίνονται με βάση δεδομένα που αντλούνται από την

πλοήγηση του χρήστη στο διαδίκτυο (Web Analytics, Clickstream, προηγούμενες συναλλαγές, Heatmaps) και πολλές φορές αποτελούν προσωπικά δεδομένα. Στη περίπτωση του παρόντος συστήματος η πληροφορία που χρησιμοποιείται είναι δημόσια, η άντληση της γίνεται με την θέληση του χρήστη και δεν αφορά προσωπικό απόρρητο του. Αποφεύγεται η άντληση περαιτέρω προσωπικών δεδομένων του χρήστη για χρήση τους στη διαδικασία της σύστασης, καθώς δεν είναι και αναγκαία για την ορθή και αποδοτική λειτουργία του συστήματος.

10. Η χρησιμότητα του συστήματος είναι χαρακτηριστική για όλες τις ομάδες στόχου. Καταρχήν δίνεται η δυνατότητα παροχής χρήσιμης πληροφορίας τόσο στον πελάτη, σύσταση για την επιλογή προϊόντος που ταιριάζει στις ανάγκες του, όσο και στον πάροχο υπηρεσίας/προϊόντος, προσδιορίζοντας του τις τάσεις όσον αφορά την ικανοποίηση των πελατών καθώς και σαφέστερη πληροφορία σχετικά με το προτιμησιακό μοντέλο των πελατών του.
11. Η διαφορετικότητα της πληροφορίας που παρέχεται από το σύστημα σε σχέση με τις υπάρχουσες εφαρμογές είναι ότι: α. καταρχήν βασίζεται σε πολύ μεγάλο όγκο δεδομένων και όχι σε μια στοχοθετημένη και περιορισμένη έρευνα αγοράς, β. χρησιμοποιεί εκτός από βαθμολογίες, ανάλυση κειμένου για να προσδιορίζει καλύτερα το συναίσθημα του πελάτη γ. τα δεδομένα εισόδου είναι ανώνυμα αλλά εξακριβωμένα ότι είναι αληθινά, συνεπώς εμπεριέχουν μεγάλο βαθμό αντικειμενικότητας.
12. Ένα άκρως σημαντικό στοιχείο που δίνει προστιθέμενη αξία στη συγκεκριμένη ερευνητική προσπάθεια είναι ότι όλα τα υποκειμενικά δεδομένα εισόδου στο σύστημα προέρχονται από πελάτες και δεν δίνεται η δυνατότητα για αλλοίωση των δεδομένων από τρίτο –μη πελάτη– τόσο στην είσοδο όσο και στην έξοδο του συστήματος. Είναι χαρακτηριστικό ότι η κατάταξη των τελικών συστάσεων προκύπτει με βάση τη μέση συνολική ικανοποίηση των πελατών για κάθε προϊόν.
13. Η εφαρμογή του συστήματος σε πραγματικές συνθήκες σε ταξιδιωτικό γραφείο επί δύο χρόνια, επίσης αποτελεί σημαντική πρωτοτυπία και

δοκιμή του σε πραγματικές συνθήκες πελατών. Το όφελος της χρήσης της συγκεκριμένης διαδικασίας για τεκμηρίωση λειτουργίας του συστήματος ήταν ότι απευθυνθήκαμε σε πραγματικούς χρήστες που ενδιαφέρονταν να διερευνήσουν τις επιλογές τους για κατάλυμα, ήταν δυνατό να γίνει σε μαζικό επίπεδο πελατών, και επίσης ότι ήταν δυνατή η διαδικασία ανατροφοδότησης σχετικά με την τελική επιλογή καταλύματος.

14. Μέσα από τα αποτελέσματα του πειράματος που υλοποιήθηκε, με χρήση του συστήματος από πραγματικούς πελάτες, αποδεικνύεται ότι το μεγαλύτερο ποσοστό των πελατών που χρησιμοποίησαν τις συστάσεις, μείνανε παραπάνω ικανοποιημένοι από τους πελάτες με παρόμοιο προτιμησιακό μοντέλο, για τους οποίους είχαμε αντλημένα δεδομένα.

## 2. Υφιστάμενη Κατάσταση & Θεωρητικό Υπόβαθρο

- 2.1 Απόψεις/κριτικές και Αξιολογήσεις
- 2.2 Συστήματα Συστάσεων
  - 2.2.1 Κατηγορίες Συστημάτων Συστάσεων
    - 2.2.1.1 Βασιζόμενα σε περιεχόμενο συστήματα συστάσεων
    - 2.2.1.2 Συνεργατικά συστήματα συστάσεων ή συνεργατικό φιλτράρισμα
    - 2.2.1.3 Υβριδικά Συστήματα Συστάσεων
  - 2.2.2 Πολυκριτήρια Συστήματα αποφάσεων
    - 2.2.2.1 Βασική Δομή των πολυκριτήριων συστημάτων λήψης απόφασης
    - 2.2.2.2 Μεθοδολογία Ανάπτυξης Πληροφοριακού Συστήματος Υποστήριξης Αποφάσεων
  - 2.2.3 Η Σύσταση ως ένα πολυκριτήριο πρόβλημα λήψης απόφασης
    - 2.2.3.1 Κατηγορίες Πολυκριτήριων Συστημάτων Συστάσεων
    - 2.2.3.2 Αλγόριθμοι για Πολυκριτήρια Συστήματα Συστάσεων
      - 2.2.3.2.1 Τεχνικές Πολυκριτήριας Αξιολόγησης κατά την διαδικασία της Πρόβλεψης
      - 2.2.3.2.2 Τεχνικές Πολυκριτήριας Αξιολόγησης κατά της διαδικασία της Σύστασης
    - 2.2.3.3 Αξιολόγηση της λειτουργίας Συστημάτων Συστάσεων
    - 2.2.3.4.1 Χαρακτηριστικά και μετρικές αξιολόγησης Συστήματος Συστάσεων
- 2.3 Μοντελοποίηση Προτίμησης
  - 2.3.1 Καθορισμός βαρών μέσω ιεράρχησης προτεραιότητας - (WAP)
- 2.4 Ικανοποίηση Πελατών
  - 2.4.1 Πολυκριτήρια Ανάλυση Ικανοποίησης (MUSA)
- 2.5 Ανάλυση περιεχομένου και συναισθήματος
  - 2.5.1 Ταξινόμηση συναισθήματος σε επίπεδο εγγράφου
  - 2.5.2 Ταξινόμηση συναισθήματος με εποπτευόμενη μάθηση
  - 2.5.3 Ταξινόμηση συναισθήματος με χρήση μη-εποπτευόμενης μάθησης
  - 2.5.4 Ταξινόμηση Υποκειμενικότητας

## 2. Υφιστάμενη Κατάσταση & Θεωρητικό Υπόβαθρο

### 2.1 Απόψεις/κριτικές και Αξιολογήσεις

Ο ορισμός της εξόρυξης άποψης είναι αρκετά ευρύς. Η παρούσα έρευνα ασχολείται, με το πεδίο της εξόρυξης άποψης, όπου αναλύεται το κείμενο προκειμένου να αποδειχτεί αν ένα αντικείμενο προκαλεί αρνητικά, θετικά ή ουδέτερα συναισθήματα καθώς και οι βαθμολογίες των κριτικών για προσδιορισμό της ικανοποίησης πελατών. Πιο συγκεκριμένα αναλύονται σχόλια χρηστών για προϊόντα ή υπηρεσίες, που είναι διαθέσιμα σε διαδικτυακές πλατφόρμες, και τα αποτελέσματα της ανάλυσης συναισθήματος και της



ανάλυσης ικανοποίησης διοχετεύονται, με συγκεκριμένο τρόπο, στον μηχανισμό που αναπτύσσεται κατά το παρόν μεθοδολογικό πλαίσιο (Liu, 2012).

Απόψεις (opinions) είναι δυνατόν να εκφράζονται για οτιδήποτε π.χ. ένα προϊόν, μια υπηρεσία, ένα πρόσωπο, έναν οργανισμό ή ένα γεγονός. Είθισται γενικά, να χρησιμοποιείται ο όρος «αντικείμενο» (object) για να υποδηλώνεται η οντότητα που αποτελεί το στόχο σχολιασμού. Ένα αντικείμενο μπορεί να έχει ένα σύνολο στοιχείων-μερών (parts) που το αποτελούν καθώς και ένα σύνολο χαρακτηριστικών (Attributes). Κάθε μέρος είναι δυνατόν να αποτελείται από υπο-μέρη με αντίστοιχα χαρακτηριστικά. Η προαναφερθείσα σχέση θα μπορούσε να περιγραφεί ως εξής:  $o:(T,A)$  όπου  $o$  το αντικείμενο,  $T$  η ιεραρχία των μερών που αποτελούν το αντικείμενο και  $A$  το σύνολο των χαρακτηριστικών του  $o$ .

Ο εκφραστής της άποψης χαρακτηρίζεται ως κάτοχος της άποψης (opinion holder) ή πηγή άποψης (opinion source) και μπορεί να είναι άνθρωπος ή οργανισμός. Στην περίπτωση των σχολίων για προϊόντα, κάτοχοι της άποψης είναι οι συγγραφείς τους. Η άποψη για ένα χαρακτηριστικό ή για ολόκληρο το αντικείμενο είναι δυνατόν να έχει τρεις προσανατολισμούς (opinion orientation/polarity): Θετικό, Αρνητικό ή Ουδέτερο.

Ένα έγγραφο απόψεων (review document)  $d$  περιέχει απόψεις ενός συνόλου κατόχων άποψης  $\{h_1, h_2, \dots, h_p\}$  για ένα σύνολο αντικειμένων  $\{o_1, o_2, \dots, o_q\}$ . Οι απόψεις για κάθε αντικείμενο  $o_j$  εκφράζονται για ένα υποσύνολο  $F_j$  των χαρακτηριστικών του  $o_j$ . Η άποψη είναι δυνατό να έχει μια από τις δύο παρακάτω μορφές:

1. Άμεση άποψη (Direct Opinion) η οποία περιγράφεται από τα εξής πέντε στοιχεία  $(o_j, f_{jk}, oo_{ijkl}, h_i, t_i)$ , όπου  $o_j$  ένα αντικείμενο,  $f_{jk}$  ένα χαρακτηριστικό του αντικειμένου  $o_j$ ,  $oo_{ijkl}$  ο προσανατολισμός της άποψης στο χαρακτηριστικό  $f_{jk}$  του αντικειμένου  $o_j$ ,  $h_i$  ο κάτοχος της άποψης και  $t_i$  η χρονική στιγμή που εκφράζεται η άποψη. Ο προσανατολισμός της άποψης  $oo_{ijkl}$  είναι δυνατό να είναι θετικός, αρνητικός ή ουδέτερος.
2. Συγκριτική άποψη (Comparative Opinion) η οποία εκφράζει μια σχέση ομοιοτήτων ή διαφορών ανάμεσα σε δύο ή περισσότερα αντικείμενα και

προτίμηση αντικειμένου από τον κάτοχο άποψης που στηρίζεται σε συγκεκριμένα χαρακτηριστικά του.

Η ανάλυση των διαδικτυακών απόψεων/κριτικών και βαθμολογιών, που αποτελούν υλικό παραγόμενο από χρήστες διαδικτυακών πλατφορμών κρατήσεων, προκειμένου να αναλυθεί η συμπεριφορά των τουριστών, οι απόψεις τους καθώς και τα συναισθήματά τους, αποτελεί βασικό αντικείμενο έρευνας τα τελευταία χρόνια (Kasper and Vela, 2011).

Οι online κριτικές και βαθμολογίες είναι αποτέλεσμα μιας στάνταρ διαδικασίας, σύμφωνα με την οποία, οι πελάτες ενός ξενοδοχείου ερωτώνται να αξιολογήσουν τον προορισμό, σχολιάζοντας τον σύμφωνα με τις προσωπικές τους εμπειρίες και βαθμολογώντας τον χρησιμοποιώντας μια κλίμακα 1 έως 5 αστεριών, με το ένα αστέρι να υποδεικνύει «φτωχή εμπειρία» και τα 5 αστέρια «τέλεια εμπειρία» (Mudambi and Schuff, 2010).

Ένας από τους βασικούς λόγους της όλο και αυξανόμενης σημαντικότητας της ανάλυσης των σχολίων των ταξιδιωτών που αναρτώνται στο διαδίκτυο, τα τελευταία χρόνια, είναι η ριζική αλλαγή του τρόπου κρατήσεων καταλυμάτων και λοιπών υπηρεσιών διακοπών και τουρισμού. Εν αντιθέσει με το παρελθόν, που τα ταξιδιωτικά πρακτορεία καθόριζαν τους προορισμούς, τις επισκέψεις, τα καταλύματα, και οι ταξιδιώτες απλά επέλεγαν ένα από τα προσφερόμενα ταξιδιωτικά πακέτα, στην παρούσα χρονική περίοδο η επιλογή όλων των λεπτομερειών ενός ταξιδιού έχει μετατραπεί σε προσωπική υπόθεση για κάθε ταξιδιώτη.

Οι διαδικτυακές ιστοσελίδες με απόψεις πελατών, πλέον επιτρέπουν στους ταξιδιώτες να επισημάνουν και να χαρακτηρίσουν κάθε μια από τις διαφορετικές παραμέτρους ενός ταξιδιού συμπεριλαμβανομένων καταλυμάτων, εστιατορίων, προορισμών και άλλων συναφών με τον τουρισμό προϊόντων (Burton and Khammash, 2010). Οι online βαθμολογίες (online ratings) βοηθούν τους πελάτες να ξεπερνούν τους έμφυτους περιορισμούς των διαδικτυακών αγορών, όπως η αδυναμία να αξιολογήσουν προϊόντα ή υπηρεσίες οι ίδιοι ή να επικοινωνήσουν απευθείας με το προσωπικό πωλήσεων (Simonson and Rosen, 2014).

Σύμφωνα με τη Verma et al., (2012), η ξενοδοχειακή βιομηχανία γνωρίζει καλά ότι οι βαθμολογίες των πελατών ασκούν μια ισχυρή επιρροή στην επιθυμία άλλων πελατών να κάνουν κράτηση σε ένα ξενοδοχείο. Όταν ένα ξενοδοχείο έχει αρνητικά σχόλια, οι ερωτηθέντες δίνουν μια πιθανότητα 2 στα 5 ότι θα κάνουν κράτηση. Όταν βλέπουν μια θετική κριτική, υπάρχει μια πιθανότητα 3.5 με 4 στα 5 να κάνουν κράτηση στο συγκεκριμένο ξενοδοχείο.

Μία έρευνα από τους Gretzel and Yoo, (2008) έδειξε ότι περισσότερο από 74 τις εκατό των ταξιδιωτών κάνουν χρήση των σχολίων των πελατών, σαν πηγές πληροφοριών όταν οργανώνουν ταξίδια αναψυχής. Οι Ye et al., (2011), κατέληξαν ότι η σχέση μεταξύ της συμπεριφοράς αξιολόγησης ξενοδοχείων από τους τουρίστες και το ποσοστό κρατήσεων δωματίων είναι πολύ σημαντική, αποδεικνύοντας ότι μια αύξηση κατά 10% στις αξιολογήσεις που προέρχονται από επισκέπτες προκαλεί αύξηση στις κρατήσεις τουλάχιστον κατά 5%.

Μια χαρακτηριστική προσέγγιση που χρησιμοποιείται για την αξιολόγηση ικανοποίησης ενός προορισμού είναι το HOLSAT μοντέλο (Tribe and Snaith, 1998). Το συγκεκριμένο μοντέλο υιοθετεί την φιλοσοφία του μοντέλου SERVQUAL (Parasuraman et al., 1985) και τα αποτελέσματα του στοχεύουν στην διαφορά μεταξύ των βαθμολογιών «προσδοκίας» και «εμπειρίας» για κάθε χαρακτηριστικό, το οποίο δίνει ένα ποσοτικό μέτρο του επιπέδου ικανοποίησης που παρουσιάζεται στους πελάτες (Truong and Foster, 2006). Η μέθοδος πολυκριτήριας ανάλυσης ικανοποίησης πελατών (MUSA) (Grigoroudis and Siskos, 2002, 2010) εφαρμόστηκε επιπρόσθετα, από τους Arabatzis and Grigoroudis, (2010) για να μελετήσουν το επίπεδο ικανοποίησης των επισκεπτών του Εθνικού Πάρκου Δάδια-Λευκίμης Σουφλίων.

Αρκετοί ερευνητές έχουν μελετήσει τα χαρακτηριστικά που συντελούν στην ικανοποίηση πελατών ξενοδοχείων (Chathoth, 2007; Jay Kandampully and Dwi Suhartanto, 2000; Skogland and others, 2004) και κατέληξαν στα ακόλουθα συμπεράσματα: «Η ικανοποίηση για ένα ξενοδοχείο είναι βασικής σημασίας για τη μακροχρόνια επιτυχία», «η ικανοποίηση είναι απόλυτα συνδεδεμένη με τις υπηρεσίες, την καθαριότητα και την τιμή» και «οι υπηρεσίες αποτελούν τις πιο σημαίνουσες παραμέτρους».

Εστιάζοντας στην σημαντικότητα των παραμέτρων, ο Atkinson βρήκε ότι η καθαριότητα, η ασφάλεια, η αντιστοίχιση αξίας προς τιμή και η ευγένεια του προσωπικού καθορίζουν την ικανοποίηση ενός πελάτη (Atkinson, 2010). Οι Choi and Chu, (2001) επίσης κατέληξαν ότι η ποιότητα του προσωπικού, οι ανέσεις του δωματίου και η αξία είναι οι τρεις παράμετροι μεγαλύτερης επιρροής για την ικανοποίηση ταξιδιωτών σύμφωνα με την ανάλυση σημαντικότητας- απόδοσης. Οι παραδοσιακές μελέτες ικανοποίησης πελατών συνήθως χρησιμοποιούν τις απαντήσεις πελατών σε διαμορφωμένα εκ των προτέρων ερωτηματολόγια. Στην περίπτωση αυτή οι ερωτήσεις είναι πολύ καλά δομημένες, αντιστοιχούν σε συγκεκριμένα κριτήρια ή υποκριτήρια, τα οποία χρησιμοποιούνται για την εφαρμογή αλγοριθμικών αναλύσεων. Το μειονέκτημα σε αυτή την προσέγγιση παραγωγής δεδομένων είναι ότι ο όγκος των δεδομένων που είναι δυνατόν να προκύψουν είναι μικρός και το δείγμα των ερωτηθέντων επικεντρώνεται σε δίκτυα φίλων των διενεργούντων την έρευνα, με αποτέλεσμα μη αντιπροσωπευτικές απαντήσεις.

### Ποιότητα των Κριτικών

Ένα από τα βασικά προαπαιτούμενα για την αποτελεσματική χρήση και επεξεργασία των κριτικών είναι ο καθορισμός της ποιότητας, της εξυπηρετικότητας και της χρηστικότητας μιας κριτικής (Archak et al., 2007), (Ghose and Ipeirotis, 2011). Έχει υψηλή σημασία η συγκεκριμένη διαδικασία, καθώς είναι επιθυμητό από το χρήστη να του παρέχεται κατάταξη των κριτικών βασισμένη στην ποιότητα τους. Η συγκεκριμένη διαδικασία, παροχής αξιολόγησης των κριτικών, παρέχεται από αρκετές ιστοσελίδες που φιλοξενούν ή πωλούν προϊόντα, όπως η Amazon, όπου ο χρήστης αναφέρει αν η κριτική του φάνηκε χρήσιμη. Εν συνεχεία, τα αποτελέσματα που προκύπτουν τα συνοψίζουν, αναφέροντας δίπλα στην κριτική, ότι π.χ. 9 στους 10 πελάτες θεώρησαν την κριτική χρήσιμη.

Η ποιότητα των κριτικών έχει συχνά προσεγγιστεί σαν ένα πρόβλημα παλινδρόμησης (regression). Το μοντέλο μάθησης προσδίδει ένα σκορ

ποιότητας σε κάθε κριτική, το οποίο εν συνεχεία χρησιμοποιείται για την κατάταξη των κριτικών. Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση και το τεστ των μοντέλων, συνήθως αποτελούν οι αξιολογήσεις των κριτικών που παρέχονται από τους πελάτες ή χρήστες στις διαδικτυακές πλατφόρμες. Οι Zhang and Varadarajan, (2006) χρησιμοποίησαν παραμέτρους όπως η αξιολόγηση της κριτικής, το μήκος της, λέξεις που δηλώνουν συναίσθημα, td-idf βαθμούς βαρύτητας, σύγκριση με τις προδιαγραφές του προϊόντος, μέτρηση πλήθους εμφάνισης συγκεκριμένων λέξεων ή φράσεων προκειμένου να εκπαιδεύσουν και τεστάρουν το μοντέλο παλινδρόμησης που ανέπτυξαν. Οι Ghose and Ipeirotis, (2011), χρησιμοποίησαν, πέραν των προαναφερθέντων χαρακτηριστικών και στοιχεία του προφίλ και του ιστορικού των χρηστών που έκαναν τις κριτικές προκειμένου να βελτιστοποιήσουν το αποτέλεσμα του μοντέλου τους.

Οι O'Mahony and Smyth, (2009) πρότειναν επίσης μια μέθοδο ταξινόμησης των κριτικών σε χρήσιμες και μη, για την περίπτωση των ξενοδοχείων. Ανάμεσα στα στοιχεία που χρησιμοποίησαν ήταν τα εξής: χαρακτηριστικά φήμης (Μέσος όρος και τυπική απόκλιση της χρησιμότητας των κριτικών), χαρακτηριστικά περιεχομένου (μέγεθος κριτικής, ποσοστό κεφαλαίων και πεζών χαρακτήρων), κοινωνικά χαρακτηριστικά (αριθμός κριτικών από συγκεκριμένο χρήστη), χαρακτηριστικά συναισθήματος (το σκορ αξιολόγησης της συγκεκριμένης κριτικής).

Οι κριτικές πελατών σε διαδικτυακές πλατφόρμες κρατήσεων, όπως το hotels.com, booking.com, αποτελούν τεράστιες βάσεις δεδομένων και κριτικών που προέρχονται από πελάτες με διαφορετικά υπόβαθρα χαρακτηριστικά καθώς και τοποθεσίες. Η ανάλυση τέτοιου είδους δεδομένων αποδεικνύει ότι έχουμε πιο αντιπροσωπευτικά δείγματα απόψεων και οι κριτές είναι πιο ρεαλιστικοί στις απαντήσεις τους. Ένα μειονέκτημα των διαδικτυακών κριτικών, είναι οι ψεύτικες κριτικές (fake reviews) που έχουν παρουσιαστεί σαν φαινόμενο, και ουσιαστικά αποτελούν κριτικές που έχουν δημοσιευτεί από μη έγκυρους «πελάτες», και οδηγούν στην διαφοροποίηση της πραγματικής εικόνας ενός προϊόντος ή υπηρεσίας. Επίσης οι κριτικές που προέρχονται από διαδικτυακές πλατφόρμες,

λόγω του μεγάλου όγκου τους καθώς και των πολλών διαφορετικών δομών στις οποίες τις συναντάμε, παρουσιάζουν μεγάλη δυσκολία στην ανάλυση και επεξεργασία τους.

## 2.2 Συστήματα Συστάσεων

Οι Malone et al., (1987) παρουσίασαν συστήματα που προσφέρουν έξυπνη πληροφορία, αναφερόμενοι σε μια θεμελιώδη κατηγοριοποίηση των συστημάτων που γενικά υποστηρίζουν την πρόσβαση σε πολύ δυναμικές πηγές πληροφορίας. Πιο συγκεκριμένα ξεχώρισαν δύο κατηγορίες συστημάτων: 1. Γνωστικά συστήματα Φιλτραρίσματος (Cognitive Filtering Systems) τα οποία χαρακτηρίζουν τα περιεχόμενα μιας πληροφοριακής πηγής και τις πληροφοριακές ανάγκες των πιθανών χρηστών των αντικειμένων, και στη συνέχεια χρησιμοποιούν αυτές τις απεικονίσεις προκειμένου ευφυώς να ταιριάζουν χρήστες με αντικείμενα, 2. Κοινωνιολογικά Συστήματα Φιλτραρίσματος (Sociological filtering Systems) τα οποία δουλεύουν με βάση τις προσωπικές και οργανωσιακές αλληλεξαρτήσεις των ατόμων σε μια κοινότητα. Τα πρώτα συστήματα διαμοιρασμού πληροφορίας ανήκαν στην 1<sup>η</sup> κατηγορία και βασίζονταν σε φιλτράρισμα κειμένου. Τα συστήματα συστάσεων παρουσιάστηκαν σαν αντιπροσωπευτικά της 2<sup>ης</sup> κατηγορίας και ορίστηκαν σαν τα συστήματα που «βοηθούν τους ανθρώπους να κάνουν επιλογές βασισμένες σε απόψεις άλλων ανθρώπων» (Goldberg et al., 1992).

Ο όρος Συστήματα Συστάσεων (Recommender Systems) αρχικά περιγράφηκε σαν συστήματα που «οι άνθρωποι παρέχουν πληροφορίες σαν εισόδους, τις οποίες στη συνέχεια το σύστημα συναθροίζει και τις κατευθύνει στους κατάλληλους αποδέκτες» (Resnick and Varian, 1997). Μέσα στο χρόνο εξελίχθηκε ο ορισμός και η πιο πρόσφατη εκδοχή του είναι: «Οποιοδήποτε σύστημα που παράγει εξατομικευμένες συστάσεις σαν έξοδο του, ή έχει τη αποτέλεσμα να οδηγεί το χρήστη με ένα προσωποποιημένο τρόπο σε

ενδιαφέροντα ή χρήσιμα για αυτόν αντικείμενα ανάμεσα σε ένα πλήθος πιθανών επιλογών» (Burke, 2002).

Παρακάτω παρατίθενται οι πιο ενδιαφέροντες ορισμοί για τα συστήματα συστάσεων που προέκυψαν με την πάροδο του χρόνου.

1. “Collaborative filtering simply means that people collaborate to help one another perform filtering by recording their reactions to documents they read.” (Goldberg et al., 1992)
2. “Collaborative filters help people make choices based on the opinions of other people.” (Resnick et al., 1994)
3. “Social information filtering essentially automates the process of ‘word-of-mouth’ recommendations: items are recommended to a user based upon values assigned by other people with similar taste.” (Shardanand and Maes, 1995)
4. “In a typical recommender system people provide recommendations as inputs, which the system then aggregates and directs to appropriate recipients.” (Resnick and Varian, 1997)
5. “The term ‘collaborative filtering’ describes techniques that use the known preferences of a group of users to predict the unknown preferences of a new user; recommendations for the new users are based on these predictions. Other terms that have been proposed are ‘social information filtering’ and ‘recommender system’.” (Pennock et al., 2000)
6. “In collaborative filtering, entities are recommended to a new user based on the stated preferences of other, similar users.” (Cohen and Fan, 2000)
7. “Recommender systems use product knowledge—either hand-coded knowledge provided by experts or ‘mined’ knowledge learned from the behavior of consumers to guide consumers through the often-overwhelming task of locating products they will like”. (Schafer et al., 1999)
8. “Recommender systems use the opinions of members of a community to help individuals in that community identify the information or products



- most likely to be interesting to them or relevant to their needs.” (Konstan, 2004)
9. “Recommender systems use the opinions of a community of users to help individuals in that community more effectively identify content of interest from a potentially overwhelming set of choices.” (Konstan, 2004)
  10. “Recommender systems—a personalized information filtering technology used to either predict whether a particular user will like a particular item (prediction problem) or to identify a set of N items that will be of interest to a certain user (top-N recommendation problem).” (Deshpande and Karypis, 2004)
  11. “Recommender system is a system that helps users to find their wanted items by making recommendations based on either the content of the recommended items (content-based filtering), or ratings of similar users on the recommended items (collaborative filtering).” (Han et al., 2004)
  12. “A personalized recommendation system can provide one-to-one service to customers based on customers’ past behavior and through inference from other users with similar preferences. The aim of personalization is to offer customers what they want without asking explicitly and to capture the social component of interpersonal interaction.” (Hung, 2005)
  13. “Recommender systems suggest items of interest to users based on their explicit and implicit preferences, the preferences of other users, and user and item attributes.” (Schein et al., 2005)
  14. “A recommender system is the information filtering that applies data analysis techniques to the problem of helping customers find the products they would like to purchase by producing a predicted likeness score or a list of recommended products for a given customer”. (Manouselis and Costopoulou, 2007a)



### 2.2.1 Κατηγορίες Συστημάτων Συστάσεων

Τα σύγχρονα συστήματα συστάσεων, χρησιμοποιούν τη βαθμολογία αξιολόγησης σε αντικείμενα προκειμένου να αξιολογήσουν την άποψη προτίμησης του χρήστη. Ένας χρήστης είναι δυνατόν να εκφέρει την άποψη του με βάση κάποια συγκεκριμένα χαρακτηριστικά του αντικειμένου. Για πιο ακριβείς συστάσεις είναι απαραίτητο να ληφθούν υπ' όψιν όλα τα κριτήρια που καθοδηγούν την απόφαση του χρήστη. Με χρήση της MCDA λαμβάνονται υπ' όψιν όλα τα δεδομένα ενδιαφέροντος σχετικά με την απόφαση του χρήστη. (Chapphannarungsri and Maneeroj, 2009)

Τα συστήματα συστάσεων συνήθως κατατάσσονται σε τρεις κατηγορίες με βάση την προσέγγιση για την σύσταση: με βάση το περιεχόμενο, τα συνεργατικά και τα υβριδικά (Balabanović and Shoham, 1997a).

Τα με βάση το περιεχόμενο συστήματα συστάσεων συστήνουν αντικείμενα παρόμοια με εκείνα που έχει προτιμήσει ο χρήστης κατά το παρελθόν.

Τα Συνεργατικά (ή συνεργατικό φιλτράρισμα- Collaborative filtering) συστήματα συστάσεων συστήνουν αντικείμενα που χρήστες με παρόμοιες προτιμήσεις έχουν επιλέξει στο παρελθόν.

Τέλος, οι υβριδικές (hybrid) προσεγγίσεις συνδυάζουν content based και collaborative filtering μεθόδους (Adomavicius and Tuzhilin, 2005a).

Επιπλέον τα συστήματα συστάσεων μπορούν να ταξινομηθούν με βάση τη φύση της αλγοριθμικής τεχνικής τους σε βασιζόμενες στη μνήμη (memory based) και σε βασιζόμενες στο μοντέλο (model based) προσεγγίσεις (Breese et al., 1998a).

Ειδικότερα, οι τεχνικές βασιζόμενες στη μνήμη συχνά υπολογίζουν συστάσεις σε μικρό χρόνο στηριζόμενες στις δραστηριότητες του προηγούμενου χρήστη.

Αντίθετα οι τεχνικές που βασίζονται στα μοντέλα χρησιμοποιούν προηγούμενες δραστηριότητες του χρήστη ώστε να σχηματίσουν και να «μάθουν» ένα μοντέλο πρόβλεψης το οποίο θα χρησιμοποιήσουν στη συνέχεια για να κάνουν τη σύσταση.

Είναι σημαντικό να σημειωθεί ότι η πλειοψηφία των συστημάτων συστάσεων τυπικά χρησιμοποιούν ένα μοναδικό κριτήριο για να απεικονίσουν τη χρησιμότητα ενός αντικειμένου σε ένα χρήστη στον δισδιάστατο χώρο Χρήστες-Αντικείμενα.

Η διαδικασία σύστασης ξεκινάει με την εξειδίκευση των αρχικών αξιολογήσεων που είτε προέρχεται «εξωτερικά» από τους χρήστες είτε «εσωτερικά» από το σύστημα. Μόλις εξειδικευτούν οι αρχικές αξιολογήσεις, το σύστημα συστάσεων προσπαθεί να υπολογίσει την συνάρτηση αξιολόγησης:

$$R, R: \text{Users} \times \text{Items}$$

για τα ζεύγη χρήστη-αντικείμενου που δεν έχουν ακόμα βαθμολογηθεί.

Αφότου υπολογιστεί το  $R$ , ένα σύστημα συστάσεων μπορεί να προτείνει το υψηλότερα αξιολογημένο αντικείμενο σε κάθε χρήστη. Παρόλο που τα single-rating συστήματα συστάσεων παρουσιάζονται αρκετά αποτελεσματικά και επιτυχημένα σε ένα πλήθος εφαρμογών, τα πολυκριτήρια συστήματα αξιολογήσεων διεισδύουν όλο και περισσότερο στην αγορά. Οδηγοί εστιατορίων, όπως ο Zagat's Guide, παρέχουν τρία κριτήρια για αξιολόγηση εστιατορίου. Επίσης γνωστά online shopping malls χρησιμοποιούν multicriteria ratings για τα ηλεκτρονικά είδη που πουλάνε. Επίσης το Yahoo!Movies παρέχει μια υπηρεσία συστάσεων που χρησιμοποιεί τα multicriteria ratings για κάθε ταινία, κάτι που αποδεικνύει ότι τα πολυκριτήρια δεδομένα είναι δυνατόν να είναι χρήσιμα τόσο σε παρόχους διαδικτυακών δεδομένων όσο και σε πελάτες και μπορούν να εξελιχθούν σε απαραίτητα στοιχεία διαφορετικών εφαρμογών σε διαφορετικές εφαρμογές που έχουν στόχο την εξατομίκευση (personalization) (Kwon, 2011).

#### **2.2.1.1 Βασιζόμενα σε περιεχόμενο συστήματα συστάσεων**

Τα βασιζόμενα σε περιεχόμενο συστήματα συστάσεων χρησιμοποιούν πραγματικά χαρακτηριστικά του περιεχομένου προκειμένου να παράξουν συστάσεις. Στα συγκεκριμένα συστήματα, η σύσταση ξεκινάει με την άντληση

του περιεχομένου, όπου χρησιμοποιούνται τεχνικές εξαγωγής χαρακτηριστικών και ταξινόμησης πληροφορίας για εξαγωγή των δεδομένων περιεχομένου.

Η πιο συνήθης διαδικασία σύστασης για τα content based, είναι σύσταση ενός αντικειμένου σε χρήστη, με βάση την περιγραφή του αντικειμένου και το προφίλ του χρήστη που προκύπτει από τις προτιμήσεις και ενδιαφέροντα του. Παρόλο που το προφίλ είναι δυνατόν να εισαχθεί απευθείας από τον χρήστη, συνήθως πραγματοποιείται μια διαδικασία μάθησης-ανατροφοδότησης του συστήματος με επιλογές του χρήστη για άλλα αντικείμενα. Για τη διαδικασία αυτή έχουν αναπτυχθεί ένα πλήθος αλγορίθμων, προκειμένου να παραχθούν ή να γίνουν γνωστά τα χαρακτηριστικά του προφίλ ενός χρήστη. Η επιλογή του κατάλληλου αλγόριθμου γίνεται ανάλογα με το διαθέσιμο περιεχόμενο (Pazzani and Billsus, 2007).

Παρακάτω αναφέρονται τα πιο σημαντικά βήματα για τη διαδικασία της σύστασης με βάση το περιεχόμενο:

1. Απόκτηση δεδομένων (Data acquisition): Απόκτηση των δεδομένων προτίμησης των χρηστών από πληροφορίες που προκύπτουν από ιστορικό περιήγησης, συναλλαγών, ή αρχεία ιστοτόπων.
2. Επεξεργασία Δεδομένων: Τα δεδομένα αποκτούνται με φιλτράρισμα ή διαλογή.
3. Επεξεργασία σύστασης: Μοντέλο σύστασης και παραγωγή κατωφλίου από τη σύγκριση των δεδομένων.
4. Αποτελέσματα σύστασης: Λίστα των αποτελεσμάτων για σύσταση προς τους χρήστες.

Τα κύρια πλεονεκτήματα των συγκεκριμένων συστημάτων σύστασης κυρίως σχετίζονται με το γεγονός ότι προσωποποιούν και εξατομικεύουν τη σύσταση στο χρήστη, ότι δεν απαιτούν πληροφορία για τις επιλογές άλλων χρηστών, μπορούν προτείνουν καινούρια προϊόντα/υπηρεσίες στο χρήστη, και επίσης μπορούν να αιτιολογήσουν την επιλογή σύστασης συγκεκριμένου προϊόντος. Για αυτό το λόγο χρησιμοποιούνται ευρέως για εξατομικευμένες υπηρεσίες

Οι κύριες δυσκολίες που πιθανόν να συναντήσει κάποιος με τη χρήση τέτοιου είδους συστημάτων συστάσεων είναι οι παρακάτω:

1. Δυσκολία ανάλυσης πολυμεσικών δεδομένων (Kim et al., 2005).
2. Απαίτηση ανατροφοδότησης από το χρήστη. Σε πολλές περιπτώσεις ο χρήστης είναι διστακτικός να κάνει ανατροφοδότηση.
3. Η συνωνυμία ή πολυσημία λέξεων ή όρων είναι σύνηθες να καταλήγει σε ανακριβή αποτελέσματα (Balabanović and Shoham, 1997b).
4. Το πρόβλημα του νέου χρήστη: Αν ο χρήστης είναι νέος και δεν προϋπάρχει πληροφορία για αυτόν είναι δύσκολο να προκύψει ακριβής σύσταση με πραγματικού χρόνου ανάλυση δεδομένων. Αρκετοί ερευνητές έχουν προβεί σε προτάσεις για την επίλυση του συγκεκριμένου προβλήματος όπως απάντηση συγκεκριμένων ερωτήσεων από το χρήστη, που θα διευκολύνει την εύρεση προτιμησιακού προφίλ του χρήστη, πριν την αρχική είσοδο στο σύστημα (Burke, 2002).
5. Ο περιορισμός αντικειμένου. Οι χρήστες μπορούν να δεχθούν συστάσεις μόνο για προϊόντα που υπάρχουν από το παρελθόν στη βάση προϊόντων, χωρίς να υπάρχει συσχέτιση του ιστορικού του προτιμησιακού μοντέλου του χρήστη με το νέα προϊόντα που θα μπορούσαν να του συσταθούν.
6. Περιορισμός στο φιλτράρισμα. Δεν υπάρχει δυνατότητα φιλτραρίσματος σε στοιχεία όπως η ποιότητα, το στυλ ή η οπτική γωνία σε ένα περιεχόμενο.

#### **2.2.1.2 Συνεργατικά συστήματα συστάσεων ή συνεργατικό φιλτράρισμα**

Τα συνεργατικά συστήματα συστάσεων βασίζονται στην ιδέα ότι τα συγκεκριμένα συστήματα παρέχουν συστάσεις στους χρήστες χρησιμοποιώντας πληροφορία από μεμονωμένους χρήστες ή γκρουπ χρηστών που έχουν παρόμοια ενδιαφέροντα ή χαρακτηριστικά. Για να υλοποιηθεί το προαναφερθέν, ακολουθούνται τα παρακάτω βασικά βήματα:

1. Παράγεται το προφίλ του χρήστη. Κατασκευή του προφίλ του χρήστη με βάση το ιστορικό συναλλαγών, αλληλεπιδράσεων ή αξιολογήσεων του.

2. Γειτονικά δεδομένα. Το σύστημα μπορεί να ψάξει για γκρουπ χρηστών που μοιάζουν με το χρήστη στόχο χρησιμοποιώντας στατιστικά μοντέλα ή μηχανική μάθηση. Αυτά τα γκρουπ είναι γνωστά σαν οι κοντινότεροι γείτονες και ουσιαστικά έχουν προτιμησιακή ομοιότητα με το χρήστη.
3. Παραγωγή σύστασης. Πρόταση προϊόντων στο χρήστη, με βάση το ποιά από αυτά έχουν αγοραστεί από άλλα μέλη του γκρουπ χρηστών με ταυτόσημα προτιμησιακά χαρακτηριστικά.

Τα συνεργατικά συστήματα συστάσεων χρησιμοποιούνται ευρέως, καθώς δεν απαιτούν ανάλυση περιεχομένου, που αποτελεί αρκετά πολύπλοκη διαδικασία σε σχέση με την ομαδοποίηση χρηστών, με βάση τη προτιμησιακή συμπεριφορά τους.

Οι Schafer et al., (1999), πρότειναν ότι η τεχνολογία συστάσεων όπου υπάρχει συσχέτιση ανθρώπων με άλλους ανθρώπους (People-to-People Correlation) είναι συνεργατικό σύστημα σύστασης που βασίζεται στη συσχέτιση μεταξύ χρηστών. Με βάση τους Carenini et al., (2003), Bobadilla et al., (2009), Breese et al., (1998b), τα συνεργατικά συστήματα συστάσεων διαχωρίζονται στις εξής δύο κατηγορίες: Βασιζόμενα στη Μνήμη (Memory based) και Βασιζόμενα στο Μοντέλο (Model based) συστήματα.

Τα βασιζόμενα στη μνήμη συνεργατικά συστήματα συστάσεων, χρησιμοποιούν τα αρχεία των βάσεων δεδομένων για ανάλυση, υπολογίζοντας μέτρα ομοιότητας και συσχέτισης, για να καταλήξουν σε αντικείμενα που έχουν την μεγαλύτερη συσχέτιση με την προτίμηση του χρήστη. Η πιο διαδεδομένη μέθοδος για τη συγκεκριμένη κατηγορία είναι οι κοντινότεροι γείτονες (Nearest Neighbors) (Breese et al., 1998b), (Pazzani and Billsus, 2007), (Goldberg et al., 1992).

Τα βασιζόμενα στο μοντέλο συνεργατικά συστήματα συστάσεων, χρησιμοποιούν μαθηματικά μοντέλα ή μοντέλα μηχανικής μάθησης για να υπολογίσουν ή να προβλέψουν ποια αντικείμενα προτιμούν οι χρήστες στόχοι (Breese et al., 1998b).

Πέραν της προαναφερθείσας κατηγοριοποίησης, οι Yang and Li, (2009) πρότειναν και την κατηγοριοποίηση με βάση τη συσχέτιση χρηστών, και τη

συσχέτιση αντικειμένων. Η βασική παραδοχή είναι, ότι οι πελάτες που παρουσιάζουν ίδια αγοραστική συμπεριφορά θα αγοράσουν και ίδια προϊόντα και ότι τα προϊόντα που αγοράζονται από συγκεκριμένους πελάτες παρουσιάζουν συσχέτιση, αντίστοιχα.

Οι βασικοί περιορισμοί των συνεργατικών συστημάτων συστάσεων είναι κυρίως η σποραδικότητα (sparsity), η επεκτασιμότητα (Scalability), η συνωνυμία (Synonymy) και η κρύα εκκίνηση (Cold Start).

Η σποραδικότητα παρουσιάζεται λόγω του μεγάλου πλήθους προϊόντων ή αντικειμένων που είναι αντικείμενα αξιολόγησης ή σύστασης. Ο πίνακας αξιολόγησης καθώς και τα προτιμησιακά προφίλ των χρηστών είναι δυνατόν να είναι πολύ διεσπαρμένα, με αποτέλεσμα να μην είναι δυνατός ο υπολογισμός της ομοιότητας. Έχουν προταθεί αρκετές λύσεις στον συγκεκριμένο περιορισμό από τους: Sarwar et al., (2001), Cho et al., (2002), Chen et al., (2008), Jeong et al., (2009), Wang, (2012), Kim and Ahn, (2017).

Η επεκτασιμότητα είναι ένα σύννηθες πρόβλημα όταν έχουμε να κάνουμε με μεγάλες βάσεις δεδομένων. Η μεγάλη ποσότητα δεδομένων σε μια βάση προκαλεί υπερφόρτωση του συστήματος (Chen et al., 2008), (Linden et al., 2003).

Η συνωνυμία αναφέρεται στο πρόβλημα όπου διαφορετικές λέξεις έχουν την ίδια ετυμολογία. Αυτό έχει σαν αποτέλεσμα, ένα συνεργατικό σύστημα σύστασης να αποτυγχάνει να διακρίνει την σημαντικότητα όρων που χρησιμοποιούνται για αξιολόγηση προϊόντων, ή να εντοπίσει την συσχέτιση μεταξύ προϊόντων πανομοιότυπων.

Η κρύα εκκίνηση αναφέρεται στην περίπτωση όπου δεν προϋπάρχουν δεδομένα σχετικά με κάποιο χρήστη στο σύστημα. Τότε το σύστημα δυσκολεύεται ή δεν μπορεί να προβεί σε σύσταση, καθώς δεν έχει τρόπο να υπολογίσει την ομοιότητα (Burke, (2002), Hahsler, (2011), Jeong et al., (2009)).

Τα προβλήματα κρύας εκκίνησης διακρίνονται σε τρεις κατηγορίες:

α. Κρύα εκκίνηση νέου συστήματος, που περιλαμβάνει την περίπτωση που το σύστημα μόλις ξεκινάει, και δεν έχει αρκετά δεδομένα αξιολογήσεων χρηστών

καθώς προφίλ χρηστών. Αυτό έχει σαν αποτέλεσμα να μην μπορεί να υπολογιστεί η ομοιότητα με ακρίβεια.

β. Κρύα εκκίνηση νέου χρήστη, όπου λόγω ανεπαρκών πληροφοριών για το προφίλ του χρήστη, είναι πιθανό η σύσταση να μην είναι αντίστοιχη του προτιμησιακού του μοντέλου (Burke, 2002).

γ. Κρύα εκκίνηση νέου αντικειμένου, όπου το νέο αντικείμενο στη βάση, δεν έχει λάβει κριτικές ή αγορασθεί από χρήστες.

### 2.2.1.3 Υβριδικά Συστήματα Συστάσεων

Τα υβριδικά συστήματα συστάσεων αποτελούν συνδυασμό συστημάτων συστάσεων βασιζόμενων στο περιεχόμενο (content based) και συνεργατικού φιλτραρίσματος (collaborative filtering). Το σημαντικό πλεονέκτημα τους είναι ότι αναλόγως του προβλήματος, χρησιμοποιούν αλγόριθμους ή τεχνικές διαφορετικού τύπου σύστασης, προκειμένου να καταλήξουν στη βέλτιστη σύσταση.

Τα υβριδικά συστήματα συστάσεων διαχωρίζονται σε 4 κατηγορίες:

- α. Ξεχωριστή εφαρμογή content based και collaborative filtering, πάνω στα δεδομένα και εν συνεχεία συνδυασμός των αποτελεσμάτων τους για τη σύσταση.
- β. Χρήση κάποιων από τα content based χαρακτηριστικά με collaborative filtering σύσταση.
- γ. Χρήση κάποιων από τα collaborative filtering χαρακτηριστικά με content based σύσταση.
- δ. Χρήση χαρακτηριστικών τόσο από collaborative filtering όσο και από content based σύσταση για την ανάπτυξη ενός νέου μοντέλου. (Burke, 2002)

### 2.2.2 Πολυκριτήρια Συστήματα αποφάσεων

Στον τομέα την επιστήμης των αποφάσεων ο βασικός στόχος της πολυκριτήριας ανάλυσης αποφάσεων είναι: να «βοηθήσει έναν αποφασίζοντα να επιλέξει την καλύτερη εναλλακτική δεδομένων των συνθηκών όταν υπάρχουν πολλά αντικρουόμενα κριτήρια που ανταγωνίζονται μεταξύ τους, εναρμονιζόμενη με τις προτιμήσεις του αποφασίζοντος».

Επιπρόσθετα στο τομέα του μάρκετινγκ η αγορά ενός προϊόντος μπορεί να θεωρηθεί ως ένα πολυκριτηριακό πρόβλημα απόφασης (Matsatsinis and Siskos, 1999, Siskos et al., 2001, Matsatsinis et al., 2003, Tsafarakis et al., 2010). Για παράδειγμα, για την αγορά ενός αυτοκινήτου λαμβάνουμε υπόψη πολλαπλά κριτήρια όπως η τιμή, η μάρκα, το χρώμα. Στο τομέα αυτό η πιο συνήθης μέθοδος για λύση τέτοιου είδους πολυκριτήριων προβλημάτων είναι η Conjoint (Green et al., 2001). Το συγκεκριμένο μοντέλο καθορίζει τα βάρη καθώς και τις τιμές των χαρακτηριστικών των προϊόντων. Εν συνεχεία υπολογίζονται οι προτιμήσεις των πελατών με ένα γραμμικό συνδυασμό των βαρών και των τιμών που έχουν καθοριστεί.

Οι Jelassi et al., (1990) διέκριναν τα Πολυκριτήρια συστήματα αποφάσεων σε πέντε γενιές, βασιζόμενοι στο χρόνο ανάπτυξης τους και επομένως στις διαθέσιμες για την εποχή δυνατότητες της τεχνολογίας των Η/Υ, στην πληρότητα, την αυτονομία, τον βαθμό ολοκλήρωσης τους καθώς και στο τρόπο επικοινωνίας μοντέλου-αποφασίζοντος.

Στην πρώτη γενιά τους τα πολυκριτήρια συστήματα περιλαμβάνουν μόνο μια πολυκριτήρια μέθοδο, όπου η επικοινωνία μηχανής-αποφασίζοντες είναι υποτυπώδης και η δομή των αρχείων δεδομένων εξαρτάται από το εκάστοτε πρόβλημα. Στα συστήματα αυτά υπήρχε πρόβλημα προσαρμοστικότητας (adaptivity), καθώς για κάθε διαφορετική εφαρμογή απαιτείτο επαναπρογραμματισμός τμημάτων. Επιπρόσθετα η διάσπαση τους σε πολλαπλά ανεξάρτητα τμήματα και η μη παραμετροποίηση των διαλόγων δυσχέραινε την αποτελεσματική χρησιμοποίησή τους.



Στη δεύτερη γενιά έγινε μια προσπάθεια βελτίωσης του συστήματος διαχείρισης των δεδομένων. Τα συστήματα αυτά συνοδεύονται από μια βάση δεδομένων καθώς και σύστημα διαχείρισης της.

Στη τρίτη γενιά εισάγεται και ένα ακόμα νέο κομμάτι, η βάση διαλόγου και το σύστημα διαχείρισης της, προκειμένου να βελτιωθεί η επικοινωνία διαχειριστή οθονών επιλογής διαδικασιών, το διαχειριστή διαλόγων διαχείρισης των δεδομένων και το γραφικό σύστημα για την οπτική παρουσίαση των αποτελεσμάτων. Τα συστήματα αυτής της γενιάς είναι και τα πιο διαδεδομένα.

Η τέταρτη γενιά εμπλουτίστηκε από περισσότερα του ενός μοντέλα και με λειτουργικότητα για δημιουργία νέων μοντέλων που να αποτελούν επέκταση των προϋπαρχόντων.

Τέλος στην πέμπτη γενιά ενσωματώθηκαν τεχνικές Τεχνητής νοημοσύνης. Συνοδευτικά των μοντέλων πολυκριτήριας ανάλυσης, με στόχο να μειώσουν τους περιορισμούς της δύσκολης επικοινωνίας μηχανής ανθρώπου που είναι ιδιαίτερα αισθητή στα πολύπλοκα προβλήματα, να συμβουλεύσουν στην επιλογή τους καταλληλότερου μοντέλου για κάθε διαφορετικό πρόβλημα, να αναβαθμίσουν την ποιότητα και αξιοπιστία του αλληλεπιδραστικού διαλόγου του συστήματος με τον αποφασίζοντα και να λειτουργήσουν συνεργατικά στις διαδικασίες λήψης αποφάσεων.

#### **2.2.2.1 Βασική Δομή των πολυκριτήριων συστημάτων λήψης απόφασης**

Η βασική δομή των πολυκριτήριων συστημάτων λήψης απόφασης συνίσταται σε τρία υποσυστήματα Ματσατσίνης, (2010):

- A. Το υποσύστημα Δεδομένων – Βάση δεδομένων και σύστημα διαχείρισης της.
- B. Το υποσύστημα Μοντέλων – Βάση μοντέλων και το σύστημα διαχείρισης της.
- Γ. Το υποσύστημα Διαλόγου – Βάση διαλόγου και το σύστημα διαχείρισης της.

#### A. Υποσύστημα Δεδομένων

Όλα τα πολυκριτήρια συστήματα λήψης αποφάσεων διαθέτουν ειδικό σύστημα διαχείρισης των δεδομένων. Τα δεδομένα προέρχονται:

- α. είτε από την μορφοποίηση των προβλημάτων και την μοντελοποίηση των κριτηρίων,
- β. είτε είναι αποτελέσματα της επεξεργασίας του συστήματος,
- γ. είτε έμμεσες ή άμεσες πληροφορίες που αντλούνται από τον αποφασίζοντα κατά τη διάρκεια της εφαρμογής τους.

#### Χαρακτηριστικά του συστήματος διαχείρισης βάσεων δεδομένων

1. Δυνατότητα συνδυασμού των δεδομένων μέσα από διαδικασίες ταξινόμησης και ανάκλησης με ή χωρίς συνθήκες.
2. Ταχεία διαχείριση των δεδομένων μέσα από διαδικασίες που παρέχει το σύστημα.
3. Δυνατότητα σύνθεσης των μεταβλητών ή ανάλυση τους σε περισσότερους μέσα από διαδικασίες παρεχόμενες από το σύστημα.
4. Διαχείριση πολλαπλών πηγών άντλησης δεδομένων.
5. Δυνατότητα άμεσης προσπέλασης στα δεδομένα της βάσης.
6. Επικοινωνία και διασύνδεση με τα άλλα υποσυστήματα.

#### B. Υποσύστημα Μοντέλων

Η βάση μοντέλων και το σύστημα διαχείρισης της αποτελούν το σημείο διαφοροποίησης των Πολυκριτήριων Συστημάτων Υποστήριξης Αποφάσεων από τα κλασικά Πληροφοριακά Συστήματα Δεδομένων. Τα περισσότερα διαθέτουν ένα σύνολο μοντέλων, τα οποία είτε είναι άμεση εφαρμογή της μεθοδολογίας και απαντούν στο σύνολο των προβληματικών, είτε συνοδεύουν το σύστημα προκειμένου να υποστηρίξουν τον αποφασίζοντα ή τον αναλυτή αποφάσεων στη διαμόρφωση και μοντελοποίηση του προβλήματος και σε τοπικά σημεία κρίσεων.

Χαρακτηριστικά της βάσης μοντέλων και του συστήματος διαχείρισης της

1. Δυνατότητα της άμεσης και αμφίδρομης επικοινωνίας με τα άλλα υποσυστήματα.
2. Παραμετρική δόμηση του προκειμένου να εφαρμόζονται σε όσο το δυνατόν περισσότερα προβλήματα διαφορετικής φύσης και προβληματικής.
3. Κατηγοριοποίηση των μοντέλων σύμφωνα με την προβληματική και τη φύση των προβλημάτων.
4. Επικοινωνία και συνεργασία των μοντέλων μεταξύ τους με τη δυνατότητα ανταλλαγής των δεδομένων και αποτελεσμάτων της επεξεργασίας τους.

Γ. Υποσύστημα Διαλόγου

Το υποσύστημα διαλόγου διαχειρίζεται την αμφίδρομη επικοινωνία του συστήματος με τον αποφασίζοντα διατηρώντας το ρόλο του συνδετικού κρίκου σε όλες τις φάσεις των επεξεργασιών. Η πληρότητα σε επιλογές και η φιλικότητα του είναι σημαντικοί παράγοντες για την λειτουργικότητα των Συστημάτων. Ειδικότερα οι παράγοντες αυτοί είναι καθοριστικοί στη διαχείριση των δεδομένων, στην υποστήριξη του αποφασίζοντος για την κατανόηση των προβλημάτων, στις αλληλεπιδραστικές διαδικασίες για την άντληση αξιόπιστων πληροφοριών από τον αποφασίζοντα και στην ανάλυση των αποτελεσμάτων παρέχοντας τα κατανοητά στη γλώσσα του αποφασίζοντος. Η λειτουργικότητα του υποσυστήματος εξαρτάται επίσης από την γνωστική μορφή του αποφασίζοντος καθώς υπάρχουν διαφοροποιήσεις που οφείλονται στην πολυπλοκότητα της γνωστικής μορφής και στην προσωπικότητα του.

Το υποσύστημα διαλόγου διαχωρίζεται σε τρία βασικά τμήματα:

α. Την ενεργό γλώσσα (Action Language), η οποία αποτελείται από το σύνολο των ενεργειών που μπορούν να γίνουν από τον αποφασίζοντα προκειμένου να επικοινωνήσει με το σύστημα, όπως το πληκτρολόγιο, τα ειδικά πλήκτρα οι οθόνες επιλογών, το ποντίκι.

β. Τη γλώσσα παρουσίασης (Presentation Language), που αποτελεί το σύνολο των εξόδων του συστήματος προς τον αποφασίζοντα. Το τμήμα αυτό περιλαμβάνει πίνακες, μηνύματα, γραφήματα. Η ποικιλία των μέσων και των μορφών παρουσίασης συνιστούν το σημαντικότερο παράγοντα για την αξιοπιστία των πληροφοριών και της αλληλεπίδρασης του αποφασίζοντος με το σύστημα.

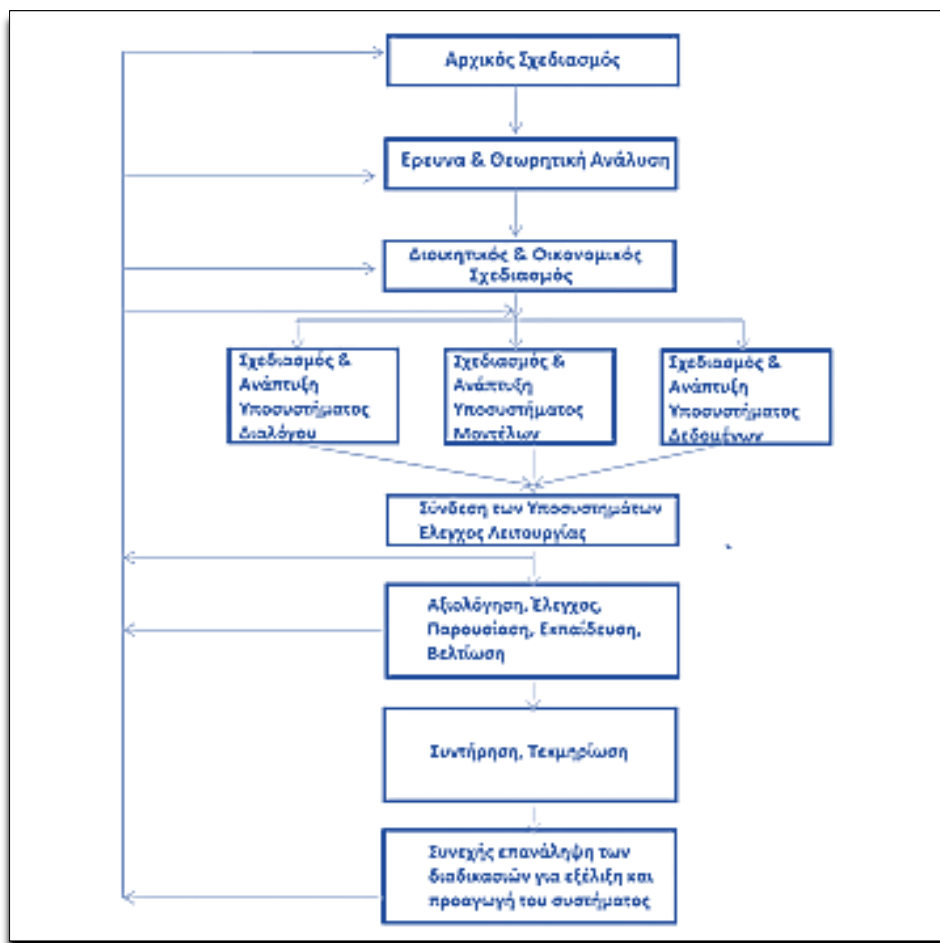
γ. Τη βάση υποστήριξης (Support Base), που αφορά όλα όσα σχετίζονται με τη υποστήριξη χειρισμού του συστήματος.

Χαρακτηριστικά του υποσυστήματος διαλόγου

1. Παρουσίαση των πληροφοριών και αποτελεσμάτων στη γλώσσα του αποφασίζοντος.
2. Πλουραλισμός στην επικοινωνία με τον αποφασίζοντα, ώστε να καλύπτονται οι διαφοροποιήσεις των γνωστικών μορφών και των αποφασιζόντων.
3. Ευκολία στην μάθηση και χρήση του συστήματος
4. Καθοδήγηση του τελικού χρήστη στη διαχείριση του συστήματος και στις διαδικασίες της μεθοδολογίας.
5. Μείωση της ασάφειας που δημιουργείται στο χειριστή σε τοπικά σημεία.

#### **2.2.2.2 Μεθοδολογία Ανάπτυξης Πληροφοριακού Συστήματος Υποστήριξης Αποφάσεων**

Οι διαδικασίες ανάπτυξης Πληροφοριακού Συστήματος Υποστήριξης Αποφάσεων, περικλείουν μεγάλο βαθμό πολυπλοκότητας καθώς για την ολοκληρωμένη και τελική έκδοση τους συμμετέχουν πολλοί παράγοντες, όπως το υλικό (Hardware), η τεχνική υποδομή, η συμπεριφορά των αποφασιζόντων, η επικοινωνία ανθρώπου μηχανής, το περιβάλλον (Ματσατσίνης, 2010). Στο παρακάτω σχήμα (Σχήμα 2-1) παρουσιάζονται οι φάσεις ανάπτυξης συστήματος με βάση τις μεθόδους που παρουσιάστηκαν από τους Meador et al., 1984.



Σχ. 2-1 Φάσεις Ανάπτυξης Πληροφοριακού Συστήματος Υποστήριξης Αποφάσεων

Οι βασικές φάσεις που ακολουθούνται, όπως φαίνεται και στο σχήμα 2-1, είναι οι ακόλουθες:

- Α. Αρχικός σχεδιασμός έργου (Planning): Στη φάση αυτή καθορίζονται οι στόχοι και τα αντικείμενα του υπό ανάπτυξη συστήματος
- Β. Έρευνα και θεωρητική ανάλυση (Searching): Καταγραφή και διερεύνηση του θεωρητικού υποβάθρου του προβλήματος. Καταγραφή των απαιτήσεων .
- Γ. Διοικητικός και οικονομικός σχεδιασμός (Design): Μελέτη και καταγραφή της βέλτιστης προσέγγισης για την υλοποίηση του έργου. Πρόβλεψη και σχεδιασμός των πόρων (Προσωπικό, οικονομικά, Υλικά) και καταμερισμός τους μέσα στο διάγραμμα του έργου.
- Δ. Ανάπτυξη του συστήματος (Development): Καθορισμός των τεχνικών χαρακτηριστικών και λεπτομερής λογικός σχεδιασμός των εκάστοτε

υποσυστημάτων. Επιλογή εργαλείων ανάπτυξης λογισμικού, μεθοδολογίας και γλώσσας προγραμματισμού.

Ε. Σύνδεση των υποσυστημάτων (Integration) και ταυτόχρονος έλεγχος λειτουργίας και συνεργασίας τους για το ολοκληρωμένο σύστημα.

ΣΤ. Έλεγχος Λειτουργίας (Test): Συλλογή στοιχείων από τη συμπεριφορά του συστήματος και διασταύρωση με τις απαιτήσεις σχεδιασμού.

Ζ. Αξιολόγηση του συστήματος (Evaluation): Αξιολόγηση του συστήματος σε σχέση με την πληρότητα κάλυψης των αναγκών των χρηστών. Η διαδικασία αυτή επαναλαμβάνεται συνεχώς σε κάθε αλλαγή ή βελτίωση του συστήματος.

Η. Παρουσίαση και εγκατάσταση του συστήματος (Presentation and Installation): Παρουσίαση του συστήματος και εγκατάσταση για πιλοτική εφαρμογή. Εκπαίδευση των χρηστών του.

Θ. Συντήρηση – Τεκμηρίωση (Maintenance): Συγγραφή εγχειριδίων χρήσης και τεκμηρίωσης καθώς και αναμορφώσεις/αναβαθμίσεις με βάση σχόλια των χρηστών.

Με βάση την προαναφερθείσα προσέγγιση ο Courbon, (1987) πρότεινε τη μέθοδο Πρωτοτύπου (Prototype) ή Επαναληπτική (Iterative) που ενδείκνυται για ανάπτυξη Πληροφοριακών συστημάτων. Η συγκεκριμένη μέθοδος αποτελείται από τα παρακάτω βήματα τα οποία είναι επαναλαμβανόμενα:

1. Κατασκευή ενός πρωτόλειου, δηλαδή ενός μικρού και απλού συστήματος ακολουθώντας τα στάδια ανάπτυξης Πληροφοριακών Συστημάτων Υποστήριξης αποφάσεων.
2. Αξιολογείται το Σύστημα σε συνεργασία αποφασιζόντων και κατασκευαστών. Προτείνονται αλλαγές και βελτιώσεις.
3. Επανασχεδιάζεται το σύστημα με βελτιώσεις και επεκτάσεις ως την τελειοποίηση του.

Κατά τη μέθοδο αυτή οι πραγματικές απαιτήσεις των αποφασιζόντων προσεγγίζονται σταδιακά με παράλληλη υλοποίηση τους στους επαναλαμβανόμενους κύκλους. Η επικοινωνία και διαδραστική σχέση μεταξύ κατασκευαστών και τελικών χρηστών περιέχει συνεχείς παρεμβάσεις διόρθωσης και επέκτασης μέχρι την ολοκλήρωση του συστήματος.

Τα πλεονεκτήματα της παραπάνω μεθόδου είναι τα παρακάτω:

1. Μικρός χρόνος υλοποίησης
2. Μεγάλη αποδοχή του τελικού προϊόντος από τους τελικούς χρήστες.
3. Μεγάλη σύγκλιση ως προς τις πραγματικές απαιτήσεις των τελικών χρηστών.
4. Αυξημένη συμμετοχή των τελικών χρηστών στο σχεδιασμό και ανάπτυξη του συστήματος.

Το κύριο μειονέκτημα της μεθόδου είναι ο κίνδυνος απώλειας του ελέγχου κατά τις επαναλήψεις των σταδίων, κυρίως όταν οι ομάδες υλοποίησης είναι μεγάλες, ιδιαίτερα από πλευράς χρηστών.

Τέλος αναφέρεται το ιδιαίτερα σημαντικό περίγραμμα προσέγγισης Representation, Operations, Memory aids και Control Mechanism approach (ROMC) για το σχεδιασμό και την ανάλυση των Πληροφοριακών Συστημάτων Υποστήριξης αποφάσεων που είχε προταθεί από τον Sprague, 1980. Στόχος του περιγράμματος είναι ο καθορισμός των χαρακτηριστικών και δυνατοτήτων του υπό ανάπτυξη συστήματος που πραγματοποιείται με συνεργασία των κατασκευαστών και των χρηστών. Το περίγραμμα ROMC που είναι ανεξάρτητο διαδικασιών προσανατολίζει τις αναζητήσεις κατασκευαστών και χρηστών στις τέσσερις βασικές παρακάτω οντότητες.

A. Αναπαραστάσεις (Representations): Εκφράζει τη δυνατότητα να παρέχονται παρουσιάσεις από το σύστημα για την κατανόηση του προβλήματος και την επικοινωνία μοντέλου- αποφασίζοντα.

B. Λειτουργίες (Operations): Περιλαμβάνει όλες τις διεργασίες που αναλύουν και διαχειρίζονται τις παρουσιάσεις.

Γ. Βοηθητικά μνήμης (Memory aids): Βοηθούν στη σύνδεση παρουσιάσεων και διεργασιών.

Δ. Μηχανισμοί ελέγχου (Control Mechanisms): Συντελούν στο έλεγχο και την εύρυθμη λειτουργία του συστήματος, διασφαλίζοντας την ομαλή λειτουργία των προαναφερθεισών οντοτήτων.

### 2.2.3 Η Σύσταση ως ένα πολυκριτήριο πρόβλημα λήψης απόφασης

Για να υλοποιηθεί η ανάλυση ενός προβλήματος λήψης απόφασης, σύμφωνα με τη μεθοδολογία του Roy, (1990) ακολουθούνται τα ακόλουθα βήματα:

1. Προσδιορισμός του αντικειμένου της απόφασης. Δηλαδή ορισμός του συνόλου των εναλλακτικών για τις οποίες θα ληφθεί απόφαση καθώς και η λογική/προβληματική της απόφασης.
2. Ορισμός μια συνεπούς οικογένειας κριτηρίων. Δηλαδή προσδιορισμός ενός συνόλου συναρτήσεων που δηλώνουν τις προτιμήσεις του αποφασίζοντα σε σχέση με τις υπάρχουσες εναλλακτικές.
3. Ανάπτυξη ενός συνολικού μοντέλου προτίμησης. Δηλαδή ορισμός της συνάρτησης που συνθέτει τις μερικές προτιμήσεις σε κάθε κριτήριο και ανάπτυξη του συνολικού μοντέλου προτίμησης ενός αποφασίζοντα σε σχέση με μια υποψήφια εναλλακτική.
4. Επιλογή της διαδικασίας υποστήριξης απόφασης. Δηλαδή σχεδιασμός της διαδικασίας, των μεθόδων ή λογισμικών που θα υποστηρίξουν τον αποφασίζοντα όταν λαμβάνει την απόφαση σχετικά με το σύνολο των εναλλακτικών, με βάση τα αποτελέσματα των παραπάνω βημάτων.

#### Αντικείμενο Απόφασης

Στα συστήματα συστάσεων το αντικείμενο απόφασης είναι το αντικείμενο  $i$  που ανήκει στο σύνολο των υποψήφιων αντικειμένων. Τα στοιχεία αυτού του συνόλου καλούνται εναλλακτικές δραστηριότητες ή εναλλακτικές στην υπάρχουσα βιβλιογραφία. Για να εκφράσει το σκεπτικό πίσω από την απόφαση ο Roy, (1990) αναφέρεται στον όρο προβληματική της απόφασης. Γενικά υπάρχουν τέσσερις προβληματικές που καλύπτουν το σύνολο των πρακτικών περιπτώσεων:

1. Επιλογή (Choice). Αφορά την επιλογή μιας ή περισσότερων εναλλακτικών οι οποίες θεωρούνται ως οι πλέον κατάλληλες ανάμεσα στις υποψήφιες.



2. Ταξινόμηση (Sorting). Αφορά την ταξινόμηση των εναλλακτικών δραστηριοτήτων σε προκαθορισμένες κατηγορίες.
3. Κατάταξη (Ranking). Αφορά την κατάταξη των εναλλακτικών δραστηριοτήτων από τις καλύτερες προς τις χειρότερες.
4. Περιγραφή (Description). Αφορά την περιγραφή των εναλλακτικών δραστηριοτήτων βάσει των επιδόσεων τους στα επιμέρους κριτήρια αξιολόγησης.

Και οι 4 προβληματικές απόφασης θεωρούνται έγκυρες για ένα πρόβλημα σύστασης:

1. Επιλογή και σύσταση ενός ή περισσότερων αντικειμένων ως καταλληλότερα για ένα συγκεκριμένο χρήστη.
2. Ταξινόμηση όλων των διαθέσιμων αντικειμένων σε προκαθορισμένες κατηγορίες σύμφωνα με την καταλληλότητα τους π.χ. προτείνονται για αγορά, προτείνονται για απλή ματιά.
3. Κατάταξη των διαθέσιμων αντικειμένων από το καταλληλότερο προς το λιγότερο κατάλληλο για ένα συγκεκριμένο χρήστη και παρουσίαση της λίστας κατάταξης στον χρήστη.
4. Περιγραφή της καταλληλότητας ενός αντικειμένου για ένα συγκεκριμένο χρήστη με βάση πόσο έχει βαθμολογηθεί για κάθε κριτήριο.

### Οικογένεια Κριτηρίων

Η επίδοση των εναλλακτικών σε σχέση με το σύνολο των κριτηρίων αναλύεται για κάθε χρήστη προκειμένου να μοντελοποιηθούν όλα τα χαρακτηριστικά τους, οι επιπτώσεις τους και η συνέπειες τους. Στα συστήματα συστάσεων τα κριτήρια είναι δυνατόν να αναφέρονται σε πολλαπλά χαρακτηριστικά ενός αντικειμένου ή σε πολλαπλές διαστάσεις με βάση τις οποίες αξιολογείται το αντικείμενο.

Για τη λήψη ορθολογικών αποφάσεων με πολλαπλά κριτήρια θα πρέπει να διασφαλίζεται ότι το σύνολο των εξεταζόμενων κριτηρίων διαμορφώνει μια συνεπή οικογένεια κριτηρίων. Μια οικογένεια κριτηρίων θεωρείται συνεπής όταν διαθέτει τις ακόλουθες τρεις ιδιότητες:

1. Μονοτονία. Για την περίπτωση της αύξουσας κλίμακας: “Όσο μεγαλύτερη είναι η τιμή του κριτηρίου για μία εναλλακτική ενέργεια, τόσο η ενέργεια αυτή προτιμάται”, ενώ για την περίπτωση της φθίνουσας κλίμακας: “Όσο μικρότερη είναι η τιμή του κριτηρίου για μία εναλλακτική ενέργεια, τόσο η ενέργεια αυτή προτιμάται”.

Έστω  $c_i(a)$ , η τιμή του κριτηρίου  $c_i$ , ως αποτέλεσμα της ενέργειας  $a$ . Τότε, η μονοτονία έχει τις εξής ιδιότητες :

- αν  $c_i(a) > c_i(b) \leftrightarrow a > b$  (η  $a$  προτιμάται της  $b$ : αύξουσα κλίμακα)
- αν  $c_i(a) > c_i(b) \leftrightarrow b > a$  (η  $b$  προτιμάται της  $a$ : φθίνουσα κλίμακα)
- αν  $c_i(a) = c_i(b) \leftrightarrow a \sim b$  (αδιαφορία)

Η μονοτονία αναφέρεται σε κάθε κριτήριο ξεχωριστά. Κάθε χαρακτηριστικό, που δεν πληροί τη συνθήκη της μονοτονίας, δεν μπορεί να αποτελέσει κριτήριο.

2. Επάρκεια. “Μέσα σε μια συνεπή οικογένεια των κριτηρίων, υπάρχουν όλα τα κριτήρια. Δεν λείπει κριτήριο”. Με άλλα λόγια, εξετάζουμε το πρόβλημα απόφασης με πλήρη και σφαιρικό τρόπο, φροντίζοντας να μη λείπει κάποιο σημαντικό κριτήριο. Η επάρκεια αναφέρεται στο σύνολο των κριτηρίων.
3. Μη πλεονασμός. Ένα σύνολο κριτηρίων διαθέτει την ιδιότητα του μη πλεονασμού εάν και μόνο εάν η διαγραφή ενός οποιουδήποτε κριτηρίου οδηγεί σε παραβίαση των ιδιοτήτων της μονοτονίας ή της επάρκειας. Επίσης, τα κριτήρια δεν πρέπει να είναι αλληλεξαρτώμενα (αρχή της ανεξαρτησίας). Στην περίπτωση που υπάρχει πλεονασμός, σημαίνει ότι κάποια άποψη θεώρησης του προβλήματος απόφασης υπερβαροδοτείται (δίνεται πολύ μεγάλη σημαντικότητα), σε βάρος των υπολοίπων απόψεων (κριτηρίων) θεώρησης.

### Συνολικό Μοντέλο Προτίμησης

Η ανάπτυξη ενός συνολικού μοντέλου προτίμησης παρέχει ουσιαστικά ένα τρόπο για να συναθροιστούν οι τιμές για κάθε κριτήριο  $c_k$  (όπου  $c = 1, \dots, k$ ) ώστε να εκφραστούν οι προτιμήσεις ανάμεσα στις διαφορετικές εναλλακτικές

για το σύνολο των αντικειμένων, ανάλογα με την προβληματική απόφασης που έχει επιλεγεί. Σύμφωνα με τους (Horst et al., 1995) προκύπτει η ακόλουθη κατηγοριοποίηση:

1. Πολυκριτήριο μαθηματικός προγραμματισμός
2. Θεωρία πολυκριτήριας Αξίας/Χρησιμότητας
3. Θεωρία των σχέσεων υπεροχής
4. Αναλυτική- Συνθετική Προσέγγιση.

Μεθοδολογίες από όλες τις προαναφερθείσες κατηγορίες είναι δυνατό να χρησιμοποιηθούν για τη δημιουργία μοντέλου ολικής προτίμησης για συστήματα συστάσεων, ανάλογα με την προβληματική απόφασης που έχει επιλεγεί και το περιβάλλον στο οποίο αναμένεται να λειτουργήσει το σύστημα συστάσεων.

#### Διαδικασία Υποστήριξης Απόφασης

Στο στάδιο αυτό λαμβάνεται μια απόφαση για ένα δεδομένο πρόβλημα MCDM επιλέγοντας την κατάλληλη μέθοδο από αυτές που περιγράφηκαν παραπάνω.

Τα πολυκριτήρια συστήματα συστάσεων κατηγοριοποιούνται ανάλογα με την προβληματική απόφασης που υποστηρίζουν, τους τύπους κριτηρίων που χρησιμοποιούν και το μοντέλο ολικής προτίμησης που χρησιμοποιούν.

Είναι σημαντικό να σημειωθεί ότι το μεγαλύτερο βάρος της έρευνας έχει πέσει στη προβληματική απόφασης της κατάταξης αντικειμένων. Υπάρχουν επίσης αρκετά συστήματα που υποστηρίζουν την ταξινόμηση αντικειμένων σύμφωνα με τη καταλληλότητα τους για τον χρήστη. Πολύ λίγα συστήματα υποστηρίζουν την προβληματική της επιλογής και της περιγραφής. Επιπρόσθετα υπάρχουν κάποια συστήματα που χρησιμοποιούν ασαφή, τακτικά και πιθανολογικά δεδομένα για να εκφράσουν τις προτιμήσεις των χρηστών σε αντίστοιχα αντικείμενα. Ελάχιστα συστήματα συστάσεων χρησιμοποιούν πολυκριτήριο μαθηματικό προγραμματισμό ή θεωρία των σχέσεων υπεροχής για την δημιουργία του ολικού μοντέλου προτίμησης. Η πλειονότητα των συστημάτων χρησιμοποιεί τη θεωρία της πολυκριτήριας αξίας και υπολογίζει την πρόβλεψη με βάση μια αθροιστική συνάρτηση αξιών (Adomavicius and Kwon, 2015, Manouselis and Costopoulou, 2007b).

Τέλος πρέπει να σημειωθεί ότι τα πειραματικά αποτελέσματα δείχνουν ότι τα πολυκριτήρια συστήματα συστάσεων επιτυγχάνουν ικανοποιητικές επιδόσεις ακόμα και σε περιπτώσεις που η μοντελοποίηση δεν έχει ακολουθήσει όλους τους περιορισμούς.

### 2.2.3.1 Κατηγορίες Πολυκριτήριων Συστημάτων Συστάσεων

Τα Πολυκριτήρια Συστήματα Συστάσεων κατηγοριοποιούνται ως ακολούθως: content based, collaborative filtering, knowledge based, hybrid, (Adomavicius and Tuzhilin, 2005b, Li et al., 2012).

- Τα content based στηρίζονται στην παρελθοντική επιλογή του χρήστη προκειμένου να προβούν σε σύσταση, ενώ τα collaborative filtering προτείνουν με βάση πληροφορίες για άλλους χρήστες με παρόμοιες προτιμήσεις.
- Τα knowledge based χρησιμοποιούν γνώση σχετική με τους χρήστες και τα αντικείμενα και βρίσκουν αντικείμενα που ταιριάζουν στους χρήστες.
- Τέλος τα hybrid συνδυάζουν όλες τις προαναφερθείσες μεθοδολογίες .

Τα πολυκριτήρια Συστήματα συστάσεων είναι δυνατόν επίσης να κατηγοριοποιηθούν ως εξής:

1. Multi-attribute content preference modeling. Τα συγκεκριμένα συστήματα προσπαθούν να αντιληφθούν και να μοντελοποιήσουν τα κοινά στοιχεία του πολυπαραγοντικού περιεχομένου ανάμεσα στα αντικείμενα που επέλεξε ο χρήστης στο παρελθόν και να προτείνουν στον χρήστη αυτά τα αντικείμενα που ταιριάζουν περισσότερο στις επιλογές του.
2. Multi-attribute content search and filtering. Τα συστήματα αυτά επιτρέπουν στο χρήστη να καθορίσει τις γενικές προτιμήσεις του σε στοιχεία περιεχομένου των αντικειμένων μέσα από μεθόδους αναζήτησης ή φιλτραρίσματος και προτείνουν στο χρήστη τα αντικείμενα που είναι

πιο κοντά στις προτιμήσεις και ταυτόχρονα ικανοποιούν τις καθορισμένες συνθήκες αναζήτησης ή φιλτραρίσματος.

3. Multi-criteria rating-based preference elicitation. Τα συγκεκριμένα συστήματα επιτρέπουν στο χρήστη να καθορίσει τις προσωπικές του προτιμήσεις βαθμολογώντας το κάθε αντικείμενο σε πολλαπλά κριτήρια και προτείνει στο χρήστη τα αντικείμενα που αντανakλούν τις προτιμήσεις του στηριζόμενα σε πολυκριτήριες αξιολογήσεις που παρέχονται από τον ίδιο και άλλους χρήστες.

#### **2.2.3.2 Αλγόριθμοι για Πολυκριτήρια Συστήματα Συστάσεων**

Τα συστήματα συστάσεων συνήθως υπολογίζουν και παρέχουν συστάσεις χρησιμοποιώντας την ακόλουθη διαδικασία:

1. Πρόβλεψη. Φάση κατά την οποία υπολογίζεται η πρόβλεψη της προτίμησης ενός χρήστη. Στη φάση αυτή το σύστημα σύστασης υπολογίζει τη συνάρτηση χρησιμότητας  $R$  για το σύνολο ή για μέρος τους σετ Χρήστες-Αντικείμενα στηριζόμενα σε γνωστές αξιολογήσεις και πιθανώς σε άλλη πληροφορία.
2. Σύσταση. Φάση κατά την οποία η υπολογισμένη πρόβλεψη χρησιμοποιείται για να υποστηρίξει την απόφαση του χρήστη με κάποια διαδικασία σύστασης (Adomavicius et al., 2011a, Liu et al., 2011).

Η πληροφορία που έχει προκύψει από την πολυκριτήρια αξιολόγηση μπορεί να χρησιμοποιηθεί και στις 2 φάσεις με διαφορετικούς τρόπους. Έτσι διαχωρίζουμε τις υπάρχουσες τεχνικές για πολυκριτήρια συστήματα συστάσεων ως εξής: 1. Τεχνικές που χρησιμοποιούνται κατά τη διαδικασία της πρόβλεψης και 2. Τεχνικές που χρησιμοποιούνται κατά τη διαδικασία της σύστασης.

### 2.2.3.2.1 Τεχνικές Πολυκριτήριας Αξιολόγησης κατά την διαδικασία της Πρόβλεψης

Οι τεχνικές πολυκριτήριας αξιολόγησης για συστήματα συστάσεων διαχωρίζονται με βάση τη συνάρτηση αξίας σε 2 κατηγορίες: 1. Ευρετικές (heuristic based), 2. Βασιζόμενες στο Μοντέλο (Model based).

Οι heuristic-based υπολογίζουν την αξία κάθε αντικειμένου για ένα χρήστη στηριζόμενες στα δεδομένα που προέρχονται από το χρήστη και τυπικά βασίζονται σε μια συγκεκριμένα ευρετική υπόθεση.

Οι model-based μαθαίνουν ένα μοντέλο πρόβλεψης, χρησιμοποιώντας στατιστικές ή machine learning μεθόδους και μπορούν να ερμηνεύσουν καλύτερα τα υπάρχοντα δεδομένα. Στη συνέχεια χρησιμοποιούν το μοντέλο που έχουν μάθει για να υπολογίσουν την αξία των άγνωστων αντικειμένων για συστάσεις.

#### Heuristic Based τεχνικές

Το πρώτο στάδιο στη διαδικασία πρόβλεψης είναι η επιλογή της μεθόδου υπολογισμού ομοιότητας ώστε να βρεθεί ένα σύνολο γειτόνων με κοινά χαρακτηριστικά για κάθε χρήστη. Στα συστήματα συστάσεων με ένα μόνο κριτήριο αξιολόγησης, οι πιο γνωστές μέθοδοι είναι η συσχέτιση (correlation-based) και η συν-ημιτονοειδής (cosine-based). Θεωρώντας ότι  $R(u,i)$  είναι η αξιολόγηση που ο χρήστης  $u$  κάνει για το αντικείμενο  $p$ , και  $P(u,u')$  είναι τα κοινά αντικείμενα που έχουν αξιολογήσει 2 χρήστες  $u, u'$ , δύο γνωστά μέτρα ομοιότητας (similarity measures) που χρησιμοποιούνται είναι τα εξής:

Pearson Correlation-based:

$$\text{sim}(u, u') = \frac{\sum_{p \in P(u, u')} (R(u, p) - \bar{R}(u)) (R(u', p) - \bar{R}(u'))}{\sqrt{\sum_{p \in P(u, u')} (R(u, p) - \bar{R}(u))^2} \sqrt{\sum_{p \in P(u, u')} (R(u', p) - \bar{R}(u'))^2}} \quad (2-1)$$

Cosine Based:

$$\text{sim}(u, u') = \frac{\sum_{p \in P(u, u')} (R(u, p) - \bar{R}(u)) (R(u', p) - \bar{R}(u'))}{\sqrt{\sum_{p \in P(u, u')} (R(u, p) - \bar{R}(u))^2} \sqrt{\sum_{p \in P(u, u')} (R(u', p) - \bar{R}(u'))^2}} \quad (2-2)$$

Τα πολυκριτήρια συστήματα συστάσεων δεν μπορούν να χρησιμοποιήσουν απευθείας τα παραπάνω μέτρα καθώς το  $R(u,i)$  περιέχει μια συνολική αξιολόγηση  $r_o$  και  $k$  πολυκριτηριακές αξιολογήσεις  $r_1, \dots, r_k$ . Συνεπώς υπάρχουν  $k+1$  τιμές αξιολογήσεων για κάθε ζευγάρι  $(u,p)$  αντί για μία απλή αξιολόγηση. Οι Adomavicius et al., (2011b) προτείνουν δύο μεθόδους για μέτρηση της ομοιότητας (similarity): 1. Μέση Ομοιότητα (Average Similarity), 2. Ομοιότητα χειρότερης περίπτωσης (Worst Case Similarity). Οι Tang and McCalla, (2009) υπολογίζουν μια συνολική ομοιότητα ως σταθμισμένο άθροισμα ξεχωριστών ομοιοτήτων για πολλά κριτήρια. Στην προσέγγιση τους το βάρος για κάθε κριτήριο  $c$ , ορίζεται σαν  $w_c$ , και δείχνει τη σημαντικότητα του κριτηρίου για την τελική σύσταση.

- Average Similarity

$$sim_{avg}(u, u') = \frac{1}{k+1} \sum_{c=0}^k sim_c(u, u') \quad (2-3)$$

- Worst-Case Similarity:

$$sim_{min}(u, u') = \min sim_c(u, u') , c = 0 \dots k \quad (2-4)$$

- Aggregate Similarity:

$$sim_{aggregate}(u, u') = \sum_{c=0}^k w_c sim_c(u, u') \quad (2-5)$$

Οι Manouselis and Costopoulou, (2007c) επίσης προτείνουν τρεις διαφορετικούς αλγόριθμους υπολογισμού της ομοιότητας μεταξύ χρηστών σε πολυκριτήριες αξιολογήσεις: 1. Similarity per priority, 2. Similarity per Evaluation, 3. Similarity per partial Utility. Ο 1<sup>ος</sup> υπολογίζει τις ομοιότητες μεταξύ των χρηστών σε συνάρτηση με τα βάρη του κάθε χρήστη για κάθε κριτήριο. Έτσι δημιουργεί μια γειτονιά χρηστών που έχουν παρόμοιες προτιμήσεις με τον χρήστη-στόχο. Στη συνέχεια προσπαθεί να προβλέψει τη συνολική αξία ενός αντικειμένου για τον χρήστη βασιζόμενος στις συνολικές αξίες των χρηστών που γειτνιάζουν. Ο 2<sup>ος</sup> υπολογίζει την ομοιότητα με βάση τις αξιολογήσεις που δεν έχουν βάρος και τις οποίες έχουν κάνει οι χρήστες για κάθε κριτήριο. Ο 3<sup>ος</sup> υπολογίζει την ομοιότητα με βάση τις σταθμισμένες αξιολογήσεις που έχουν κάνει οι χρήστες για κάθε κριτήριο.

Μετά τον υπολογισμό των ομοιοτήτων μεταξύ των χρηστών με βάση τις πολυκριτήριες αξιολογήσεις, η υπόλοιπη διαδικασία σύστασης είναι ίδια όπως για τα μονοκριτηριακά συστήματα.

### Model-Based Τεχνικές

Οι συγκεκριμένες τεχνικές κατασκευάζουν ένα μοντέλο πρόβλεψης για να υπολογίσουν άγνωστες αξιολογήσεις μαθαίνοντας από τα υπάρχοντα δεδομένα. Οι βασικές προσεγγίσεις που εφαρμόζονται για τις περιπτώσεις των πολυκριτήριων συστημάτων συστάσεων αυτής της κατηγορίας είναι οι εξής:

1. Aggregation Function Approach,
2. Probabilistic Modeling,
3. Multi-linear Singular Value Decomposition(MSVD).

Η 1<sup>η</sup> ουσιαστικά υποθέτει ότι η συνολική αξιολόγηση προκύπτει από ένα συνταίριασμα των πολυκριτήριων αξιολογήσεων. Η aggregation function προκύπτει από στατιστικές τεχνικές όπως η γραμμική και μη γραμμική παλινδρόμηση ή από machine learning τεχνικές όπως τα τεχνητά νευρωνικά δίκτυα.

Η 2<sup>η</sup> υποθέτει ότι η συνολική αξιολόγηση για ένα άγνωστο αντικείμενο για ένα συγκεκριμένο χρήστη προκύπτει από 2 βήματα: 1. Εκμάθηση, 2. Πρόβλεψη. Στο 1<sup>ο</sup> βήμα όλες οι παράμετροι του μοντέλου υπολογίζονται χρησιμοποιώντας των αλγόριθμο μεγιστοποίησης προσδοκίας (Expectation Maximization Algorithm). Με χρήση των παραμέτρων αυτών στο 2<sup>ο</sup> βήμα προβλέπεται η συνολική αξιολόγηση για ένα δεδομένο άγνωστο αντικείμενο σαν η πιο πιθανή τιμή.

Η 3<sup>η</sup> προσέγγιση χρησιμοποιείται για δεδομένα πολλαπλών διαστάσεων και ουσιαστικά μειώνει τις διαστάσεις των πολυκριτήριων δεδομένων. Με τη χρήση της MSVD είναι δυνατόν ανακαλυφθούν οι σχέσεις μεταξύ χρηστών, αντικειμένων και κριτηρίων και στη συνέχεια χρησιμοποιείται αυτή η πληροφορία για να εντοπιστούν οι πλησιέστεροι γείτονες για κάθε χρήστη ώστε τελικά να πραγματοποιηθούν οι συστάσεις.(Adomavicius et al., 2011b)



### 2.2.3.2.2 Τεχνικές Πολυκριτήριας Αξιολόγησης κατά τη διαδικασία της Σύστασης

Τα πολυκριτήρια συστήματα συστάσεων μπορούν να επιλέξουν να μοντελοποιήσουν είτε την αξία ενός αντικειμένου για ένα χρήστη ενσωματώνοντας τη συνολική αξιολόγηση και τις αξιολογήσεις του κάθε κριτηρίου ενός αντικειμένου, είτε να ενσωματώσουν αξιολογήσεις ξεχωριστών κριτηρίων. Στην πρώτη περίπτωση η διαδικασία σύστασης που ακολουθείται είναι η εξής: αφού γίνει η πρόβλεψη όλων των άγνωστων αξιολογήσεων, το σύστημα σύστασης χρησιμοποιεί τη συνολική αξιολόγηση των αντικειμένων για να επιλέξει τα πλέον προτιμητέα για κάθε χρήστη. Στην δεύτερη περίπτωση το πρόβλημα είναι αρκετά πιο σύνθετο, καθώς δεν είναι προφανής η κατάταξη των αντικειμένων. Ανάμεσα στις προσεγγίσεις που χρησιμοποιούνται στη βιβλιογραφία των συστημάτων συστάσεων για τη δεύτερη περίπτωση κάποιες προσπαθούν να σχεδιάσουν μια συνολική κατάταξη των αντικειμένων και να καταλήξουν σε μια ολική βέλτιστη λύση για κάθε χρήστη ενώ κάποιες άλλες παίρνουν μία από τις πιθανές μερικές κατατάξεις των αντικειμένων και υπολογίζουν πολλαπλές λύσεις (Adomavicius et al., 2011b).

#### Πολυκριτήρια Βελτιστοποίηση

Οι πιο συχνές τεχνικές που χρησιμοποιούνται σε πολυκριτήρια προβλήματα βελτιστοποίησης και είναι δυνατόν να εφαρμοστούν σε συστήματα συστάσεων είναι οι εξής:

1. Εύρεση των βέλτιστων λύσεων κατά Pareto.
2. Γραμμικός συνδυασμός των πολλαπλών κριτηρίων ώστε το πρόβλημα να αναχθεί σε μονού κριτηρίου πρόβλημα βελτιστοποίησης.
3. Βελτιστοποίηση μόνο του πιο σημαντικού κριτηρίου και μετατροπή όλων των υπολοίπων κριτηρίων σε περιορισμούς.

Βελτιστοποίηση ενός κριτηρίου κάθε φορά μετατρέποντας την βέλτιστη λύση σε περιορισμό και επανάληψη για τα υπόλοιπα κριτήρια.

### 2.2.3.3 Αξιολόγηση της λειτουργίας Συστημάτων Συστάσεων

Οι τρεις βασικοί τρόποι που χρησιμοποιούνται για τη διαδικασία αξιολόγησης ενός συστήματος σύστασης είναι οι εξής: 1. Offline πειράματα (offline experiments), 2. Μελέτες χρηστών (user studies), 3. Online πειράματα με πραγματικούς χρήστες.

1. Τα Offline πειράματα είναι τυπικά πιο εύκολο να πραγματοποιηθούν, καθώς δεν απαιτούν αλληλεπίδραση με πραγματικούς χρήστες. Ένα offline πείραμα πραγματοποιείται με χρήση προεπιλεγμένων δεδομένων χρηστών που έχουν αξιολογήσει ή επιλέξει αντικείμενα. Με την χρήση αυτών των δεδομένων, μπορούμε να προσομοιώσουμε τη συμπεριφορά των χρηστών που αλληλεπιδρούν με το σύστημα σύστασης. Για να εφαρμόσουμε το παραπάνω υποθέτουμε ότι η συμπεριφορά του χρήστη τώρα παραμένει η ίδια με αυτή που ήταν όταν πραγματοποιήθηκε η συλλογή των δεδομένων, ώστε να προβαίνουμε σε αξιόπιστες αποφάσεις σύμφωνα με τη προσομοίωση.

Το σημαντικό προτέρημα των offline πειραμάτων είναι ότι δεν απαιτούν αλληλεπίδραση με πραγματικούς χρήστες και έτσι επιτρέπουν τη σύγκριση με ένα μεγάλο εύρος αλγορίθμων σε χαμηλό κόστος.

Το μειονέκτημα αυτών των πειραμάτων είναι ότι μπορούν μόνο να απαντήσουν σε ένα αρκετά μικρό σύνολο ερωτήσεων σχετικά με τον αλγόριθμο, που συνήθως έχουν να κάνουν με τις δυνατότητες πρόβλεψης.

Συνήθως τα offline πειράματα χρησιμοποιούνται προκειμένου να φιλτράρουν ένα μεγάλο αριθμό αλγοριθμικών προσεγγίσεων, δίνοντας ως αποτέλεσμα ένα μικρό σετ υποψήφιων αλγορίθμων για να εξεταστούν με μελέτες χρηστών ή online πειράματα.

Τα δεδομένα που χρησιμοποιούνται στην συγκεκριμένη περίπτωση πρέπει να προσεγγίζουν όσο το δυνατόν περισσότερο τα πραγματικά δεδομένα που θα χρησιμοποιηθούν, σε μια online λειτουργία του συστήματος.

Προκειμένου να προσομοιωθεί η συμπεριφορά του χρήστη συνήθως χρησιμοποιούνται ιστορικά δεδομένα χρήσης.

2. Στις μελέτες χρηστών, ζητάμε από ένα μικρό γκρουπ υποκειμένων να χρησιμοποιήσει το σύστημα σε ένα ελεγχόμενο περιβάλλον και στο τέλος να περιγράψουν την εμπειρία τους. Κατά τη διάρκεια χρήσης του συστήματος, παρατηρούμε και καταγράφουμε τη συμπεριφορά τους. Σε αυτού του είδους τα πειράματα μπορούμε να αντλήσουμε τόσο ποιοτική όσο και ποσοτική πληροφορία για τα συστήματα, αλλά απαιτείται προσοχή ώστε να ληφθούν υπόψη διάφοροι παράγοντες που είναι δυνατόν να επηρεάσουν το περιβάλλον του πειράματος.

Το πλεονέκτημα των μελετών χρηστών είναι ότι μπορούν να απαντήσουν σε ένα ευρύ σκετ ερωτήσεων, ότι μπορούμε να αντλήσουμε ποιοτικά δεδομένα που είναι πολύ συχνά χρήσιμο για την ανάλυση ποσοτικών δεδομένων καθώς και ότι είναι δυνατό να αντληθούν μεγάλα σκετ ποσοτικών δεδομένων με δεδομένο ότι οι χρήστες παρακολουθούνται πολύ στενά όταν τρέχουν το σύστημα.

Το κύριο μειονέκτημα των ερευνών χρηστών είναι ότι είναι πολύ κοστοβόρες συνήθως, καθώς ίσως και να απαιτείται και αποζημίωση των υποκειμένων που χρησιμοποιούνται στο πείραμα. Συνεπώς είναι δυνατό να τεθεί ζήτημα περιορισμού του μεγέθους του σκετ υποκειμένων. Επιπρόσθετα αρκετές φορές απαιτούνται επαναλαμβανόμενα πειράματα για να καταλήξουμε σε αξιόπιστα συμπεράσματα.

Κατά τις μελέτες χρηστών χρησιμοποιούνται σε μεγάλο βαθμό και ερωτηματολόγια. Οι ερωτήσεις μπορούν να παρέχουν πληροφορία για ιδιότητες ή χαρακτηριστικά του υποκειμένου που είναι δύσκολο να μετρηθούν.

3. Τα online πειράματα θεωρούνται τα πιο έγκυρα καθώς συμμετέχουν πραγματικοί χρήστες, όπου τυπικά είναι ανυποψίαστοι για το πείραμα. Σε τέτοιου είδους πειράματα είναι δυσκολότερο να συλλεχθούν συγκεκριμένα είδη δεδομένων, αλλά είναι τεράστιας σημασίας το γεγονός ότι αποτελούν πραγματικά δεδομένα. Επιπρόσθετα σε κάποιες

περιπτώσεις τέτοιου είδους πειράματα περιέχουν ρίσκο, καθώς στην περίπτωση που οι συστάσεις είναι άσχετες μπορεί να αποθαρρύνουν τους χρήστες από το να ξαναχρησιμοποιήσουν το σύστημα.

Για τον παραπάνω λόγο, είθισται να εφαρμόζεται ένα online σύστημα αξιολόγησης αφού έχει προηγηθεί ενδελεχής και μακροσκελής offline μελέτη, που αποδεικνύει ότι οι συστάσεις είναι τουλάχιστον λογικές.

Το πλεονέκτημα των online πειραμάτων είναι ότι μπορούν να μετράνε απευθείας σε πραγματικό χρόνο τους δείκτες-στόχους του συστήματος, όπως το κέρδος ή αφοσίωση χρηστών.

#### 2.2.3.4.1 Χαρακτηριστικά και μετρικές αξιολόγησης Συστήματος Συστάσεων

##### A. Προτίμηση/Αποδοχή Χρήστη

Ένας βασικός τρόπος αξιολόγησης ενός συστήματος είναι η προτίμηση/αποδοχή του χρήστη να χρησιμοποιήσει ένα συγκεκριμένο σύστημα συστάσεων. Αυτός ο τρόπος αξιολόγησης δεν υπόκειται σε περιορισμούς όσον αφορά τα υποκείμενα και τα χαρακτηριστικά τους και είναι γενικά ευκολότερος τρόπος αξιολόγησης από την βαθμολόγηση, ανάλογα με την εμπειρία χρήσης. Η αξιολόγηση του συστήματος προκύπτει με βάση το άθροισμα των θετικών ψήφων/ αποδοχών των χρηστών.

Ο τρόπος αυτός αξιολόγησης εμπεριέχει αρκετούς κινδύνους σφάλματος καθώς, δεν υπάρχει βαρύτητα σημαντικότητας στις αποδοχές των χρηστών, και επίσης δεν προκύπτει ποιοτική πληροφορία που θα οδηγήσει σε λήψη απόφασης.

##### B. Ακρίβεια πρόβλεψης

Η κοινή παραδοχή για ένα σύστημα συστάσεων είναι ότι όσο περισσότερη ακρίβεια πρόβλεψης έχει, τόσο πιο πιθανό είναι να προτιμηθεί από το χρήστη. Η ακρίβεια πρόβλεψης είναι τυπικά ανεξάρτητη από το περιβάλλον χρήστη, και

μπορεί να μετρηθεί με χρήση offline πειράματος. Η συγκεκριμένη μέτρηση έχει διαφορετική έννοια από την πρόβλεψη της συμπεριφοράς του χρήστη χωρίς συστάσεις και θεωρείται αρκετά ρεαλιστικός τρόπος αξιολόγησης λειτουργίας συστήματος.

Οι βασικές κατηγορίες μετρικών ακρίβειας πρόβλεψης είναι οι εξής: α. Μέτρηση της ακρίβειας προβλέψεων αξιολογήσεων, β. Μέτρηση της ακρίβειας προβλέψεων χρήσης (Shani and Gunawardana, 2011).

#### α. Μέτρηση της ακρίβειας προβλέψεων αξιολογήσεων

Οι μετρικές που χρησιμοποιούνται για την ακρίβεια προβλέψεων αξιολογήσεων επιτρέπουν τη σύγκριση των σφαλμάτων όταν ένας χρήστης αξιολογεί ένα αντικείμενο που του έχει συσταθεί.

Η πρώτη από αυτές τις μετρικές είναι το μέσο απόλυτο σφάλμα MAE (Mean Absolute Error)(Breese et al., 1998b), που υπολογίζει το μέσο της απόλυτης απόκλισης μεταξύ της προβλεπόμενης και της πραγματικής βαθμολογίας/αξιολόγησης από τους χρήστες του συστήματος.

$$\text{MAE} = \frac{\sum_{i=1}^N |q_i - r_i|}{N} \quad (2-6)$$

Όπου  $q_i$  είναι η προβλεπόμενη βαθμολογία,  $r_i$  είναι η πραγματική βαθμολογία και  $N$  είναι ο συνολικός αριθμός προβλέψεων.

Προκειμένου να επιλυθούν σχέσεις με μεγαλύτερες αποκλίσεις από τις πραγματικές βαθμολογίες χρησιμοποιείται το μέσο τετραγωνικό σφάλμα MSE (Mean Square Error)(Shardanand and Maes, 1995).

$$\text{MSE} = \frac{\sum_{i=1}^N (q_i - r_i)^2}{N} \quad (2-7)$$

Η πιο σημαντική μετρική που χρησιμοποιείται για την αξιολόγηση ακρίβειας προβλέψεων αξιολογήσεων είναι η ρίζα μέσω τετραγωνικού σφάλματος RMSE (Root Mean Square Error)(Bennett and Lanning, 2007).

$$\boxed{\text{RMSE} = \sqrt{\text{MSE}}} \quad (2-8)$$

Επίσης στην ερευνητική βιβλιογραφία εμφανίζονται οι κανονικοποιημένες μορφές των παραπάνω μετρικών Normalized RMSE (NRMSE) και Normalized MAE (NMAE), που είναι κανονικοποιημένες σε συγκεκριμένο εύρος βαθμολογιών ( $r_{\min}$ - $r_{\max}$ ). Αν θέλουμε ένα μέτρο χαρακτηριστικό του σφάλματος πρόβλεψης για κάθε αντικείμενο, τότε είναι προτιμότερο να υπολογίσουμε το MAE και RMSE ξεχωριστά για κάθε αντικείμενο και μετά να υπολογίσουμε το μέσο όρο όλων των αντικειμένων.

Τέλος αναφέρουμε τη μετρική ποσοστιαίου σφάλματος που δεν χρησιμοποιεί απόλυτες τιμές και για την περίπτωση της αξιολόγησης ενός συστήματος συστάσεων κρίνεται άκρως παραστατική.

$$\boxed{\text{PE} = \frac{e_i - t_i}{e_i} * 100} \quad (2-6)$$

Όπου  $e_i$  είναι η πειραματική βαθμολογία και  $t_i$  είναι η θεωρητική βαθμολογία.

### β. Μέτρηση της ακρίβειας προβλέψεων χρήσης

Σε πολλές εφαρμογές συστημάτων σύστασης δε γίνεται προσπάθεια πρόβλεψης των προτιμήσεων χρηστών αλλά συστήνονται στους χρήστες αντικείμενα που θα μπορούσαν να χρησιμοποιήσουν. Στην περίπτωση αυτή δεν ενδιαφερόμαστε για την ακριβή πρόβλεψη βαθμολογίας αλλά αν το σύστημα θα προβλέψει ότι ο χρήστης θα αξιολογήσει τα αντικείμενα σαν σχετικά ή μη σχετικά. Οι μετρικές που χρησιμοποιούνται σε αυτή την περίπτωση είναι: η ακρίβεια (Precision) και η ανάκληση (Recall) και δείχνουν την ποιότητα της μεθόδου σύστασης.

Το Precision είναι το κλάσμα των συστημένων αντικειμένων που είναι σχετικά.

$$\boxed{\text{Precision} = \frac{|\text{relevant items recommended}|}{|\text{Items in the list}|}} \quad (2-10)$$

Ο αριθμός των αντικειμένων που συστήνονται μπορεί να είναι πολύ υψηλός ανάλογα με τη μέθοδο σύστασης και το μέγεθος του σετ δεδομένων.

Το Recall είναι το κλάσμα των σχετικών συστάσεων που παρουσιάζονται στο χρήστη.

$$\text{Recall} = \frac{|\text{relevant items recommended}|}{|\text{relevant items}|} \quad (2-11)$$

Άλλες μετρικές που χρησιμοποιούνται συχνά είναι οι εξής:

#### Κάλυψη(Coverage)

Σε πολλές περιπτώσεις η ακρίβεια πρόβλεψης αυξάνει με την αύξηση του όγκου των δεδομένων. Αυτό έχει σαν αποτέλεσμα κάποιοι αλγόριθμοι σύστασης να παρέχουν πιο υψηλής ποιότητας συστάσεις αλλά μόνο για ένα μικρό σύνολο των αντικειμένων για τα οποία έχουν πολλά δεδομένα. Ο όρος κάλυψη αναφέρεται στο ποσοστό των αντικειμένων που το σύστημα συστάσεων μπορεί να προτείνει. Συναντάται συχνά με τον όρο catalog coverage. Ουσιαστικά προσδίδονται βαρύτητες στα αντικείμενα ανάλογα με τη δημοφιλία τους ή τη χρήση τους. Εν συνεχεία το σύστημα δεν μπορεί να προτείνει κάποια αντικείμενα που δεν χρησιμοποιούνται συχνά.

Η Κάλυψη επίσης συναντάται σαν το ποσοστό των χρηστών ή των αλληλεπιδράσεων χρηστών για το οποίο το σύστημα μπορεί να προβεί σε σύσταση. Σε αρκετές εφαρμογές το σύστημα δεν προβαίνει σε σύσταση σε χρήστες για τους οποίους δεν έχει αρκετά δεδομένα για το προφίλ τους.

#### Εμπιστοσύνη (Confidence)

Εμπιστοσύνη στη σύσταση προσδιορίζεται ως η εμπιστοσύνη του συστήματος στις συστάσεις ή προβλέψεις που μπορεί να προβεί. Στα περισσότερα συστήματα ανάλογα με τον όγκο των δεδομένων αυξάνει και η ακρίβεια σύστασης. Ταυτόχρονα με την ακρίβεια σύστασης αυξάνει και η αυτοπεποίθηση του συστήματος. Σε αρκετές περιπτώσεις ο χρήστης μπορεί να επωφεληθεί παρακολουθώντας τις τιμές εμπιστοσύνης. Αν το σύστημα εμφανίζει χαμηλή αυτοπεποίθηση ο χρήστης μπορεί να ψάξει περισσότερο την πιθανή επιλογή πριν αποφασίσει (McLaughlin and Herlocker, 2004).

### Καινοτομία (Novelty)

Καινοτόμες συστάσεις είναι συστάσεις για αντικείμενα που ο χρήστης δεν γνώριζε στο παρελθόν (Konstan and Riedl, 2012). Σε τέτοιου είδους εφαρμογές φιλτράρονται συνήθως τα αντικείμενα που έχουν προηγουμένως βαθμολογηθεί ή χρησιμοποιηθεί από το χρήστη.

### Ποικιλία (Diversity)

Η ποικιλία προσδιορίζεται σαν το αντίθετο της ομοιότητας. Σε κάποιες περιπτώσεις το να συστήνεις σε ένα χρήστη ένα σύνολο όμοιων αντικειμένων μπορεί να μην είναι χρήσιμο για το χρήστη καθώς απαιτείται περισσότερος χρόνος για να εξερευνήσει τις λεπτομέρειες των αντικειμένων. Συνήθως υπολογίζεται χρησιμοποιώντας την ομοιότητα μεταξύ αντικειμένων, που στηρίζεται στο περιεχόμενο του αντικειμένου. Εν συνεχεία η ποικιλία υπολογίζεται βασιζόμενη στο άθροισμα, τη μέγιστη ή ελάχιστη απόσταση μεταξύ των ζευγών αντικειμένων (Ziegler et al., 2005).

### Ευρωστία (Robustness)

Η Ευρωστία προσδιορίζεται ως η σταθερότητα της σύστασης όταν υπάρχει η παρουσία μια ψεύτικης πληροφορίας, που συνήθως εισάγεται επίτηδες για να επηρεάσει τις συστάσεις (O'Mahony and Smyth, 2009).

### Ιδιωτικότητα (Privacy)

Σε ένα σύστημα συστάσεων, ο χρήστης εισάγει τις προτιμήσεις του για προϊόντα στοχεύοντας σε χρήσιμες συστάσεις. Είναι σημαντικό για τους χρήστες όμως οι προτιμήσεις τους να παραμένουν μυστικές, ώστε να μην διαρρέει το προτιμησιακό τους προφίλ σε τρίτους (Lam et al., 2006).

### Επεκτασιμότητα (Scalability)

Η επεκτασιμότητα σε πραγματικά και μεγάλα δεδομένα είναι ένας από τους βασικούς παράγοντες επιτυχίας ενός συστήματος συστάσεων. Με την εκθετική αύξηση του μεγέθους των σετ δεδομένων πολλοί αλγόριθμοι επιβραδύνουν ή απαιτούν αυξημένους υπολογιστικούς πόρους. Βασικός παράγοντας για την



κατανόηση και την μέτρηση της επεκτασιμότητας ενός συστήματος συστάσεων είναι ο υπολογισμός της κατανάλωσης των πόρων του συστήματος σε μεγάλα σετ δεδομένων (Sarwar et al., 2001).

### 2.3 Μοντελοποίηση Προτίμησης

Η πολυκριτηριακή ανάλυση αποφάσεων περιλαμβάνει ένα σύνολο τεχνικών και μεθοδολογικών προσεγγίσεων για την εκτίμηση των βαρυτήτων των κριτηρίων που αντανακλούν τις προτιμήσεις του αποφασίζοντα. Οι βαρύτητες των κριτηρίων χρησιμοποιούνται για την εύρεση διακριτών εναλλακτικών από τις περισσότερες πολυκριτηριακές μεθόδους ανάλυσης αποφάσεων όπως οι ELECTRE (Roy, 1990), η Simos (José Figueira and Roy, 2002), (Simos, 1990a), (Simos, 1990b) και η WAP (Tsotsolas et al., 2016b).

#### 2.3.1 Καθορισμός βαρών μέσω ιεράρχησης προτεραιότητας (WAP)

Στο τομέα της υποστήριξης αποφάσεων, η γνώση των προτιμήσεων του αποφασίζοντα και ο καθορισμός των βαρυτήτων των κριτηρίων, αποτελούν άκρως σημαντικές και δύσκολες ερωτήσεις. Στη βιβλιογραφία συναντώνται αρκετές μέθοδοι που προσδίδουν μια κατάλληλη τιμή στις τιμές των κριτηρίων. Ο J. Simos πρότεινε μια αρκετά απλή διαδικασία, όπου με τη χρήση καρτών, επιτρέπει τον έμμεσο προσδιορισμό αριθμητικών τιμών στα βάρη. (Figueira and Roy, 2002) Στο παρόν κεφάλαιο αφού περιγράψουμε τις μεθόδους Simos και Revised Simos, θα αναλύσουμε τη μέθοδο WAP, που αποτελεί παραλλαγή της Simos και την οποία χρησιμοποιούμε στα πλαίσια της μεθοδολογίας για τη μοντελοποίηση της προτίμησης ενός αποφασίζοντα.

Με τη Simos (Simos, 1990a), ουσιαστικά δίνεται η δυνατότητα στον αποφασίζοντα να σκεφτεί και να εκφράσει τον τρόπο που θα επιθυμούσε να

ιεραρχηθούν τα διάφορα κριτήρια μια συγκεκριμένη ομάδας, στα πλαίσια ενός συγκεκριμένου προβλήματος. Η μέθοδος στοχεύει επίσης να μεταδώσει στον αναλυτή όλες τις πληροφορίες που απαιτούνται προκειμένου να οριστούν αριθμητικές τιμές στις βαρύτητες κάθε κριτηρίου (Figueira and Roy, 2002). Η συγκεκριμένη τεχνική έχει βρει εφαρμογή σε πολλές περιπτώσεις και έχει αποδειχθεί ότι γίνεται πολύ εύκολα αποδεκτή από το χρήστη (αποφασίζοντα), καθώς με τις πληροφορίες που παρέχει βοηθά το χρήστη να εκφράσει καλύτερα την προτίμηση του.

Η κύρια καινοτομία της Simos έγκειται στη συσχέτιση μίας «κάρτας» με κάθε κριτήριο. Το γεγονός ότι ο χρήστης πρέπει να χειριστεί τις κάρτες ώστε να τις κατατάξει, εισάγοντας ορισμένες άσπρες κάρτες, επιτρέπει τη βαθύτερη κατανόηση του σκοπού της διαδικασίας.

Η συγκέντρωση των απαραίτητων πληροφοριών γίνεται σε τρία στάδια σύμφωνα με τους Jose Figueira and Roy, (2002):

1. Δίνεται στο χρήστη ένα πακέτο με  $n$  κάρτες. Πάνω σε κάθε κάρτα γράφεται το όνομα κάθε κριτηρίου που ανήκει στο σύνολο κριτηρίων  $F$  μαζί με οποιαδήποτε συμπληρωματική πληροφορία που κρίνεται απαραίτητη. Ταυτόχρονα του δίνεται και ένα πακέτο με άσπρες κάρτες, ο αριθμός των οποίων εξαρτάται από τις ανάγκες του χρήστη.
2. Ο χρήστης καλείται να κατατάξει τις κάρτες με αύξουσα σειρά από τη λιγότερο σημαντική στην πιο σημαντική, ανάλογα της σημαντικότητας που θέλει να προσδώσει στο κριτήριο. Το πρώτο κριτήριο στην κατάταξη είναι το λιγότερο σημαντικό και το τελευταίο το πιο σημαντικό. Στην περίπτωση που κάποια κριτήρια είναι εξίσου σημαντικά για το χρήστη, θα πρέπει να οριστεί ένα υποσύνολο καρτών.
3. Ο χρήστης καλείται να σκεφτεί το γεγονός ότι η σημαντικότητα των διαδοχικών κριτηρίων μπορεί να είναι ίδια. Στον καθορισμό των βαρών πρέπει να ληφθεί υπόψη αυτή η ελάχιστη διαφορά. Για αυτό το λόγο ο χρήστης καλείται να εισάγει τόσες περισσότερες άσπρες κάρτες μεταξύ των δύο διαδοχικών καρτών, όσο μεγαλύτερη θεωρεί τη διαφορά της σημαντικότητας μεταξύ των κριτηρίων.

Ο υπολογισμός των βαρών, με άθροισμα 100, ακολουθεί την εξής διαδικασία:

- Αριθμός καρτών: Για κάθε κλάση που δημιουργείται (συμπεριλαμβάνονται και οι κλάσεις των λευκών καρτών) καταγράφονται οι κάρτες που την αποτελούν και υπολογίζεται το άθροισμά τους.
- Θέσεις: Για κάθε κάρτα γίνεται η αρίθμηση – θέση 1,2,3,... αρχίζοντας από την ουρά της κατάταξης μέχρι την κεφαλή. Η θέση της τελευταίας κάρτας είναι προφανώς ο συνολικός αριθμός των καρτών.
- Μη κανονικοποιημένα βάρη: Ως βάρος της κάθε κλάσης υπολογίζεται το άθροισμα των θέσεων της κλάσης δια του αριθμού των καρτών της κλάσης.
- Κανονικοποιημένα βάρη – στρογγυλοποίηση: Τα μη κανονικοποιημένα βάρη διαιρούνται δια του αθροίσματος των θέσεων, όπου δεν συνυπολογίζονται οι θέσεις των λευκών καρτών και πολλαπλασιάζονται επί 100. Εν συνεχεία, τα βάρη αυτά στρογγυλεύονται στον πλησιέστερο ακέραιο.

Οι Maystre et al., (1994) ανέλυσαν με ένα παράδειγμα (βλέπε παρακάτω Πίνακα 2-1) τον τρόπο που πρότεινε ο Simos για την επεξεργασία των συγκεντρωμένων πληροφοριών.

Τα παραπάνω βήματα της μεθόδου υλοποιούνται στήλη – στήλη στον πίνακα, όπου ένας αποφασίζων έχει κατατάξει δώδεκα κριτήρια  $F = \{c_1, c_2, \dots, c_{12}\}$  σε 7 ισόβαρες κλάσεις (μια από τις κλάσεις αυτές κατέχεται από μια λευκή κάρτα). Στην κατάταξη αυτή, η κλάση κριτηρίων  $\{c_3, c_7, c_{12}\}$  είναι η λιγότερο σημαντική (ουρά) ενώ η κλάση  $\{c_{11}\}$  η περισσότερη σημαντική (κεφαλή κατάταξης).

Πίνακας 2-1:Υπολογισμός βαρών 12 κριτηρίων με τη μέθοδο Simos

Κλάση	Αριθμός καρτών	Θέσεις	Μη κανονικοποιημένο βάρος	Κανονικοποιημένο βάρος	Σύνολο
$\{c_3, c_7, c_{12}\}$	3	1,2,3	$\frac{1+2+3}{3} = 2$	$\frac{2}{86} \times 100 = 2,3 \rightarrow 2$	$3 \times 2 = 6$

$\{c_4\}$	1	4	4	$\frac{4}{86} \times 100 = 4,7 \rightarrow 5$	$1 \times 5 = 5$
Λευκή	1	(5)	-	-	-
$\{c_2, c_6, c_9, c_{10}\}$	4	6,7,8,9	$\frac{6+7+8+9}{4} = 7,5$	$\frac{7,5}{86} \times 100 = 8,7 \rightarrow 9$	$4 \times 9 = 36$
$\{c_5\}$	1	10	10	$\frac{10}{86} \times 100 = 11,6 \rightarrow 12$	$1 \times 12 = 12$
$\{c_1, c_8\}$	2	11,12	$\frac{11+12}{2} = 11,5$	$\frac{11,5}{86} \times 100 = 13,4 \rightarrow 13$	$2 \times 13 = 26$
$\{c_{11}\}$	1	13	13	$\frac{13}{86} \times 100 = 15,1 \rightarrow 15$	$1 \times 15 = 15$
Σύνολο	13	86	-	-	100

Παρατηρήσεις:

1. Η κανονικοποίηση των βαρών στη μέθοδο ακολουθείται, από στρογγυλοποίηση στον πλησιέστερο ακέραιο, χωρίς αυτό να προσφέρει κάτι χρήσιμο στις μεθόδους υπεροχής. Τούτο μπορεί εύκολα να οδηγήσει σε άθροισμα διαφορετικό του 100.

Βασιζόταν σε μια μη ρεαλιστική υπόθεση, με αποτέλεσμα να οδηγεί πολλές φορές σε σαφώς ορισμένα αποτελέσματα (Schärlig, 1996). Το προηγούμενο πρόβλημα καθώς και κάποια άλλα πιο σημαντικά «διορθώθηκαν» σε μια αναθεωρημένη έκδοση της μεθόδου που δημιουργήσαν οι Figueira & Roy (2002).

Η κατάταξη των κριτηρίων, που χρησιμοποιεί η μέθοδος ως αρχική πληροφορία, πρέπει μάλλον να εκληφθεί από τους αναλυτές ως ένα από τα ζητούμενα των μεθόδων αυτών παρά ως παράδειγμα. Η νέα αναθεωρημένη μέθοδος του Simos διαφέρει από την αρχική στα ακόλουθα τρία σημεία:

- i. Συλλέγει νέο είδος πληροφοριών.
- ii. Επεξεργάζεται τις πληροφορίες, προκειμένου να αποδώσει κανονικοποιημένες βαρύτητες στα κριτήρια.
- iii. Χρησιμοποιεί μια διαφορετική τεχνική με την οποία κανονικοποιεί τις βαρύτητες, ελαχιστοποιώντας λάθη της στρογγυλοποίησης. Σκοπός της

αναθεωρημένη μεθόδου είναι αποδώσει την πραγματική βαρύτητα σε κάθε κριτήριο, χωρίς να χρειάζεται να γνωρίζει ο χρήστης την κλίμακα κατάταξης που ορίζεται κάθε φορά, προκειμένου να εκφράσει την προτίμηση του στην κατάταξη. Άλλωστε σε πραγματικές εφαρμογές, ο αποφασίζων θέλει απλώς να εκφράσει την επιθυμία του ως προς την κατάταξη των κριτηρίων, χωρίς να απαιτείται να γνωρίζει την κλίμακα που μπορεί να έχει οριστεί, ούτε τη διαδικασία 'αποκωδικοποίησης' αυτής.

Μια νέα πληροφορία που δίνεται στην αναθεωρημένη Simos σε σχέση με την απλή, είναι ένα νέο είδος πληροφορίας που ζητείται από τον χρήστη. Του ζητείται να ορίσει πόσες φορές πιο σημαντικό είναι το τελευταίο κριτήριο της κατάταξης σε σχέση με το πρώτο.

Η WAP (Weights Assessment through Prioritization) παρέχει ένα εύκολο και ευέλικτο τρόπο για τη μοντελοποίηση της προτίμησης ενός αποφασίζοντα και τον υπολογισμό των βαρών για το σύστημα αξιών, τα οποία αντικατοπτρίζουν το όραμα, τη στρατηγική και τις προτιμήσεις του αποφασίζοντα.

Ο υπολογισμός των κριτηρίων με τη συγκεκριμένη μέθοδο βασίζεται σε δύο προτιμησιακές πληροφορίες που παρέχονται από τον αποφασίζοντα. Η πρώτη είναι η προτεραιότητα των κριτηρίων σε τάξεις σημασίας από το σημαντικότερο, στο λιγότερο σημαντικό. Η δεύτερη έχει να κάνει με τη σχετική σημαντικότητα των κριτηρίων μεταξύ τους, σύμφωνα πάντα με τον αποφασίζοντα. Για αυτό γίνεται χρήση των δεικτών  $z$  (το  $z = w_k / w_r$  προκύπτει από το πηλίκο του βάρους του πιο σημαντικού κριτηρίου  $w_k$  με το βάρος του λιγότερο σημαντικού κριτηρίου  $w_r$ ) για κάθε ζεύγος διαδοχικών κριτηρίων, που έχουν προηγουμένως ταξινομηθεί με κάποιου είδους κατάταξη από τον αποφασίζοντα.

Σημαντική παράμετρος κατά τη χρήση της WAP είναι η δυνατότητα εξάλειψης των white cards που χρησιμοποιούνται από τη μέθοδο SIMOS και των λοιπών παραλλαγών της. Επιπρόσθετα στην WAP δεν ζητείται από τον αποφασίζοντα να δώσει συγκεκριμένες τιμές, αλλά ένα εύρος τιμών  $[z_{\min_r}, z_{\max_r}]$ , όπου η τιμή του  $z_r$  μπορεί να ποικίλει.

Για δύο διαδοχικά κριτήρια  $c_r, c_{r+1}$  το εύρος τιμών  $[z_{\min_r}, z_{\max_r}]$  είναι τέτοιο ώστε να ισχύει  $z_{\min_r} \leq z_r \leq z_{\max_r}$  και  $c_r = z_r^* c_{r+1}$ . Στη συγκεκριμένη προσέγγιση

ο αποφασίζοντας ρωτάται να κατατάξει τα  $n$  κριτήρια σε  $m$  κλάσεις ( $m \leq n$ ) και να προσδιορίσει  $z(m+1)$  τιμές για τα  $z_{\min_r}$  και  $z_{\max_r}$  όπου ( $r=1,2,\dots,m-1$ ). Το τελικό  $z$  μπορεί να υπολογιστεί από το γινόμενο των  $z_r$  δεικτών:

$$\text{Δηλαδή: } z = z_1 z_2 \dots z_{m-1} = (c_1/c_2)(c_2/c_3) \dots (c_{m-1}/c_m) = c_1/c_m \quad (2-12)$$

Επίσης για τα ελάχιστα και μέγιστα  $z$  ισχύει:

$$\left[ \prod_{i=1}^{m-1} z_{\min_i} \leq z \leq \prod_{i=1}^{m-1} z_{\max_i} \right] \quad (2-13)$$

Η προσέγγιση της WAP περιλαμβάνει τα εξής βήματα:

1. Κατάταξη των κριτηρίων από τον αποφασίζοντα.
2. Τα κριτήρια ταξινομούνται σύμφωνα με την κατάταξη του χρήστη, από το σημαντικότερο στο λιγότερο σημαντικό σε κλάσεις που το πλήθος τους είναι δυνατόν να είναι ίσο με τον αριθμό των κριτηρίων.
3. Καλείται ο αποφασίζοντας να εκφράσει τα όρια των ευρών της σχετικής σημαντικότητας μεταξύ των διαδοχικών κριτηρίων για όλα τα ζεύγη αυτών, προσδιορίζοντας τη διαφορά σημαντικότητας επιρροής του καθενός από τα διαδοχικά κριτήρια, στην τελική του απόφαση. Στη φάση αυτή υπολογίζονται τα  $[z_{\min_r}, z_{\max_r}]$ .
4. Κατόπιν κατασκευάζεται και επιλύεται το παρακάτω γραμμικό πρόβλημα:

$$\text{Min } c_i \text{ \& Max } c_b \text{ for } i = 1, 2, \dots, n \quad (2-14)$$

$c_i - c_{i+1} = 0$ , αν  $g_{i+1}$  ακολουθείται από το  $g_b$  και το  $g_{i+1}$  ανήκει στην ίδια τάξη σημαντικότητας ( $j$ ) όπως  $g_b$  (2-15)

$c_i - c_{i+1} \geq z_{\min_j}$ ,  $c_i - c_{i+1} \leq z_{\max_j}$ , if  $g_r$  ακολουθείται από το  $g_{i+b}$ ,  $g_i$  ανήκει στην  $j$  και το  $g_{i+1}$  ανήκει στη τάξη  $j+1$  (2-16)

$$c_1 + c_2 + \dots + c_n = 1 \quad (2-17)$$

$$c_i \geq 0, c_2 \geq 0, \dots, p_n \geq 0 \quad (2-18)$$

$$w_1 \geq 0, w_2 \geq 0, \dots, w_k \geq 0 \quad (2-7)$$

Οι λύσεις από το παραπάνω γραμμικό πρόβλημα οδηγούν στον εντοπισμό των ελάχιστων και μέγιστων τιμών των βαρυτήτων των κριτηρίων, που είναι

φραγμένα από το m-διάστατο πολύεδρο των διανυσμάτων των βαρυτήτων που ικανοποιούν τις προτιμήσεις του αποφασίζοντα.

5. Υπολογίζεται η ευρωστία (robustness) της εκτιμώμενης λύσης χρησιμοποιώντας δύο δείκτες. Ο πρώτος τύπος δεικτών είναι το εύρος μεταξύ των μέγιστων και ελάχιστων τιμών των βαρών των κριτηρίων για κάθε κριτήριο. Αυτός δείχνει το δείκτη ευρωστίας για κάθε κριτήριο. Για το i-th κριτήριο ο δείκτης υπολογίζεται ως εξής:

$$\mu = (\max(w_{ij}) - \min(w_{ij})) (2-20)$$

όπου  $w_{ij}$  είναι το βάρος το i κριτηρίου για το j,  $i=1,2,\dots,n$ ,  $j=1,2,\dots,m$ , όπου n είναι ο αριθμός των κριτηρίων και m ο αριθμός των αξόνων του υπερπολύεδρου.

Ο δεύτερος δείκτης αποτελεί την κανονικοποιημένη τυπική απόκλιση των διαφορετικών λύσεων, με την τιμή 1 να αντιστοιχεί σε πλήρη ευρωστία και το 0 σε μη-ευρωστία των προτιμησιακών μοντέλων (Hurson and Siskos 2014). Αυτός ο δείκτης ονομάζεται Μέσος δείκτης Ευστάθειας (Average Stability Index (ASI)) και υπολογίζεται ως εξής:

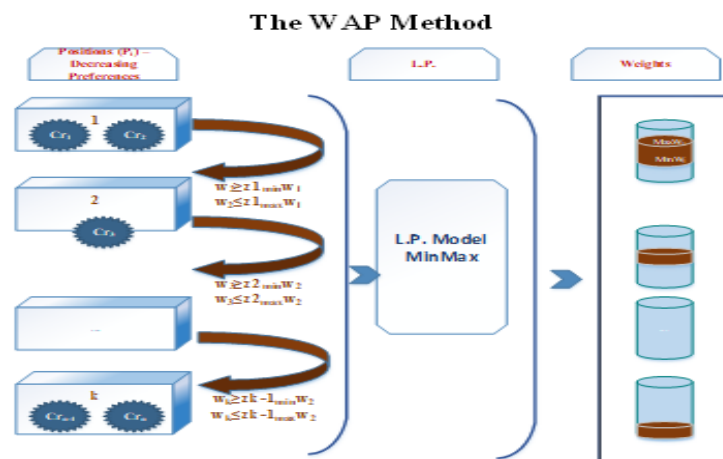
$$s = 1 - \frac{1}{n} \sum_{i=1}^m \sqrt{\frac{m \sum_{j=1}^n p_{ij} - \left( \sum_{j=1}^n p_{ij} \right)^2}{\frac{m\sqrt{n-1}}{n}}} \quad (2-8)$$

Όπου n, ο αριθμός των κριτηρίων, m ο αριθμός των αξόνων του υπερπολύεδρου.

Τέλος υπολογίζεται το βαρύκεντρο που αποτελεί τη μέση τιμή των βαρυτήτων για κάθε κριτήριο. Αν οι τιμές θεωρούνται ικανοποιητικές από τον αποφασίζοντα σε συνάρτηση με το επίπεδο ευστάθειας, τότε είναι δυνατό να χρησιμοποιηθεί σαν διάνυσμα βαρυτήτων για τα επόμενα βήματα. Σε αντίθετη περίπτωση, υπάρχουν δύο εναλλακτικές ενέργειες. Κατά την πρώτη, τροποποιούνται οι αρχικές προτιμήσεις, που σχετίζεται πρακτικά με την αλλαγή των τιμών  $z_{\min}$  και  $z_{\max}$ , ή με την αλλαγή στην κατάταξη των κριτηρίων. Η δεύτερη ενέργεια αφορά τις ανατροφοδοτήσεις με τις οποίες μπορεί να γίνει

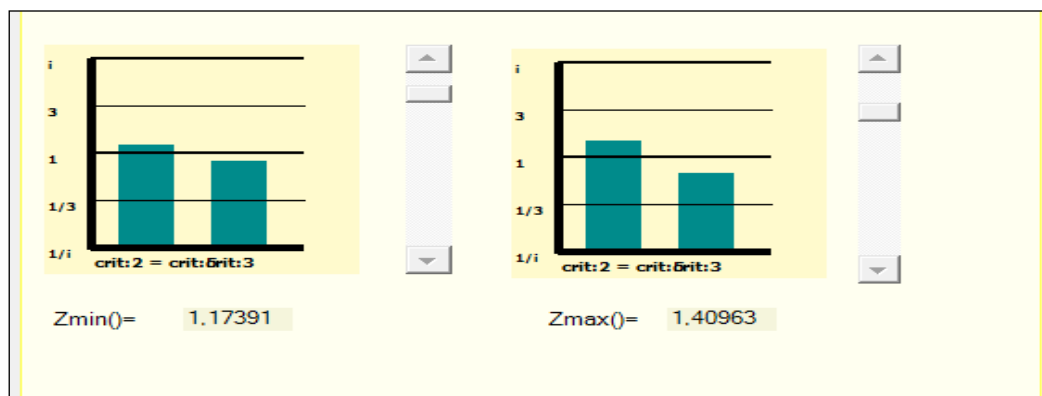
ανάλυση της ευρωστίας χρησιμοποιώντας τη τομογραφική προσέγγιση (Spyridakos et al., 2015), που απαιτεί πρόσθετες προτιμησιακές πληροφορίες από τη μεριά του αποφασίζοντα, ώστε να καταλήξουμε σε υπερ-πολύεδρο με υψηλότερα επίπεδα ευρωστίας.

Παρακάτω παρατίθεται σχηματικό διάγραμμα (Σχ.2.2) της λειτουργίας της Μεθόδου WAP που περιλαμβάνει όλα τα βήματα που περιγράφηκαν αναλυτικά παραπάνω.



Σχ. 2-2 Σχηματικό Διάγραμμα της λειτουργίας της Μεθόδου WAP

Επίσης για την απλοποίηση της διαδικασίας για τον αποφασίζοντα έχει αναπτυχθεί λογισμικό, όπου με χρήση ράβδου κύλισης, ο χρήστης μπορεί να προσδιορίσει τη διαφορά σημαντικότητας μεταξύ των καταταγμένων κριτηρίων δηλαδή τα  $Z_{min}$ ,  $Z_{max}$ .



Σχ. 2-3 Ράβδοι κύλισης για προσδιορισμό της διαφοράς σημαντικότητας κριτηρίων με  $Z_{min}$ ,  $Z_{max}$



## 2.4 Ικανοποίηση Πελατών

Η ικανοποίηση πελάτη αποτελεί άκρως σημαντική πληροφορία για μια εταιρεία προκειμένου να αξιολογήσει τη θέση της απέναντι στον ανταγωνισμό, να σχεδιάσει της στρατηγική της για το μέλλον, να βρει ευκαιρίες στην αγορά και να εξερευνήσει λεπτομερειακά τις επιδιώξεις, τις ανάγκες και επιθυμίες των πελατών. Σύμφωνα με τους Engler et al., (2015), η ικανοποίηση πελατών εμπεριέχει τις προ-αγοραστικές επιδιώξεις σε σχέση με την πραγματική απόδοση τους προϊόντος, και κυριολεκτικά καθορίζει τις online βαθμολογίες. Ακριβώς το ίδιο επισημαίνουν και οι King et al., (1993) για την περίπτωση της ικανοποίησης τουριστών.

Η ικανοποίηση καθορίζεται από υποκειμενικούς (π.χ. ανάγκες πελατών, συναισθήματα) και αντικειμενικούς παράγοντες (π.χ. χαρακτηριστικά προϊόντος και υπηρεσίας).

Παρακάτω παρουσιάζονται συνοπτικά εναλλακτικές μεθοδολογικές προσεγγίσεις στο πρόβλημα μέτρησης και ανάλυσης της ικανοποίησης πελατών, με σημείο αναφοράς τις ποσοτικές μεθόδους ανάλυσης δεδομένων. Οι βασικές ποσοτικές μέθοδοι ανάλυσης δεδομένων είναι οι εξής:

Περιγραφική στατιστική: Με βάση τη συγκεκριμένη μέθοδο ουσιαστικά υπολογίζονται οι συχνότητες των απαντήσεων των πελατών σε συγκεκριμένες ερωτήσεις που θεωρούνται «κρίσιμες» και ανάλογα με την κλίμακα που χρησιμοποιείται, υπολογίζονται τα ποσοστά των ικανοποιημένων και των δυσαρεστημένων πελατών, τα οποία αποτελούν ένα δείκτη απόδοσης της επιχείρησης.

Ο χαρακτηρισμός και η επιλογή των βαθμίδων ικανοποίησης που αντιπροσωπεύουν ικανοποιημένους ή δυσαρεστημένους πελάτες εξαρτάται από την πολιτική της εταιρείας και τον ανταγωνισμό. Όταν υπάρχουν διαθέσιμα δεδομένα από διαφορετικές χρονικές περιόδους τα αποτελέσματα μπορούν να χρησιμοποιηθούν στις ακόλουθες αναλύσεις τάσεων: α. απεικόνιση των ποσοστών σε μια χρονολογική σειρά, β. με στόχο των προσδιορισμό των διαστάσεων ικανοποίησης στα οποία έχει επιτευχθεί βελτίωση και σε αυτά που

πρέπει να δοθεί πρόσθετη προσοχή, γ. εφαρμογή στατιστικού ελέγχου υποθέσεων για να εξεταστούν πιθανές αλλαγές στη στάση των πελατών σε σχέση με προηγούμενες χρονικές περιόδους και απεικόνιση των δεδομένων σε γραφήματα στατιστικού ελέγχου ποιότητας με προκαθορισμένα όρια διακύμανσης.

Βασικές στατιστικές μέθοδοι: Μια από τις πρώτες προσπάθειες εφαρμογής στατιστικών μεθόδων είναι η πολλαπλή ανάλυση παλινδρόμησης (multiple regression analysis). Η μέθοδος αυτή χρησιμοποιείται για τον προσδιορισμό της σχέσης ανάμεσα στο σύνολο της απόδοσης χαρακτηριστικών (ανεξάρτητες μεταβλητές) και στη συνολική αποτίμηση αξίας της ικανοποίησης των πελατών (εξαρτημένες μεταβλητές).

Μια άλλη στατιστική μέθοδος που χρησιμοποιείται ευρέως είναι η παραγοντική ανάλυση (factor analysis). Βασικός σκοπός της είναι η μελέτη του τρόπου συσχέτισης ενός συνόλου χαρακτηριστικών του προσφερόμενου προϊόντος ή υπηρεσίας.

Probit και Logit ανάλυση που αποτελούν εξειδικευμένα στατιστικά μοντέλα κατηγορικών μεταβλητών: Λογιστικό μοντέλο πιθανότητας (Logit analysis) και κανονικό μοντέλο πιθανότητας (Probit analysis).

Τα μοντέλα αυτά με βάση την αξιολόγηση του πελάτη για ένα σύνολο χαρακτηριστικών του προϊόντος, εκτιμούν μια συνάρτηση κατανομής, η οποία εκφράζει την πιθανότητα ο πελάτης αυτός να ανήκει σε κάποια συγκεκριμένη ομάδα ικανοποίησης (Feinberg, 1977, Agresti and Natarajan, 2001).

Η διακριτική ανάλυση (Discriminant analysis): Μια από τις πιο ευρέως χρησιμοποιούμενες πολυμεταβλητές μεθόδους, η οποία, δοθείσης της ικανοποίησης ή της άποψης του πελάτη για το σύνολο των χαρακτηριστικών του προϊόντος, εκτιμά αν ο συγκεκριμένος πελάτης ανήκει σε κάποια καθορισμένη ομάδα ικανοποίησης.

Ανάλυση συζυγιών (conjoint analysis): χρησιμοποιείται για να εκτιμήσει την επίδραση των παραχωρήσεων (trade-offs) του πελάτη, όταν αγοράζει ή αξιολογεί την ικανοποίησή του, σε προσφερόμενα προϊόντα ή υπηρεσίες. Η συγκεκριμένη ανάλυση αποτελεί την λογική επέκταση των ερευνών ικανοποίησης, δεδομένου ότι οι σημαντικότερες αποφάσεις παραχώρησης του πελάτη περικλείουν τις κρίσιμες διαστάσεις απόδοσης του προϊόντος που έχουν προσδιοριστεί από την έρευνα ικανοποίησης: (Green and Srinivasan, 1978, Anderson and Bettencourt, 1993).

Ανάλυση αντιστοιχιών (correspondence analysis): Ένα από τα σημαντικότερα μαθηματικά εργαλεία για την κατασκευή αντιληπτικών χαρτών (perceptual maps) σε εφαρμογές marketing.

Τέλος, άλλα στατιστικά μοντέλα και ποσοτικές μέθοδοι ανάλυσης ικανοποίησης πελατών, με μικρότερο αριθμό πρακτικών εφαρμογών είναι τα εξής:

1. Περιβάλλουσα ανάλυση δεδομένων (DEA – Data Envelopment Analysis)
2. Πολυδιάστατη ανάλυση αναλογιών (multidimensional scaling)
3. Ανάλυση ομαδοποίησης (cluster analysis)
4. Δομικά μοντέλα εξισώσεων (structural equation models)
5. Γραφικά μοντέλα πιθανότητας (probability plotting methods)

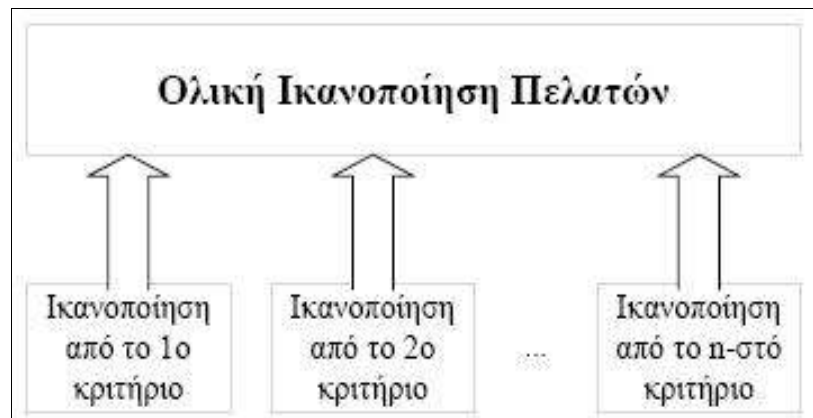
Η μεθοδολογία MUSA, η οποία χρησιμοποιείται στην παρούσα εργασία, στοχεύει στην χρήση προχωρημένων τεχνικών για την διαχείριση ποιοτικών δεδομένων, καθώς έχει τη δυνατότητα να εκτιμήσει ένα σύστημα αξιών για οποιοδήποτε σύνολο δεδομένων ικανοποίησης, το οποίο σέβεται την ποιοτική μορφή της συγκεκριμένης πληροφορίας.

#### **2.4.1 Πολυκριτήρια Ανάλυση Ικανοποίησης (MUSA)**

Το μοντέλο MUSA (MULTicriteria Satisfaction Analysis) στηρίζεται στην σύνθεση των διαφορετικών προτιμήσεων ενός συνόλου πελατών σε μια συνολική συνάρτηση αξιών, χρησιμοποιώντας προχωρημένες τεχνικές

γραμμικού προγραμματισμού (Siskos, 1985), ποιοτικής ανάλυσης παλινδρόμησης και μονότονες μεταβλητές. Το συγκεκριμένο μοντέλο ανήκει στην ευρύτερη κατηγορία των «Αναλυτικών – Συνθετικών Μοντέλων Προτίμησης (Preference Aggregation- Disaggregation Models)». Το μοντέλο εκτίμησης της ικανοποίησης πελατών προσπαθεί να αναλύσει την πολυκριτήρια συμπεριφορά ενός συνόλου ατόμων, με δεδομένες τις προτιμήσεις των τελευταίων.

Πιο συγκεκριμένα, η μέθοδος υποθέτει ότι η συνολική ικανοποίηση ενός μεμονωμένου πελάτη εξαρτάται από ένα σύνολο μεταβλητών, τα οποία αντιπροσωπεύουν τα χαρακτηριστικά του προσφερόμενου προϊόντος ή υπηρεσίας.



Σχ. 2-4 Ολική ικανοποίηση

1. Ορθολογικός καταναλωτής: Η συγκεκριμένη υπόθεση σχετίζεται με την ύπαρξη ορθολογικών (rational) πελατών και συναντάται στο σύνολο του χώρου της επιστήμης των αποφάσεων.
2. Κριτήρια ικανοποίησης: Η μέθοδος MUSA υποθέτει την ύπαρξη ενός συνόλου χαρακτηριστικών του εξεταζόμενου προϊόντος ή υπηρεσίας, σύμφωνα με τα οποία οι πελάτες αντιλαμβάνονται την ικανοποίησή τους. Το σύνολο αυτό των χαρακτηριστικών αποτελεί τα κριτήρια ικανοποίησης των πελατών.
3. Προσθετικό μοντέλο σύνθεσης: Τέλος, γίνεται η παραδοχή ύπαρξης ενός προσθετικού μοντέλου σύνθεσης του συνόλου των κριτηρίων ικανοποίησης. Στα

πλαίσια της πολυκριτήριας ανάλυσης αποφάσεων, οι συναρτήσεις αυτές πληρούν την ιδιότητα της μονοτονίας.

Τα βασικά στάδια της μεθοδολογίας υλοποίησης ερευνών ικανοποίησης είναι τα εξής:

1. Προκαταρκτική έρευνα: Το αρχικό αυτό στάδιο αφορά στο σαφή καθορισμό των στόχων της έρευνας, οι οποίοι αποτελούν τη βασική προϋπόθεση για το σχεδιασμό και την ανάπτυξη ενός προγράμματος μέτρησης ικανοποίησης. Ταυτόχρονα, το στάδιο αυτό περιλαμβάνει τόσο την προκαταρκτική ανάλυση της συμπεριφοράς των πελατών, όσο και την έρευνα του περιβάλλοντος αγοράς, με στόχο τη συλλογή της απαιτούμενης πληροφορίας για την ανάπτυξη ενός συστήματος μέτρησης της ικανοποίησης πελατών.
2. Διεξαγωγή έρευνας ικανοποίησης: Η φάση αυτή περιλαμβάνει τη σύνταξη του ερωτηματολογίου και την πραγματοποίηση της έρευνας σε ένα αντιπροσωπευτικό δείγμα πελατών της εταιρείας.
3. Αναλύσεις: Το σύνολο των αναλύσεων και της επεξεργασίας των δεδομένων της έρευνας συμπεριλαμβάνονται στο στάδιο αυτό. Οι αναλύσεις αφορούν στην εφαρμογή τόσο μεθόδων περιγραφικής στατιστικής, όσο και της πολυκριτήριας μεθόδου μέτρησης της ικανοποίησης MUSA.
4. Αποτελέσματα: Στο στάδιο αυτό εκτιμείται η εγκυρότητα των αποτελεσμάτων της έρευνας και μορφοποιούνται οι τελικές προτάσεις για τη βελτίωση ή/και διατήρηση του επιπέδου ικανοποίησης των πελατών, συνοψίζοντας τα σημαντικότερα συμπεράσματα των προηγούμενων αναλύσεων.

Η μέθοδος MUSA προσπαθεί να εκτιμήσει μια συλλογική συνάρτηση αξιών (additive collective value function)  $Y^*$  και ένα σύνολο μερικών συναρτήσεων ικανοποίησης  $X_i^*$  με βάση τις γνώμες των πελατών, έχοντας σαν αντικειμενικό σκοπό την επίτευξη της μεγαλύτερης δυνατής συμφωνίας ανάμεσα στη συνάρτηση  $Y^*$  και στις απόψεις των πελατών  $Y$ . Η συνάρτηση ανάλυσης παλινδρόμησης έχει την εξής μορφή:

$$\begin{cases} Y^* = \sum_{i=1}^n b_i X_i^* \\ \sum_{i=1}^n b_i = 1 \end{cases} \quad (2-9)$$

όπου οι συναρτήσεις αξιών  $Y^*$  και  $X_i^*$  είναι κανονικοποιημένες στο διάστημα  $[0, 100]$ ,  $n$  είναι ο αριθμός των κριτηρίων και  $b_i$  είναι το βάρος του  $i$  κριτηρίου.

Ενσωματώνοντας και μια διπλή μεταβλητή σφάλματος η παραπάνω εξίσωση παίρνει την ακόλουθη μορφή:

$$\tilde{Y}^* = \sum_{i=1}^n b_i X_i^* - \sigma^+ + \sigma^- \quad (2-23)$$

Όπου  $\tilde{Y}^*$  είναι η εκτίμηση τη συνολικής συνάρτησης αξιών  $Y^*$ , και  $\sigma^-$  και  $\sigma^+$  τα σφάλματα υποεκτίμησης και υπερεκτίμησης, αντίστοιχα.

Όπως προαναφέρθηκε οι συναρτήσεις αξιών  $Y^*$  και  $X_i^*$  είναι μονότονες συναρτήσεις κανονικοποιημένες στο διάστημα  $[0, 100]$ . Προκειμένου να μειωθεί το μέγεθος το μαθηματικού προβλήματος, με στόχο τη μείωση της υπολογιστικής δυσκολίας εύρεσης της βέλτιστης λύσης αφαιρούνται οι περιορισμοί μονοτονίας με χρήση των παρακάτω εξισώσεων μετασχηματισμού:

$$\begin{cases} Z_m = y^{*m+1} - y^{*m}, \text{για } m = 1, 2, \dots, a-1 \\ w_{ik} = b_i x_i^{*k+1} - b_i x_i^{*k} \text{για } k = 1, 2, \dots, a-1 \text{ και } i = 1, 2, \dots, n \end{cases} \quad (2-24)$$

Όπου  $y^{*m}$  είναι η τιμή του  $y^*$  επιπέδου ικανοποίησης,  $x_i^{*k}$  είναι η τιμή του  $x_i^*$  επιπέδου ικανοποίησης και  $a, \alpha_i$  είναι ο αριθμός των επιπέδων συνολικής και μερικής ικανοποίησης.

Χρησιμοποιώντας την εξίσωση (2.24), οι αρχικές μεταβλητές απόφασης του γραμμικού προγράμματος γράφονται:

$$\begin{cases} y^{*m} = \sum_{i=1}^{m-1} z_i \text{για } m = 2, 3, \dots, a \\ b_i x_i = \sum_{k=1}^{m-1} w_{ik} \text{για } k = 2, 3, \dots, a \text{ και } i = 1, 2, \dots, n \end{cases} \quad (2-25)$$

Οπότε εισάγοντας τις νέες μεταβλητές  $z_m$  και  $w_{ik}$  και χρησιμοποιώντας τις εξισώσεις (2.24), (2.25), η εξίσωση παλινδρόμησης γίνεται:

$$\sum_m z_m = \sum_i \sum_k w_{ik} - \sigma^+ + \sigma^- \quad (2-26)$$

Πιο συγκεκριμένα έστω ότι ο πελάτης  $j$  έχει εκφράσει την ικανοποίησή του με βάση τις καθορισμένες ποιοτικές κλίμακες  $Y$  και  $X_i$ .

Τότε για την εξίσωση (2.26) θα ισχύει:

$$\sum_{m=1}^{t_j-1} z_m = \sum_{i=1}^n \sum_{k=1}^{t_{ji}-1} w_{ik} - \sigma^+ + \sigma^- \forall j \quad (2-27)$$

Το βασικό μοντέλο εκτίμησης μπορεί να γραφτεί στη μορφή ενός γραμμικού προγράμματος ως εξής:

$$\left\{ \begin{array}{l} [\min] F = \sum_{j=1}^M \sigma_j^+ + \sigma_j^- \\ \text{υπό τους περιορισμούς} \\ \sum_{i=1}^n \sum_{k=1}^{t_{ji}-1} w_{ik} - \sum_{m=1}^{t_j-1} z_m - \sigma_j^+ + \sigma_j^- = 0 \quad \text{για } j=1,2,\dots,M \\ \sum_{m=1}^{\alpha-1} z_m = 100 \\ \sum_{i=1}^n \sum_{k=1}^{\alpha_i-1} w_{ik} = 100 \\ z_m \geq 0, w_{ik} \geq 0 \quad \forall m,i,k \\ \sigma_j^+ \geq 0, \sigma_j^- \geq 0 \quad \text{για } j=1,2,\dots,M \end{array} \right. \quad (2-28)$$

όπου  $M$  είναι το μέγεθος το δείγματος πελατών, και  $x_i^j, y^j$  είναι το  $j$  επίπεδο στο οποίο υπολογίζονται οι μεταβλητές  $X_i$  και  $Y$ .

Προκειμένου να ξεπεραστεί το πρόβλημα της ευστάθειας του μοντέλου εφαρμόζεται ένα στάδιο ανάλυσης μεταβελτιστοποίησης. Η τελική λύση προκύπτει, εξερευνώντας το πολύεδρο των πολλαπλών ή κοντινών βέλτιστων λύσεων, που παράγεται από τους περιορισμούς του προαναφερθέντος γραμμικού προγράμματος. Αυτή η λύση υπολογίζεται από  $n$  γραμμικά προβλήματα (ίσα με τον αριθμό των κριτηρίων). Τα γραμμικά αυτά προγράμματα μεγιστοποιούν το βάρος  $b_i$  κάθε κριτηρίου και έχουν την ακόλουθη μορφή:

$$\left\{ \begin{array}{l} [\max] F' = \sum_{k=1}^{a_i-1} w_{ik}^{*m+1}, \text{ για } i = 1,2, \dots, n \\ \text{υπό τους περιορισμούς} \\ F \leq F^* + \varepsilon \end{array} \right. \quad (2-29)$$

όπου  $\varepsilon$  είναι μικρό ποσοστό της  $F^*$  (βέλτιστη τιμή της αντικειμενικής συνάρτησης του γραμμικού προγράμματος (2-29)). Ως τελική λύση μπορεί να θεωρηθεί η μέση τιμή των λύσεων που προκύπτουν από τα  $n$  γραμμικά προβλήματα. Στην περίπτωση μη-ευστάθειας η μέση λύση είναι λιγότερο αντιπροσωπευτική

Οι δείκτες συνολικής και μερικής ικανοποίησης υπολογίζονται με βάση τις παρακάτω εξισώσεις:

$$\begin{cases} S = \frac{1}{100} \sum_{m=1}^a p^m y^{*m} \\ S_i = \frac{1}{100} \sum_{k=1}^a p_i^k x_i^{*k} \text{ για } i = 1, 2, \dots, n \end{cases} \quad (2-30)$$

όπου  $S$  και  $S_i$ , οι δείκτες συνολικής και μερικής ικανοποίησης,  $p^m$  και  $p_i^k$  είναι οι συχνότητες των πελατών που ανήκουν στο  $y^m$  και  $x_i^k$  επίπεδο ικανοποίησης, αντίστοιχα.

Οι εκτιμώμενες συναρτήσεις ικανοποίησης ουσιαστικά εκφράζουν την αξία που προσδίδει το σύνολο των πελατών σε ένα καθορισμένο ποιοτικό επίπεδο ικανοποίησης. Μέσα από τα αποτελέσματα ικανοποίησης προκύπτουν επίσης συμπεράσματα σχετικά με το επίπεδο απαιτητικότητας των χρηστών.

Ανάλογα με το βαθμό απαιτητικότητας οι πελάτες κατηγοριοποιούνται σε κανονικούς, απαιτητικούς και μη απαιτητικούς.

Το παρακάτω σχήμα παρουσιάζει 3 βασικές ομάδες πελατών με διαφορετικό βαθμό απαιτητικότητας (τα αποτελέσματα ισχύουν τόσο για την ολική, όσο και για τις μερικές συναρτήσεις ικανοποίησης):

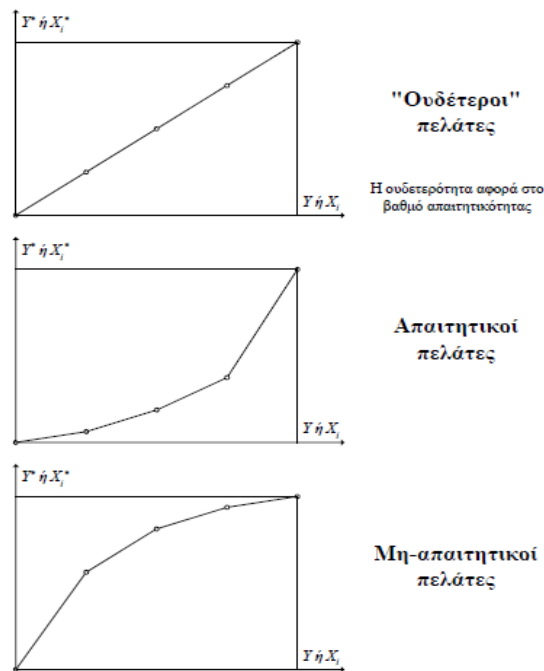
Κανονικοί πελάτες: η συνάρτηση ικανοποίησης έχει γραμμική μορφή, γεγονός που σημαίνει ότι οι συγκεκριμένοι πελάτες όσο περισσότερο ικανοποιημένοι δηλώνουν ότι είναι, τόσο μεγαλύτερο είναι το ποσοστό των προσδοκιών τους που εκπληρώνεται.

Απαιτητικοί πελάτες: η συνάρτηση ικανοποίησης έχει κυρτή μορφή, δεδομένου ότι η ομάδα αυτή των πελατών δεν είναι ικανοποιημένη παρά μόνο αν τους προσφέρεται το βέλτιστο επίπεδο υπηρεσιών.

Μη-απαιτητικοί πελάτες: η συνάρτηση ικανοποίησης έχει κοίλη μορφή, γεγονός που υποδηλώνει ότι οι συγκεκριμένοι πελάτες δηλώνουν ότι είναι



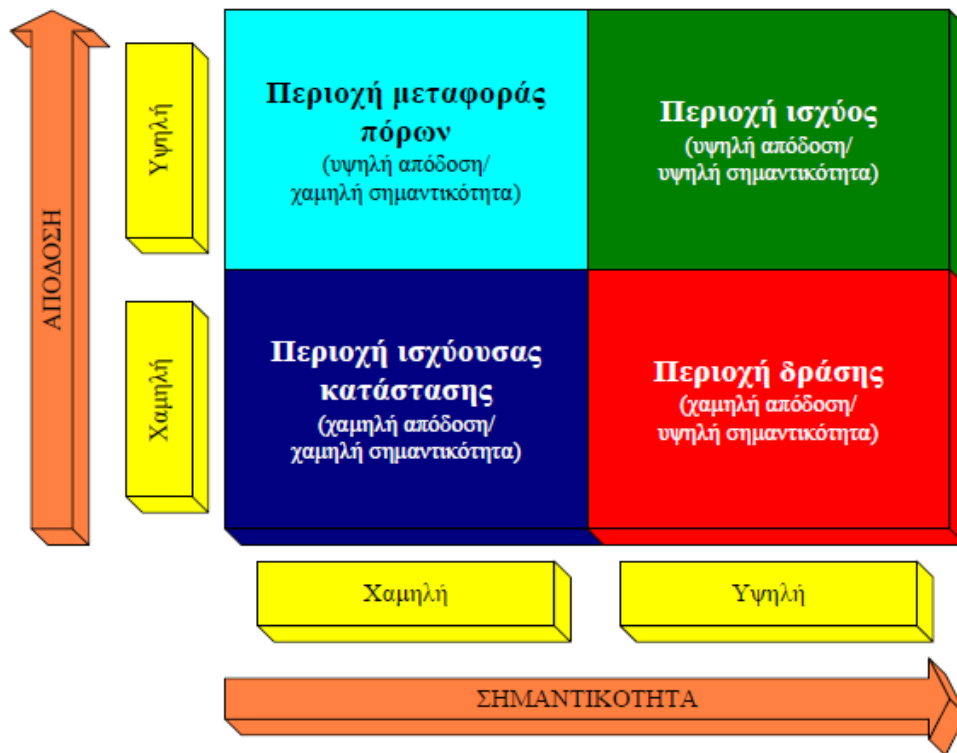
ικανοποιημένοι παρόλο που ένα μικρό ποσοστό των προσδοκιών τους εκπληρώνεται.



Σχ. 2-5 Ομάδες πελατών με διαφορετικό βαθμό απαιτητικότητας

Τα βάρη των κριτηρίων ικανοποίησης υποδηλώνουν το σχετικό βαθμό σπουδαιότητας που δίνει το σύνολο των πελατών στις αξίες των διαστάσεων ικανοποίησης που έχουν καθοριστεί. Φυσικά αυτό έχει άμεση συνάφεια και με τον αριθμό των κριτηρίων που χρησιμοποιούνται.

Συνδυάζοντας τα βάρη ικανοποίησης με τους μέσους δείκτες ικανοποίησης υπολογίζονται τα διαγράμματα δράσης. Οι διαστάσεις των συγκεκριμένων διαγραμμάτων είναι η απόδοση και η σημαντικότητα των κριτηρίων. Στη βιβλιογραφία συναντιούνται επίσης σαν χάρτες απόφασης και ως στρατηγικοί χάρτες. Κάθε διάγραμμα δράσης (Σχήμα 2-6) χωρίζεται σε 4 τεταρτημόρια. Με αυτό τον τρόπο είναι δυνατός ο προσδιορισμός των απαιτούμενων ενεργειών για τη βελτίωση ή διατήρηση του επιπέδου ικανοποίησης των πελατών. Κάθε τεταρτημόριο αποτελεί μια περιοχή διαφορετικής κατάστασης για τα κριτήρια ικανοποίησης. Στη συνέχεια παρατίθενται σχηματικά και περιγραφικά οι 4 διαφορετικές περιοχές που διαμορφώνονται.



Σχ. 2-6 Διάγραμμα Δράσης

**Περιοχή Ισχύουσας Κατάστασης (Χαμηλή απόδοση – Χαμηλή Σημαντικότητα):** Στη περίπτωση αυτή δεν χρειάζεται κάποια πρόσθετη ενέργεια από τον πάροχο υπηρεσίας για το κριτήριο που βρίσκεται στη περιοχή αυτή καθώς φαίνεται να μην έχει μεγάλη σημασία για τον πελάτη.

**Περιοχή Μεταφοράς Πόρων (Υψηλή Απόδοση - Χαμηλή Σημαντικότητα):** Οι πόροι που αφορούν το κριτήριο στη συγκεκριμένη περιοχή θα μπορούσαν να μεταφερθούν αλλού ή να χρησιμοποιηθούν με διαφορετικό τρόπο.

**Περιοχή Δράσης (Χαμηλή Απόδοση - Υψηλή Σημαντικότητα):** Στη συγκεκριμένη περιοχή ανήκουν τα κριτήρια-χαρακτηριστικά που χρήζουν άμεσης βελτίωσης προκειμένου να αυξηθεί το επίπεδο ικανοποίησης των πελατών.

**Περιοχή Ισχύος (Υψηλή Απόδοση – Υψηλή Σημαντικότητα):** Το συγκεκριμένα κριτήρια αποτελούν το συγκριτικό πλεονέκτημα του παρόχου έναντι των ανταγωνιστών του. Συνεπώς δεν υπάρχει λόγος κάποιας τροποποίησης των διαδικασιών που σχετίζονται με τα συγκεκριμένα κριτήρια.

Το παραπάνω διάγραμμα (Σχήμα 2-6) μπορεί επιπρόσθετα να καθορίσει την ιεράρχηση της σπουδαιότητας των ενεργειών βελτίωσης για τις κρίσιμες διαστάσεις ικανοποίησης:

- Η περιοχή δράσης είναι προφανώς η πρώτη προτεραιότητα της επιχείρησης, δεδομένου ότι στο συγκεκριμένο τεταρτημόριο ανήκουν σημαντικά κριτήρια για τα οποία οι πελάτες δεν είναι ικανοποιημένοι.
- Η δεύτερη προτεραιότητα θα πρέπει επικεντρωθεί στις διαστάσεις ικανοποίησης που ανήκουν στην περιοχή ισχύος, ειδικά όταν υπάρχουν περιθώρια βελτίωσης (ο δείκτης ικανοποίησης είναι κοντά στον κάθετο άξονα).
- Η περιοχή της ισχύουσας κατάστασης είναι η τρίτη κατά σειρά προτεραιότητα της επιχείρησης. Παρόλο που οι συγκεκριμένες διαστάσεις ικανοποίησης δεν είναι ιδιαίτερα κρίσιμες την περίοδο της ανάλυσης, ενδέχεται να γίνουν σημαντικές στο μέλλον, ενώ δεν πρέπει να λησμονείται το γεγονός ότι η ικανοποίηση των πελατών είναι χαμηλή για τα συγκεκριμένα κριτήρια.
- Η τελευταία προτεραιότητα της επιχείρησης θα πρέπει να είναι η περιοχή μεταφοράς πόρων, διότι περιλαμβάνει χαρακτηριστικά τα οποία αφενός δεν είναι σημαντικά για τους πελάτες και αφετέρου η απόδοση της εταιρείας είναι υψηλή.

Σε αρκετές περιπτώσεις είναι χρήσιμη η χρήση των σχετικών διαγραμμάτων δράσης. Προκειμένου να σχηματιστούν αυτά, γίνεται χρήση των μεταβλητών  $b'_i$  και  $S'_i$  (Εξισώσεις 2-31) προκειμένου να ξεπεραστεί το πρόβλημα αξιολόγησης του κατ' αποκοπή επιπέδου για τους άξονες σημαντικότητας και απόδοσης. Οι προαναφερθείσες μεταβλητές υπολογίζονται ως εξής:

$$\text{Σημαντικότητα} \quad b'_i = \frac{b_i - \bar{b}}{\sqrt{\sum_i (b_i - \bar{b})^2}}$$

$$\text{Αποδοτικότητα} \quad S'_i = \frac{s_i - \bar{s}}{\sqrt{\sum_i (s_i - \bar{s})^2}}$$

(2-31)

όπου  $\bar{b}$  και  $\bar{S}$  είναι οι μέσες τιμές των βαρών των κριτηρίων και οι μέσοι δείκτες ικανοποίησης αντίστοιχα. Αυτός ο τύπος διαγράμματος είναι χρήσιμος, αν τα σημεία είναι συσσωρευμένα σε μια μικρή περιοχή λόγω της χαμηλής διακύμανσης που εμφανίζεται στους δείκτες μέσης ικανοποίησης. Τα συγκεκριμένα διαγράμματα εμφανίζονται στη βιβλιογραφία επίσης σαν χάρτες απόφασης, στρατηγικοί χάρτες και είναι παρόμοια με τα διαγράμματα που προκύπτουν από SWOT ανάλυση.

Τα διαγράμματα δράσης μπορούν να υποδείξουν ποιες διαστάσεις ικανοποίησης πρέπει να βελτιωθούν, αλλά δεν είναι σε θέση να προσδιορίσουν ποιο θα είναι το αποτέλεσμα των ενεργειών βελτίωσης, ούτε το μέγεθος της προσπάθειας που χρειάζεται για να επιτευχθεί η προσδοκώμενη βελτίωση.

Το πρόβλημα αυτό λύνεται με την κατασκευή των διαγραμμάτων βελτίωσης (Σχήμα 2-7), όπου:

- Οι μέσοι δείκτες απαιτητικότητας δείχνουν το μέγεθος της προσπάθειας που καταβάλλεται για τη βελτίωση ενός χαρακτηριστικού, δεδομένου ότι όσο πιο απαιτητικοί είναι οι πελάτες, τόσο περισσότερο πρέπει να βελτιωθεί το επίπεδο ικανοποίησης για να εκπληρωθούν οι προσδοκίες τους.
- Το αποτέλεσμα των ενεργειών βελτίωσης εξαρτάται τόσο από τη σημαντικότητα του κριτηρίου, όσο και από τη συνεισφορά του στη μη-ικανοποίηση των πελατών.

Για το λόγο αυτό, ορίζεται ένα σύνολο μέσων δεικτών αποτελεσματικότητας  $I_i$  σύμφωνα με τις σχέσεις:

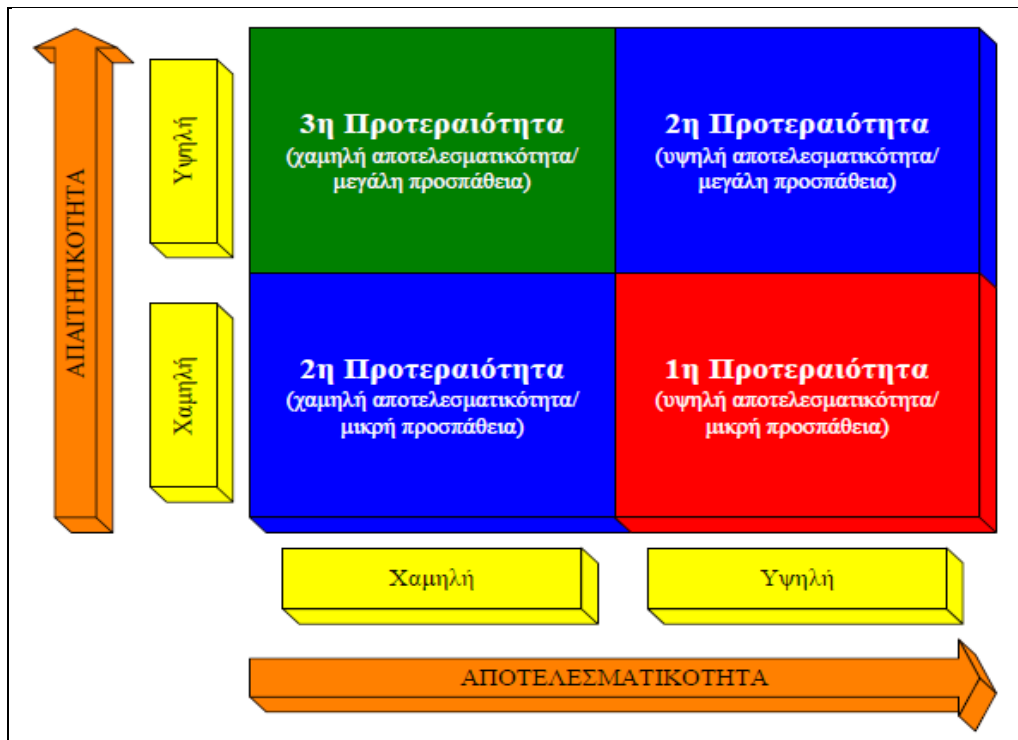
$$I_i = b_i(1 - S_i) \text{ για } i=1,2,3..n \quad (2-32)$$

Οι συγκεκριμένοι δείκτες ορίζονται στο διάστημα  $[0, 1]$  ενώ μπορεί εύκολα να αποδειχθεί ότι:

$$I_i = 1 \Leftrightarrow b_i = 1 \wedge S_i = 0 \quad (2-33)$$

$$I_i = 0 \Leftrightarrow b_i = 0 \wedge S_i = 1 \text{ για } i=1,2,3..n \quad (2-10)$$

Παρακάτω παρατίθεται σχηματικά και περιγραφικά το διάγραμμα βελτίωσης (Σχήμα 2-7) με τις 4 περιοχές προτεραιότητας που δημιουργούνται.



Σχ. 2-7 Διάγραμμα Βελτίωσης

Η επιχείρηση θα πρέπει να επικεντρώσει τις προσπάθειες βελτίωσης στις διαστάσεις ικανοποίησης που έχουν μεγάλη αποτελεσματικότητα ενώ οι πελάτες δεν εμφανίζονται ιδιαίτερα απαιτητικοί.

- Η δεύτερη προτεραιότητα των ενεργειών βελτίωσης αποτελούν τα κριτήρια που είτε παρουσιάζουν μεγάλη αποτελεσματικότητα και μεγάλο βαθμό απαιτητικότητας, είτε εμφανίζουν μικρή αποτελεσματικότητα, ενώ οι πελάτες δε φαίνονται ιδιαίτερα απαιτητικοί.

Τέλος, τα χαρακτηριστικά που παρουσιάζουν μικρή αποτελεσματικότητα και μεγάλη απαιτητικότητα αποτελούν την τελευταία προτεραιότητα βελτίωσης.

#### Εκτίμηση αποτελεσμάτων και Δείκτες Σφάλματος

Η εκτίμηση των αποτελεσμάτων της μεθόδου MUSA, σχετίζεται με τις εξής προβληματικές:

A) το βαθμό προσαρμογής του μοντέλου στα δεδομένα του προβλήματος αξιολόγησης της ικανοποίησης πελατών

B) την ευστάθεια των αποτελεσμάτων της ανάλυσης μεταβελτιστοποίησης.

Η προσαρμογή του μοντέλου σχετίζεται με την εύρεση ενός συστήματος αξιών για το σύνολο των πελατών, με τα ελάχιστα δυνατά σφάλματα. Για το λόγο αυτό, οι βέλτιστες τιμές των μεταβλητών σφάλματος υποδηλώνουν την αξιοπιστία του συστήματος αξιών.

#### A. Δείκτες Προσαρμογής

1. AFI (Average Fitting Index) που ορίστηκε για τη μέθοδο MUSA είναι ο εξής:

$$AFI = 1 - (F^*/100M) \quad (2.35)$$

όπου  $F^*$  είναι η βέλτιστη τιμή σφάλματος του αρχικού γραμμικού προγράμματος του μοντέλου MUSA και  $M$  είναι ο αριθμός των πελατών.

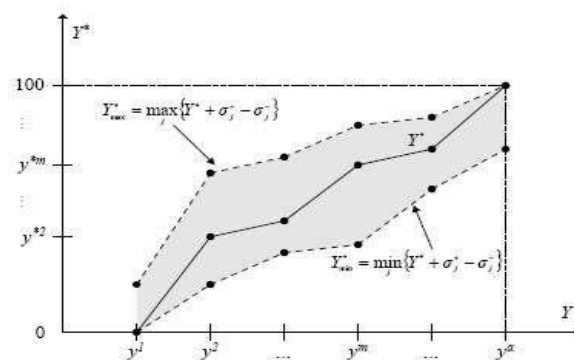
Ο μέσος δείκτης προσαρμογής AFI παίρνει την τιμή 1 μόνο όταν  $F^* = 0$ , δηλαδή όταν το μοντέλο είναι σε θέση να εκτιμήσει ένα σύνολο αξιών για τους πελάτες με μηδενικά σφάλματα. Όμοια, ο μέσος δείκτης προσαρμογής παίρνει την τιμή 0 μόνο όταν  $F^* = 100M$  δηλαδή όταν τα ζεύγη των μεταβλητών σφάλματος και  $σι^+$ ,  $σι^-$  παίρνουν τη μέγιστη τιμή τους.

2. Το διάγραμμα διακύμανσης της ολικής συνάρτησης ικανοποίησης.

Η κατασκευή του διαγράμματος διακύμανσης της ολικής ικανοποίησης βασίζεται τόσο στην εκτιμώμενη αξία ικανοποίησης, όσο και στις τιμές των μεταβλητών σφάλματος και περιλαμβάνει τα εξής βήματα:

- I. Για κάθε πελάτη υπολογίζεται η εκτίμηση της αξίας ικανοποίησης
- II. Η μέγιστη και η ελάχιστη αξία ικανοποίησης για κάθε επίπεδο ικανοποίησης.

Εν συνεχεία με το διάγραμμα διακύμανσης (Γραφική 2-1) της ολικής ικανοποίησης παρατηρούμε το εύρος των τιμών της αξίας ικανοποίησης που δίνει το σύνολο των πελατών για κάθε επίπεδο της ποιοτικής κλίμακας ικανοποίησης. Με αυτόν τον τρόπο είναι εύκολο να φανεί αν υπάρχει κάποιο συγκεκριμένο επίπεδο ικανοποίησης με μεγάλη διακύμανση της εκτιμώμενης ολικής συνάρτησης ικανοποίησης.



Γραφική 2-1 Διάγραμμα διακύμανσης ολικής συνάρτησης ικανοποίησης

3. Ο πίνακας πρόβλεψης ή εκτίμησης της ολικής ικανοποίησης.

Ο πίνακας πρόβλεψης ή εκτίμησης της ολικής ικανοποίησης (Πίνακας 2-4) υπολογίζεται σύμφωνα με τα ακόλουθα βήματα:

1. Για κάθε πελάτη υπολογίζεται η εκτίμηση της αξίας ικανοποίησης.
2. Με βάση την προηγούμενη τιμή, για κάθε πελάτη υπολογίζεται η εκτίμηση του επιπέδου ικανοποίησης.

Με βάση αυτές τις τιμές είναι δυνατό να υπολογιστεί ο αριθμός των πελατών που ανήκουν σε κάθε εκτιμώμενο και πραγματικό επίπεδο ικανοποίησης και να κατασκευαστεί ο πίνακας πρόβλεψης της ολικής ικανοποίησης.

Πίνακας 2-3 Προβλεπόμενο- Πραγματικό επίπεδο ολικής ικανοποίησης

		Προβλεπόμενο επίπεδο ολικής ικανοποίησης											
		$\bar{y}^1$				$\bar{y}^2$				$\bar{y}^j$			
Πραγματικό επίπεδο ολικής ικανοποίησης	$y^1$	$N_{11}$	$R_{11}$	$N_{12}$	$R_{12}$	...	$N_{1j}$	$R_{1j}$	...	$N_{1a}$	$R_{1a}$		
		$C_{11}$		$C_{12}$		...	$C_{1j}$		...	$C_{1a}$			
	$y^2$	$N_{21}$	$R_{21}$	$N_{22}$	$R_{22}$	...	$N_{2j}$	$R_{2j}$	...	$N_{2a}$	$R_{2a}$		
		$C_{21}$		$C_{22}$		...	$C_{2j}$		...	$C_{2a}$			
	$\vdots$												
	$y^j$	$N_{j1}$	$R_{j1}$	$N_{j2}$	$R_{j2}$	...	$N_{jj}$	$R_{jj}$	...	$N_{ja}$	$R_{ja}$		
		$C_{j1}$		$C_{j2}$		...	$C_{jj}$		...	$C_{ja}$			
	$\vdots$												
	$y^a$	$N_{a1}$	$R_{a1}$	$N_{a2}$	$R_{a2}$	...	$N_{aj}$	$R_{aj}$	...	$N_{aa}$	$R_{aa}$		
		$C_{a1}$		$C_{a2}$		...	$C_{aj}$		...	$C_{aa}$			

## Β. Δείκτης Ευστάθειας

Για την ανάλυση της ευστάθειας των αποτελεσμάτων της μεθόδου είναι δυνατό να χρησιμοποιηθεί η διακύμανση των βαρών των κριτηρίων του πίνακα

μεταβελτιστοποίησης. Έτσι, ο μέσος δείκτης ευστάθειας (Average Stability Index) θα μπορούσε να οριστεί ως η μέση τιμή της κανονικοποιημένης τυπικής απόκλισης των εκτιμώμενων βαρών  $b_i$  των κριτηρίων του προβλήματος:

$$ASI = 1 - \frac{1}{n} \sum_{i=1}^n \frac{\sqrt{\lambda \sum_{j=1}^{\lambda} (b_i^j)^2 - (\sum_{j=1}^{\lambda} b_i^j)^2}}{\sqrt{\lambda-1}} \quad (2-11)$$

όπου  $b_{ji}$  το εκτιμώμενο βάρος του κριτηρίου  $i$  κατά την επίλυση του  $j$  γραμμικού προγράμματος της φάσης μεταβελτιστοποίησης, και  $\lambda$  ο αριθμός των γραμμικών προβλημάτων που επιλύονται. Ο μέσος δείκτης ευστάθειας παίρνει τιμές στο διάστημα  $[0, 1]$ .

#### Καθορισμός παραμέτρων για την εφαρμογή MUSA

- Το επιθυμητό μέγεθος  $M$  του συνόλου δεδομένων.
- Το επίπεδο απόκλισης  $D_e$  (με  $D_e$  ανήκει  $[0,1]$ ). Επίπεδο απόκλισης είναι η μέγιστη επιτρεπτή διαφοροποίηση μεταξύ μερικής και ολικής ικανοποίησης. Το επίπεδο απόκλισης παίζει σημαντικό ρόλο στη δημιουργία ενός συνεπούς συνόλου δεδομένων.
- Ο αριθμός των κριτηρίων ικανοποίησης  $n$ .
- Ο αριθμός επιπέδων της κλίμακας ολικής ικανοποίησης  $a$ , καθώς και ο αριθμός των επιπέδων των κλιμάκων που χρησιμοποιούνται για κάθε ένα από τα κριτήρια.
- Κατώφλι προτίμησης  $\gamma$  ( $\gamma=\gamma_i=0$ ): Το κατώφλι προτίμησης εκφράζει το ελάχιστο βήμα αύξησης της συνάρτησης  $Y^*$ ,  $\forall i$ . Η τιμή του  $\gamma$  εξαρτάται από τις παραμέτρους εισόδου  $\alpha$ ,  $n$  του προβλήματος.
- Βαθμός παραχώρησης  $\varepsilon$ : Ο βαθμός παραχώρησης είναι μια μικρή προκαθορισμένη ποσότητα κατά την οποία αυξάνεται η βέλτιστη τιμή της αντικειμενικής συνάρτησης  $F^*$  στη φάση της μεταβελτιστοποίησης, για να αναζητηθούν νέες βέλτιστες λύσεις. Η τιμή του  $\varepsilon$  ορίζεται ως ποσοστό της βέλτιστης τιμής της αντικειμενικής συνάρτησης  $F^*$  που προκύπτει από την φάση της βελτιστοποίησης.



## 2.5 Ανάλυση περιεχομένου και συναισθήματος (Content and Sentiment analysis)

Η συναισθηματική Ανάλυση είναι το πεδίο των σπουδών που αναλύει τις ανθρώπινες απόψεις, συναισθήματα, αξιολογήσεις, εκτιμήσεις, συμπεριφορές και συναισθήματα προς οντότητες όπως προϊόντα, υπηρεσίες, οργανισμοί, φυσικά πρόσωπα, ζητήματα, γεγονότα και στα χαρακτηριστικά τους.

Ο όρος ανάλυση συμπεριφοράς εμφανίστηκε αρχικά στο άρθρο των Nasukawa and Yi, (2003) και ο όρος εξόρυξη άποψης στο αντίστοιχο άρθρο των Dave et al., (2003). Παρόλα αυτά η έρευνα για συναισθήματα και άποψεις εμφανίστηκε νωρίτερα (Das and Chen, 2001).

Η συναισθηματική ανάλυση ή εξόρυξη άποψης αναφέρεται στη χρήση της επεξεργασίας φυσικής γλώσσας, της ανάλυσης κειμένου, της υπολογιστικής γλωσσολογίας και της συναισθηματικής τεχνητής νοημοσύνης για τον συστηματικό εντοπισμό, την ποσοτικοποίηση και την μελέτη των συναισθηματικών καταστάσεων και της υποκειμενικής πληροφορίας που περιλαμβάνεται σε κείμενο.

Η έρευνα στο πεδίο της συναισθηματικής ανάλυσης έχει άμεση συσχέτιση με την επεξεργασία φυσικής γλώσσας και παρατηρείται αρκετή δραστηριότητα στο συγκεκριμένο πεδίο κυρίως τα τελευταία χρόνια. Αυτό οφείλεται κυρίως στην ραγδαία αύξηση εμπορικών εφαρμογών καθώς και στον τεράστιο όγκο δεδομένων που είναι διαθέσιμος στο διαδίκτυο. Η μελέτη της συναισθηματικής ανάλυσης εκτός από το πεδίο της επεξεργασίας φυσικής γλώσσας έχει και άμεση εφαρμογή στις επιστήμες μάνατζμεντ, των οικονομικών, τις κοινωνικές και πολιτικές επιστήμες.

Η συναισθηματική ανάλυση εφαρμόζεται ευρέως για την ανάλυση κριτικών πελατών (reviews) και απαντήσεων σε έρευνες αγοράς σε online κοινωνικά δίκτυα και κοινότητες, κυρίως σχετικά με τον τουρισμό και ιατρικές έρευνες από τα τμήματα μάρκετινγκ και εξυπηρέτησης πελατών.

Οι απόψεις αποτελούν πολύ βασικό κομμάτι των ανθρώπινων δραστηριοτήτων καθώς αποτελούν βασικούς παράγοντες επιρροής για τις συμπεριφορές μας. Οι

επιχειρήσεις και οι οργανισμοί πάντα θέλουν να έχουν στη διάθεση τους απόψεις πελατών για τα προϊόντα τους και οι μεμονωμένοι πελάτες αντίστοιχα, προκειμένου να αποφασίσουν για την αγορά ή όχι ενός προϊόντος. Στο παρελθόν ο τρόπος να εκμαιεύσει κάποιος απόψεις ήταν μέσω γνωστών, ή μέσω ερευνών πελατών, δημοσκοπήσεων και focus γκρουπ. Η κτήση τέτοιους είδους δεδομένων αποτελούσε ανέκαθεν μια προτεραιότητα για τις επιχειρήσεις και τους οργανισμούς εμπορικού ή πολιτικού ενδιαφέροντος. Πλέον, πριν κάποιος αγοράσει ένα προϊόν/υπηρεσία, ανατρέχει σε blogs, δημόσια φόρουμ ή άλλου είδους διαδικτυακές κριτικές.

Η δυσκολία έχει μεταφερθεί πλέον στο φιλτράρισμα και τη σύνοψη όλων των υπαρχόντων πηγών και σχετικών διαδικτυακών πυλών με απόψεις και κείμενα. Αυτό είναι το αντικείμενο στο οποίο δίνει λύσεις η ανάλυση συναισθήματος ή εξόρυξη άποψης. Οι εφαρμογές ανάλυσης συναισθήματος έχουν εξαπλωθεί και χρησιμοποιούνται σε ένα ευρύ φάσμα τομέων, τουρισμό, προϊόντα, υγεία, πολιτική, κοινωνικά γεγονότα. Τεχνολογικοί κολοσσοί όπως η Google, Apple, Microsoft, Amazon, SAP, SAS έχουν αναπτύξει δικές τους εφαρμογές για το συγκεκριμένο πεδίο.

Παράλληλα με τις επιχειρήσεις, έχουν γίνει πολλές ερευνητικές προσπάθειες στην συγκεκριμένη κατεύθυνση. Οι McGlohon et al., (2010) χρησιμοποίησαν κριτικές χρηστών προκειμένου να κατατάξουν προϊόντα και πωλητές. Οι O'Connor et al., (2010), συνέδεσαν το συναίσθημα που προκύπτει από το Twitter με δημοσκοπήσεις κοινής γνώμης, οι Tumasjan et al., (2010), με αντίστοιχα δεδομένα έκαναν πρόβλεψη αποτελεσμάτων εκλογών ενώ οι Bollen et al., (2011) προβήκανε σε προβλέψεις για το δείκτη του χρηματιστηρίου. Τέλος οι Groh and Hauffa, (2011) χρησιμοποίησαν συναισθηματική ανάλυση για να χαρακτηρίσουν κοινωνικές σχέσεις.

Οι πιο σημαντικοί δείκτες συναισθημάτων είναι οι συναισθηματικές λέξεις ή λέξεις άποψης. Αυτές οι λέξεις συνήθως χρησιμοποιούνται για να εκφράσουν θετικά ή αρνητικά συναισθήματα. Για παράδειγμα, οι λέξεις «καλό», «καταπληκτικό», «θεσπέσιο» είναι θετικές συναισθηματικές λέξεις ενώ οι λέξεις «κακό», «ελλιπές», «τραγικό» είναι αρνητικές συναισθηματικές λέξεις. Εκτός

από λέξεις, υπάρχουν και εκφράσεις ή ιδιωματισμοί που αντίστοιχα είναι θετικοί ή αρνητικοί. Μια λίστα τέτοιων λέξεων και εκφράσεων καλείται «συναισθηματικό λεξικό» ή «λεξικό άποψης». Για την προσπέλαση και την βέλτιστη χρησιμοποίηση τέτοιου είδους λεξικών, έχουν αναπτυχθεί πλήθος αλγορίθμων.

Το πιο σημαντικό ζήτημα στη συναισθηματική ανάλυση μετά την δημιουργία και προσπέλαση του λεξικού, είναι η κατανόηση τους ανάλογα με τη χρήση τους.

Για παράδειγμα, μια πρόταση είναι δυνατόν να συμπεριλαμβάνει συναισθηματικές λέξεις που δεν εκφράζουν όμως κάποιο συναίσθημα, ή να περιλαμβάνει συναισθηματικές λέξεις με αντίθετες έννοιες αναλόγως του πώς χρησιμοποιούνται. Επιπλέον πολλές προτάσεις μπορεί να υπονοούν απόψεις, το οποίο είναι δύσκολο να εντοπιστεί, ή να χρησιμοποιούνται με σαρκαστική διάθεση.

Οι απόψεις που καλούμαστε να αναλύσουμε είναι δυνατόν να έχουν εκφραστεί για οτιδήποτε, π.χ. ένα προϊόν, μια υπηρεσία, έναν οργανισμό, ένα γεγονός ή ένα θέμα συζήτησης. Η οντότητα που αποτελεί το στόχο της άποψης καλείται «αντικείμενο» (object), και είναι δυνατόν να έχει ένα σύνολο «συστατικών» ή μερών (components) που το αποτελούν καθώς και ένα σύνολο «χαρακτηριστικών» (attributes) ή ιδιοτήτων. Κάθε συστατικό είναι δυνατόν να αποτελείται από υπο-μέρη με αντίστοιχα χαρακτηριστικά.

Η προαναφερθείσα σχέση θα μπορούσε να περιγραφεί ως εξής:  $o:(T,A)$  όπου  $o$  το αντικείμενο,  $T$  η ιεραρχία των μερών που αποτελούν το αντικείμενο και  $A$  το σύνολο των χαρακτηριστικών του  $o$ .

Με βάση τα προλεχθέντα, ένα αντικείμενο είναι δυνατόν να απεικονίζεται υπό δένδροειδή μορφή. Η ρίζα του δέντρου είναι το αντικείμενο. Κάθε κόμβος διαφορετικός της ρίζας είναι ένα συστατικό του αντικειμένου. Ο κάθε κόμβος σχετίζεται με ένα σύνολο ιδιοτήτων. Μια άποψη μπορεί να εκφραστεί επί του κόμβου ή της ιδιότητας αυτού. Η χρήση των χαρακτηριστικών ή ιδιοτήτων για ένα αντικείμενο χρησιμοποιείται κυρίως όταν το αντικείμενο είναι ένα προϊόν. Στην περίπτωση που το αντικείμενο αποτελεί ένα γεγονός ή ένα θέμα συζήτησης, το χαρακτηριστικό συνήθως παίρνει τη μορφή όψης (aspect). Για

την αποφυγή σύγχυσης των εννοιών, θα χρησιμοποιηθεί εφεξής ο όρος χαρακτηριστικό (feature) για να υποδηλωθεί η ιδιότητα ή όψη ενός αντικειμένου.

Κείμενο άποψης για ένα χαρακτηριστικό: Ένα κείμενο άποψης  $d$  σχετικό με το χαρακτηριστικό  $f$  ενός αντικειμένου  $O$ , αποτελείται από ένα σύνολο συνεχόμενων προτάσεων  $s_i$ , που εκφράζει μια θετική ή αρνητική άποψη επί του χαρακτηριστικού  $f$ .

Σαφώς ορισμένο και υπονοούμενο χαρακτηριστικό (explicit and implicit feature): Αν ένα χαρακτηριστικό  $f$  ή κάποιο συνώνυμο του εμφανίζεται σε μια πρόταση  $s$ , το  $f$  καλείται σαφώς ορισμένο χαρακτηριστικό στην  $s$ . Αν το  $f$  ή κάποιο συνώνυμο του δεν εμφανίζονται σε μια πρόταση  $s$ , αλλά υπονοείται, τότε το  $f$  καλείται υπονοούμενο χαρακτηριστικό.

Κάτοχος άποψης (opinion holder): Ο κάτοχος μιας άποψης είναι το άτομο ή ο οργανισμός που εκφράζει την άποψη.

Οι κάτοχοι άποψης καλούνται επίσης πηγές άποψης (opinion sources)(Wiebe et al., 2005). Στις περιπτώσεις κριτικών για προϊόντα σε blogs οι κάτοχοι άποψης, είναι οι συγγραφείς αυτής.

Άποψη (opinion): Άποψη σε ένα χαρακτηριστικό  $f$  είναι η θετική ή αρνητική διάθεση, συναίσθημα ή οπτική προς το  $f$  από τον κάτοχο άποψης.

Προσανατολισμός άποψης (opinion orientation): Ο προσανατολισμός μια άποψης για ένα χαρακτηριστικό  $f$  υποδεικνύει αν η άποψη είναι θετική, αρνητική ή ουδέτερη. Ο προσανατολισμός άποψης συναντάται στην βιβλιογραφία επίσης ως: προσανατολισμός συναισθήματος (sentiment orientation), πολικότητα άποψης (polarity of opinion), σημασιολογικός προσανατολισμός (semantic orientation).

Άμεση άποψη (direct opinion): Μια άμεση άποψη ορίζεται από τα εξής πέντε στοιχεία  $(o_j, f_{jk}, oo_{ijkl}, h_i, t_i)$ , όπου  $o_j$  είναι το αντικείμενο ή οντότητα,  $f_{jk}$  ένα χαρακτηριστικό του αντικειμένου,  $oo_{ijkl}$  ο προσανατολισμός της άποψης ως προς το χαρακτηριστικό  $f_{jk}$  του αντικειμένου  $o_j$ ,  $h_i$  ο κάτοχος της άποψης και  $t_i$  είναι ο χρόνος που εκφράζεται η άποψη. Ο προσανατολισμός της άποψης  $oo_{ijkl}$  είναι δυνατόν να είναι θετικός, αρνητικός ή ουδέτερος. Για το χαρακτηριστικό  $f_{jk}$  που σχολιάζει ο κάτοχος άποψης  $h_i$ , επιλέγει μια λέξη ή φράση από το  $W$ , ή ένα δείκτη από το  $I$ , για να περιγράψει το χαρακτηριστικό, και εν συνεχεία εκφράζει τη θετική, αρνητική, ή ουδέτερη άποψη του.

Μοντέλο αντικειμένου: Ένα αντικείμενο ο παριστάνεται από ένα ορισμένο σύνολο χαρακτηριστικών,  $F = \{f_1, f_2, \dots, f_n\}$ , το οποίο συμπεριλαμβάνει και το ίδιο αντικείμενο σαν χαρακτηριστικό. Κάθε χαρακτηριστικό  $f_i \in F$ , μπορεί να περιγραφεί από ένα ή παραπάνω σύνολα λέξεων ή φράσεων  $W = \{w_{i1}, w_{i2}, \dots, w_{in}\}$  που είναι συνώνυμα στο χαρακτηριστικό ή υποδηλώνεται από ένα ορισμένο σύνολο δεικτών  $I = \{i_{i1}, i_{i2}, \dots, i_{in}\}$  του χαρακτηριστικού.

Αυτός ο ορισμός, παρόλο που είναι ακριβής, δεν είναι ιδιαίτερα εύχρηστος ιδίως στο τομέα των online κριτικών προϊόντων και υπηρεσιών, καθώς η πλήρης περιγραφή ενός αντικειμένου μπορεί να είναι πολύ σύνθετη, ακόμα και να μην εμφανίζεται στην ίδια πρόταση.

Συγκριτική άποψη (comparative opinion): Η συγκριτική άποψη εκφράζει μια σχέση των ομοιοτήτων ή διαφορών μεταξύ δύο ή περισσότερων αντικειμένων ή επιλογών του κατόχου άποψης, βασιζόμενη σε κάποιο από τα κοινά χαρακτηριστικά τους.

Στις άμεσες απόψεις συναντώνται δύο υπο-κατηγορίες. Στην πρώτη, οι απόψεις εκφράζονται άμεσα προς το αντικείμενο ή κάποιο χαρακτηριστικό του. Στη δεύτερη, οι απόψεις για ένα αντικείμενο εκφράζονται βασιζόμενες στην επιρροή του σε κάποια άλλα αντικείμενα.

Συναισθήματα (emotions): Τα συναισθήματα είναι τα υποκειμενικά μας αισθήματα ή σκέψεις. Σύμφωνα με τον Parrott (2001) , οι άνθρωποι έχουν 6 τύπους κύριων συναισθημάτων: αγάπη, χαρά, έκπληξη, θυμό, θλίψη και φόβο, τα οποία υποδιαιρούνται σε πολλά δευτερεύοντα συναισθήματα. Επίσης κάθε συναίσθημα έχει διαφορετική ένταση. Η δύναμη μιας άποψης είναι απόλυτα σχετική με την ένταση συγκεκριμένων συναισθημάτων. Είναι προφανές λοιπόν ότι υπάρχει ένα πλήθος γλωσσικών εκφράσεων που είναι δυνατόν να χρησιμοποιηθούν, προκειμένου να εκφραστεί ένα συναίσθημα και δει η πολικότητα του.

Η συναισθηματική ανάλυση ή εξόρυξη άποψης ουσιαστικά επιχειρεί να εντοπίσει τα ανθρώπινα συναισθήματα με βάση γλωσσικές εκφράσεις.

Η ισχύς ενός συναισθήματος ή άποψης συνήθως αναλογεί στην ένταση συγκεκριμένων αισθημάτων. Η απόψεις που μελετώνται συνήθως είναι αξιολογήσεις. Οι αξιολογήσεις διαχωρίζονται σε 2 υποκατηγορίες: λογικές αξιολογήσεις και συναισθηματικές αξιολογήσεις.

Λογική αξιολόγηση: Τέτοιου είδους αξιολογήσεις προκύπτουν από λογικά επιχειρήματα, ρεαλιστικές πεποιθήσεις και απτά συμβάντα.

Συναισθηματική αξιολόγηση: Αυτές οι αξιολογήσεις προκύπτουν από μη-απτές και συναισθηματικές απαντήσεις σε αντικείμενα που βρίσκονται βαθιά στο μυαλό των ανθρώπων.

Για την εφαρμογή των δύο παραπάνω τύπων αξιολογήσεων, χρησιμοποιούνται οι εξής μετρικές συναισθήματος:

- i. Συναισθηματικά αρνητική (emotional negative)=-2
- ii. Λογικά αρνητική (rational negative)=-1
- iii. Ουδέτερη(neutral)=0
- iv. Λογικά θετική (rational positive)=+1
- v. Συναισθηματικά θετική (emotional positive)=+2

Η συναισθηματική ανάλυση πραγματοποιείται με εφαρμογή των 6 παρακάτω βημάτων:

Βήμα 1<sup>ο</sup>: Εξαγωγή και κατηγοριοποίηση αντικειμένου. Εξαγωγή όλων εκφράσεων που χαρακτηρίζουν το αντικείμενο και κατηγοριοποίηση τους με βάση τη

συνωνυμία. Κάθε ομάδα εκφράσεων εκφράζει ένα μοναδικό αντικείμενο ή οντότητα.

Βήμα 2<sup>ο</sup>: Εξαγωγή και κατηγοριοποίηση χαρακτηριστικού. Εξαγωγή όλων των εκφράσεων για χαρακτηριστικά των αντικειμένων και κατηγοριοποίηση τους. Κάθε ομάδα εκφράσεων χαρακτηριστικών εκφράζει ένα μοναδικό χαρακτηριστικό.

Βήμα 3<sup>ο</sup>: Εξαγωγή και κατηγοριοποίηση κατόχου άποψης. Εξαγωγή των κατόχων άποψης για απόψεις στο κείμενο ή δομημένα δεδομένα και κατηγοριοποίηση τους.

Βήμα 4<sup>ο</sup>: Εξαγωγή χρόνου και προτυποποίηση. Εξαγωγή των χρόνων που δόθηκαν οι απόψεις και προτυποποίηση των διαφορετικών format χρόνου.

Βήμα 5<sup>ο</sup>: Κατηγοριοποίηση χαρακτηριστικού συναισθήματος. Καθορισμός αν η άποψη για ένα χαρακτηριστικό είναι θετική, αρνητική ή ουδέτερη, ή/και αντιστοίχιση μιας αριθμητικής αξιολόγησης στο χαρακτηριστικό.

Βήμα 6<sup>ο</sup>: Παραγωγή της άποψης με βάση τα 5 στοιχεία της. Παραγωγή όλων των στοιχείων που χαρακτηρίζουν μια άποψη ( $o_j, f_{jk}, oo_{ijkl}, h_i, t_i$ ), και έχουν αναφερθεί στο κείμενο, χρησιμοποιώντας τα αποτελέσματα από τα προηγούμενα βήματα.

Η συναισθηματική ανάλυση που βασίζεται στο παραπάνω πλαίσιο καλείται ανάλυση συναισθήματος με βάση τα χαρακτηριστικά (Hu and Liu, 2004).

### 2.5.1 Ταξινόμηση συναισθήματος σε επίπεδο εγγράφου

Για την ταξινόμηση συναισθήματος σε επίπεδο εγγράφου σαν βασική μονάδα πληροφορίας και δεδομένων, χρησιμοποιείται ένα ολόκληρο έγγραφο.

Ο στόχος είναι, για ένα δεδομένο κείμενο άποψης  $d$ , κατόπιν ανάλυσης να προκύψει το συνολικό συναίσθημα  $s$ , του κατόχου άποψης  $h$ , για το αντικείμενο  $o$  στο οποίο αναφέρεται το κείμενο. Το  $s$  ανάλογα με τον τύπο μεταβλητής και τις τιμές που μπορεί να πάρει, αποτελεί διαφορετικό είδος προβλήματος. Αν το  $s$  παίρνει κατηγορικές τιμές π.χ θετικό, αρνητικό τότε θεωρείται πρόβλημα ταξινόμησης (classification), ενώ αν το  $s$  παίρνει αριθμητικές ή

κανονικοποιημένες τιμές σε συγκεκριμένο εύρος, π.χ 1 έως 5, το πρόβλημα θεωρείται παλινδρόμησης (regression).

Η υπόθεση είναι ότι το κείμενο άποψης αφορά ένα συγκεκριμένο αντικείμενο και έχει εκφραστεί από έναν κάτοχο άποψης. Αυτή η υπόθεση στέκει κυρίως για αξιολόγηση/κριτική προϊόντος ή υπηρεσίας που έχει γραφτεί από ένα πελάτη. Αυτή η υπόθεση δεν αφορά ένα φόρουμ ολόκληρο καθώς μιλάμε για πολλαπλά αντικείμενα, πολλαπλούς κατόχους άποψης και πολλαπλές κριτικές.

### **2.5.2 Ταξινόμηση συναισθήματος με εποπτευόμενη μάθηση**

Στο συγκεκριμένο τύπο ταξινόμησης, σαν σετ δεδομένων εκπαίδευσης και τεστ χρησιμοποιείται τμήμα των κριτικών ενός προϊόντος. Στην περίπτωση αυτή γίνεται αντιστοίχιση των online αξιολογήσεων (συνήθως 1 έως 5) σε λεκτικές εκφράσεις, π.χ. θετική άποψη αν η αξιολόγηση είναι 4 ή 5.

Η ερευνητική βιβλιογραφία περιλαμβάνει ένα πλήθος εργασιών με χρήση αλγόριθμων εποπτευόμενης μάθησης. Ο Gamon,(2004) παρουσίασε κατηγοριοποίηση για δεδομένα που προέρχονται από γνώμες πελατών που συγκρίνονται με κριτικές. Οι Ng et al., (2006) πραγματοποίησαν σε ταξινόμηση με βάση γλωσσικές πηγές γνώσης ενώ οι Martineau and Finin, (2009) πρότειναν έναν νέο τρόπο για να προσδίδεται βαρύτητα σε όρους που καλείται Delta TFIDF. Οι Tokuhisa et al., (2008), ερεύνησαν την ταξινόμηση συναισθήματος σε διαφορετικές διαλογικές εκφράσεις, κατατάσσοντας το συναίσθημα σε 3 κατηγορίες (θετικό, ουδέτερο, αρνητικό), και εν συνεχεία τις θετικές και αρνητικές εκφράσεις σε 10 υποκατηγορίες συναισθήματος.

Οι Kouloumpis et al., (2011) χρησιμοποίησαν γλωσσικά στοιχεία που εντοπίζουν πληροφορία σχετικά με την ανεπίσημη και δημιουργική γλώσσα σε blogs.



### 2.5.3 Ταξινόμηση συναισθήματος με χρήση μη-εποπτευόμενης μάθησης

Ένα χαρακτηριστικό παράδειγμα μη-εποπτευόμενης μάθησης είναι η βασιζόμενη σε λεξικό μέθοδος, η οποία χρησιμοποιεί ένα λεξικό με λέξεις και εκφράσεις με τους αντίστοιχους προσανατολισμούς και δύναμη. Αυτή η μέθοδος ενσωματώνει μεθόδους προσθαφαίρεσης προκειμένου να υπολογιστεί το τελικό σκορ συναισθήματος για κάθε έγγραφο (Taboada et al., 2011, Ding et al., 2008).

### 2.5.4 Ταξινόμηση Υποκειμενικότητας

Η ταξινόμηση υποκειμενικότητας κατηγοριοποιεί τις προτάσεις σε 2 κατηγορίες, τις υποκειμενικές και τις αντικειμενικές (Bruce and Wiebe, 1999). Μια αντικειμενική άποψη εκφράζει πραγματική πληροφορία με γεγονότα, ενώ μια υποκειμενική παρουσιάζει κυρίως απόψεις, αξιολογήσεις, συναισθήματα, πεποιθήσεις.

Σε ερευνητικό επίπεδο αρκετή από τη δουλειά έχει βασιστεί στην εποπτευόμενη μάθηση. Η μέθοδος (Hatzivassiloglou and McKeown, 1997) ήταν από τις αρχικές ερευνητικές προσεγγίσεις και χρησιμοποιήθηκε για να φιλτράρει τις λάθος υποκειμενικές εκφράσεις. Οι Yu and Hatzivassiloglou, (2003) εφάρμοσαν ταξινομήσεις χρησιμοποιώντας ομοιότητα προτάσεων και το Bayes classifier. Η μέθοδος ομοιότητας πρότασης βασίζεται στην παραδοχή ότι οι υποκειμενικές απόψεις είναι πιο όμοιες σε άλλες υποκειμενικές απόψεις παρά σε αντικειμενικές. Σαν μέτρο ομοιότητας χρησιμοποιήθηκε το Simfinder και βασίστηκε σε κοινές λέξεις και εκφράσεις. Οι Pang and Lee, (2004) πρότειναν ένα αλγόριθμο για το διαχωρισμό υποκειμενικών από αντικειμενικές προτάσεις, με βάση τη χρήση γράφου για ένα κείμενο άποψης, όπου για κάθε πρόταση υπολογίζεται ένα σκορ υποκειμενικότητας. Ο γράφος χτίζεται με βάση την τοπικότητα των ετικετών, που σημαίνει ότι αν μια πρόταση είναι κοντά σε μια

άλλη τότε είναι πολύ πιθανό να είναι υποκειμενική ή αντικειμενική ανάλογα της ετικέτας της προηγούμενης.

Οι Wiebe et al., (2005) πραγματεύτηκαν τη χρησιμότητα του εντοπισμού της δύναμης-βάρους της υποκειμενικότητας σε μια πρόταση. Με τη μελέτη τους, εφάρμοσαν αυτόματη ταξινόμηση υποκειμενικότητας, όπου ταξινομούνταν οι όροι μιας πρότασης ανάλογα με τη δύναμη υποκειμενικότητας τους, σε τέσσερα επίπεδα (ουδέτερη, χαμηλή, μεσαία, υψηλή). Η ουδέτερη δείχνει την απουσία υποκειμενικότητας, ενώ οι λέξεις που δείχνουν και καθορίζουν τη δύναμη της υποκειμενικότητας, παράγονται από το δέντρο εξόρυξης εξάρτησης.

### 3. Μεθοδολογία Συστήματος

3.1	Σχηματικό Διάγραμμα
3.2	Ανάλυση και Επεξεργασία Δεδομένων
3.3	Εφαρμογή MUSA
3.4	Εφαρμογή Ανάλυσης Συναισθήματος
3.5	Εφαρμογή Μοντελοποίησης Προτίμησης
3.6	Τελική Σύσταση
3.7	Μηχανισμός Ανατροφοδότησης
3.8	Συνοπτική Μαθηματική παρουσίαση της μεθοδολογίας

### 3. Μεθοδολογία Συστήματος

Στο παρόν κεφάλαιο περιγράφεται το γενικό μεθοδολογικό πλαίσιο και αναλύονται οι διαφορετικές φάσεις υλοποίησης του.

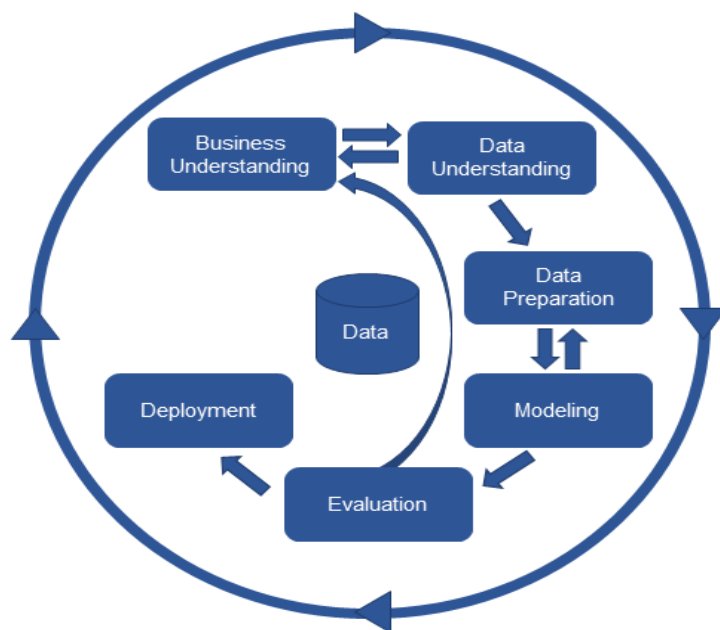
Στόχος της δημιουργίας του νέου συστήματος συστάσεων είναι να εκμεταλλευτούμε στο έπακρον τη διαθέσιμη πληροφορία από απόψεις και βαθμολογίες πελατών για προϊόντα, χρησιμοποιώντας πραγματικές απόψεις και βαθμολογίες που διατίθενται στο διαδίκτυο, ώστε να δημιουργήσουμε προφίλ πελατών που επιλέγουν συγκεκριμένα προϊόντα. Εν συνεχεία στόχος είναι η δημιουργία του προτιμησιακού προφίλ που τρέχει το σύστημα και ενδιαφέρεται για συγκεκριμένες συστάσεις που του «ταιριάζουν». Με έλεγχο της ομοιότητας ή ταύτισης των προτιμησιακών προφίλ του χρήστη του συστήματος με αυτά των χρηστών που επιλέγουν συγκεκριμένα προϊόντα, το σύστημα καταλήγει σε ταξινομημένες συστάσεις προϊόντων προς το χρήστη.

Το μεθοδολογικό πλαίσιο που χρησιμοποιήθηκε βασίστηκε στις έξι βασικές φάσεις του μοντέλου CRISP-DM (Cross Industry Standard Process for Data Mining model) (Shearer, 2000b).

Οι βασικές φάσεις του CRISP-DM (Σχήμα 3-1), είναι οι παρακάτω:

1. Business Understanding (Επιχειρηματική Κατανόηση)
2. Data Understanding (Κατανόηση των Δεδομένων)
3. Data Preparation (Προετοιμασία των Δεδομένων)

4. Modeling (Μοντελοποίηση)
5. Evaluation (Αξιολόγηση)
6. Deployment (Ανάπτυξη)



Σχ. 3-1 Διάγραμμα Φάσεων Crisp DM

Με βάση τις παραπάνω φάσεις, η μεθοδολογία για την παρούσα ερευνητική εργασία διαχωρίστηκε στις ακόλουθες φάσεις:

1. Ανάλυση και κατανόηση του προβλήματος.
2. Στόχος της έρευνας.
3. Απόκτηση δεδομένων και επεξεργασία τους.
4. Μοντελοποίηση.
5. Αξιολόγηση της μεθόδου και αποτελέσματα.

Οι πρώτες δύο φάσεις της μεθοδολογίας έχουν περιγραφεί στα αρχικά κεφάλαια της διατριβής. Παρακάτω περιγράφονται οι λοιπές φάσεις, ενώ στο επόμενο κεφάλαιο πραγματοποιείται αναλυτική περιγραφή της μελέτης περίπτωσης και του συστήματος που υλοποιήθηκε.

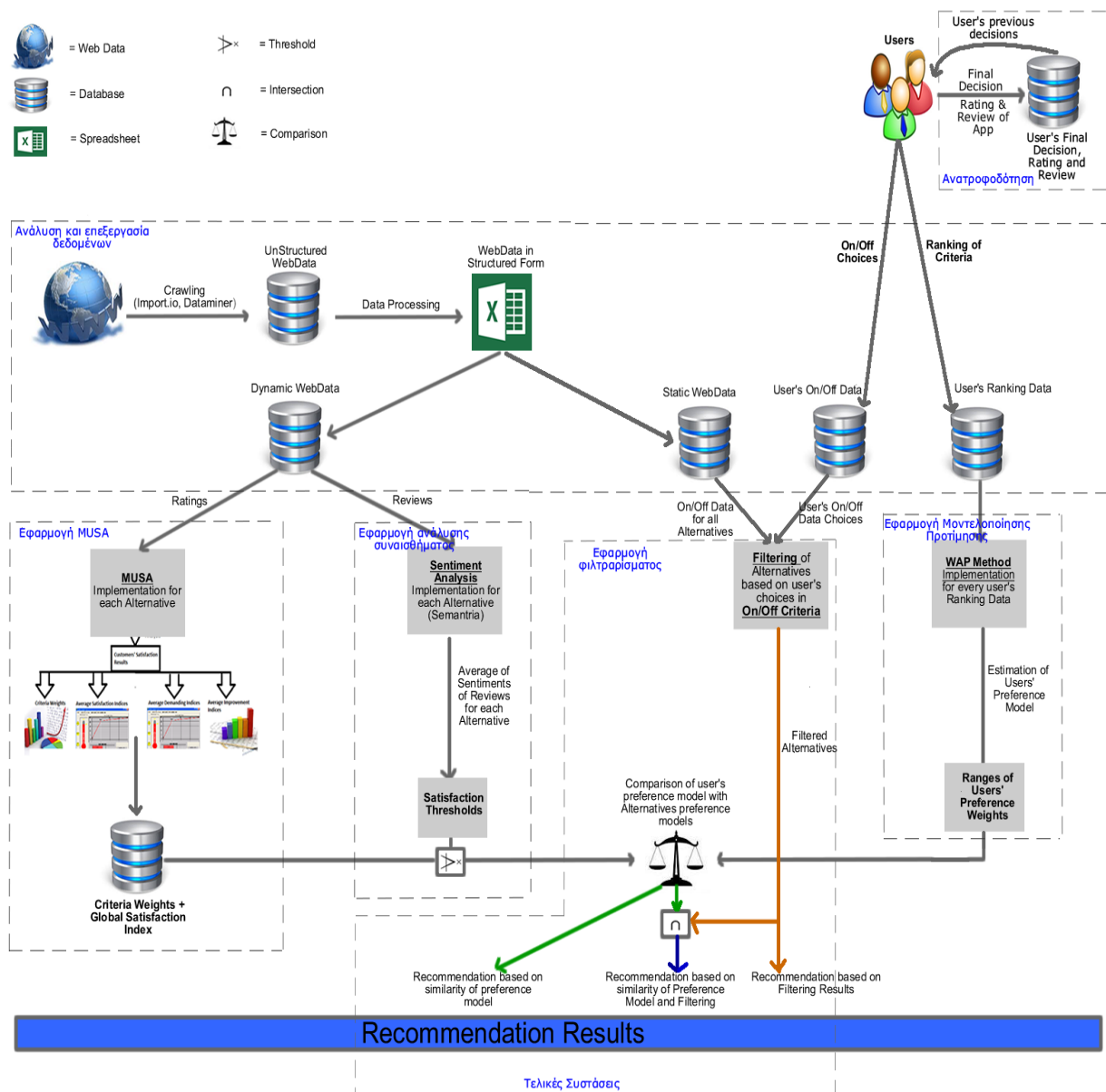
Πιο συγκεκριμένα στο παρόν κεφάλαιο αναλύονται τα παρακάτω:

- η διαδικασία άντλησης δυναμικών και στατικών δεδομένων από το Διαδίκτυο και το χρήστη,

- η διαδικασία εφαρμογής Πολυκριτήριας Ανάλυσης Ικανοποίησης Πελατών στα αριθμητικά δεδομένα που έχουν αντληθεί,
- η διαδικασία εφαρμογής ανάλυσης συναισθήματος στα δεδομένα κειμένου που έχουν αντληθεί,
- η διαδικασία ανάλυσης του προτιμησιακού προφίλ του χρήστη με εφαρμογή της μεθόδου WAP,
- η διαδικασία φιλτραρίσματος των εναλλακτικών βάση των επιλογών του χρήστη,
- η διαδικασία σύνθεσης των αποτελεσμάτων των προαναφερθεισών αναλύσεων ώστε να προκύψει τελική σύσταση για το χρήστη,
- ο μηχανισμός ανατροφοδότησης του συστήματος με τις αποφάσεις του χρήστη.

### 3.1 Σχηματικό Διάγραμμα

Παρακάτω παρουσιάζεται στο σχηματικό διάγραμμα της εικόνας (Σχήμα 3-2) το γενικό μεθοδολογικό πλαίσιο που αναπτύχθηκε με συνοπτική αναφορά στις φάσεις της μεθοδολογίας. Εν συνεχεία ακολουθεί αναλυτική περιγραφή των φάσεων.



Σχ. 3-2 Σχηματικό Διάγραμμα λειτουργιών νέου Συστήματος Συστάσεων

Οι κύριες φάσεις του μεθοδολογικού πλαισίου είναι οι παρακάτω:

- ⇒ Ανάλυση και επεξεργασία Δεδομένων (Παράγραφος 3.2)
  - Πραγματοποιείται άντληση των δυναμικών (βαθμολογιών και σχολίων) και στατικών (χαρακτηριστικά προϊόντων) δεδομένων από διαδικτυακή πηγή (π.χ πλατφόρμα κρατήσεων ξενοδοχείων).
  - Επεξεργασία και διαμόρφωση των δεδομένων σε μορφή χρησιμοποιήσιμη για τις μεθοδολογικές αναλύσεις.

- Τα δυναμικά δεδομένα αποτελούν εισόδους για τις μεθόδους MUSA (βαθμολογίες-ratings) και ανάλυσης συναισθήματος (σχόλια-reviews).
  - Τα στατικά δεδομένα θα αποτελέσουν βάση δεδομένων για τα χαρακτηριστικά των προϊόντων, που χρησιμοποιείται εκ των υστέρων, για τις On/Off επιλογές του χρήστη, επί των χαρακτηριστικών των εναλλακτικών που τον ενδιαφέρουν, για το φιλτράρισμα των τελευταίων.
- ⇒ Εφαρμογή MUSA (Παράγραφος 3.3)
- Είσοδο για την εφαρμογή της MUSA αποτελούν οι βαθμολογίες-ratings των πελατών για κάθε εναλλακτική.
  - Πραγματοποιείται ανάλυση ικανοποίησης με χρήση MUSA για καθεμία από τις εναλλακτικές.
  - Από τα αποτελέσματα της MUSA για κάθε εναλλακτική, συλλέγουμε σε βάση δεδομένων τα βάρη των κριτηρίων καθώς και το δείκτη συνολικής ικανοποίησης που προκύπτουν. Τα συγκεκριμένα αποτελέσματα αποτελούν και την έξοδο από την εφαρμογή MUSA.
  - Με τα βάρη των κριτηρίων από την εφαρμογή της MUSA έχουμε διαθέσιμο το προτιμησιακό προφίλ των πελατών που επέλεξαν συγκεκριμένα προϊόντα.
- ⇒ Εφαρμογή Ανάλυσης Συναισθήματος (Παράγραφος 3.4)
- Είσοδο για την εφαρμογή της MUSA αποτελούν τα σχόλια πελατών για κάθε εναλλακτική.
  - Πραγματοποιείται ανάλυση συναισθήματος για καθεμία από τις εναλλακτικές.
  - Υπολογίζεται η μέση τιμή συναισθήματος για κάθε εναλλακτική.
  - Εφαρμόζεται κατώφλι ικανοποίησης στις εναλλακτικές, με βάση τη μέση τιμή συναισθήματος που έχει προκύψει.
  - Εξόδους τη συγκεκριμένης φάσης της μεθοδολογίας αποτελούν οι εναλλακτικές που η μέση τιμή συναισθήματος τους, ξεπερνά το κατώφλι ικανοποίησης.
- ⇒ Εφαρμογή Φιλτραρίσματος (Παράγραφος 3.2)
- Κατά τη συγκεκριμένη φάση της μεθοδολογίας, ο χρήστης του συστήματος συστάσεων έχει τη δυνατότητα να επιλέξει ανάμεσα σε συγκεκριμένα χαρακτηριστικά των εναλλακτικών, αυτά που θεωρεί απαραίτητα να υπάρχουν στα προϊόντα που θα του συσταθούν τελικά.
  - Γίνεται σύγκριση των επιλογών του χρήστη, με τη βάση στατικών δεδομένων που περιλαμβάνει όλα τα χαρακτηριστικά των διαθέσιμων εναλλακτικών.

- Με βάση τη σύγκριση, πραγματοποιείται φιλτράρισμα των εναλλακτικών.
- ⇒ Εφαρμογή Μοντελοποίησης Προτίμησης (Παράγραφος 3.5)
  - Ο χρήστης του συστήματος συστάσεων κατατάσσει τα κριτήρια για την επιλογή εναλλακτικής. (Παράγραφος 3.2)
  - Η κατάταξη των κριτηρίων καθώς και η επιλογή της διαφοράς σημαντικότητας μεταξύ των συνεχόμενων κριτηρίων, αποτελούν τα δεδομένα εισόδου της μεθόδου WAP.
  - Το αποτέλεσμα της WAP είναι το προτιμησιακό προφίλ του χρήστη, που καθορίζεται με τα εύρη βαρών των κριτηρίων, καθώς την τιμή της ευρωστίας για κάθε χρήστη του συστήματος συστάσεων.
- ⇒ Τελική Σύσταση (Παράγραφος 3.6)
  - Για την τελική σύσταση, γίνεται σύγκριση του προτιμησιακού μοντέλου του χρήστη του συστήματος συστάσεων, που αποτελεί την έξοδο της εφαρμογής μοντελοποίησης προτίμησης, με τα προτιμησιακά μοντέλα των πελατών που έχουν επιλέξει συγκεκριμένες εναλλακτικές, που αποτελούν την έξοδο της εφαρμογής MUSA. Επιπλέον εφαρμόζεται το κατώφλι συναισθήματος από την εφαρμογή ανάλυσης συναισθήματος.
  - Στην περίπτωση σύγκλισης, προκύπτει σύσταση με βάση την ταύτιση των προτιμησιακών μοντέλων.
  - Εν συνεχεία ενσωματώνονται οι εναλλακτικές που έχουν προκύψει από το φιλτράρισμα των εναλλακτικών, με βάση τα On/Off χαρακτηριστικά που έχει επιλέξει ο χρήστης. Ως τελική σύσταση, σε αυτή τη περίπτωση προκύπτει η τομή των συνόλων των εναλλακτικών που έχουν προκύψει από το φιλτράρισμα και από την ταύτιση προτιμησιακών μοντέλων.
- ⇒ Μηχανισμός Ανατροφοδότησης (Παράγραφος 3.7)
  - Ο χρήστης του συστήματος συστάσεων, καλείται να εισάγει την τελική του απόφαση κατόπιν της σύστασης, καθώς και να βαθμολογήσει, σχολιάσει το σύστημα.
  - Με νέα είσοδο του στο σύστημα, του υπενθυμίζονται οι πρόσφατες επιλογές του σε επίπεδο προϊόντων.
  - Εφαρμόζεται ανάλυση συναισθήματος στα σχόλια του χρήστη για βελτίωση του συστήματος.



### 3.2 Ανάλυση και Επεξεργασία Δεδομένων

Τα δεδομένα που χρησιμοποιούνται στην παρούσα μεθοδολογία διαχωρίζονται σε δύο διαφορετικές κατηγορίες:

A. τα δεδομένα που προέρχονται από το διαδίκτυο και γίνεται άντληση τους με μεθόδους εξόρυξης δεδομένων,

B. τα δεδομένα που αποτελούν επιλογές του χρήστη του συστήματος, αφορούν την κατάταξη των κριτηρίων για ανάπτυξη του προτιμησιακού μοντέλου και τα κατ' αποκοπή κριτήρια, τα οποία εισάγονται από τον ίδιο το χρήστη.

I. Αρχικά πραγματοποιείται άντληση δυναμικών και στατικών δεδομένων από το Διαδίκτυο. Τα δεδομένα από το Διαδίκτυο αφορούν αξιολογήσεις προϊόντων ή υπηρεσιών που μπορεί ένας χρήστης να εντοπίσει και να αγοράσει με χρήση διαδικτυακών πλατφορμών. Δυναμικά δεδομένα στο παρόν μεθοδολογικό πλαίσιο θεωρούνται οι βαθμολογίες (ratings) και τα σχόλια (reviews) των χρηστών για ένα προϊόν ή μια υπηρεσία, ενώ στατικά δεδομένα είναι τα χαρακτηριστικά του προϊόντος ή της υπηρεσίας.

Οι εφαρμογές που χρησιμοποιήθηκαν για την άντληση των δεδομένων είναι η import.io (<https://www.import.io/>) και η dataminer (<https://data-miner.io/>).

Η import.io. είναι εφαρμογή που διατίθεται ελεύθερα στο διαδίκτυο και έχει τη μορφή ενός τυπικού προγράμματος περιήγησης όπως ο Chrome της Google. Δίνει στο χρήστη τη δυνατότητα να παραμετροποιήσει web-crawlers (προγράμματα ανίχνευσης) για την άντληση συγκεκριμένων δεδομένων σε ένα σύνολο πανομοιότυπων σελίδων που ορίζει ο ίδιος. Είναι αποδοτικό ως εργαλείο άντλησης δεδομένων, σε ιστότοπους που έχουν πολλαπλές ιστοσελίδες με παρόμοια μορφή, όπως ιστότοπους ηλεκτρονικού εμπορίου με προϊόντα ή ιστότοπους με βαθμολογίες και σχόλια χρηστών που τα χαρακτηριστικά τους εμφανίζονται στα ίδια σημεία πάνω στην ιστοσελίδα και τα url των ιστοσελίδων για τα προϊόντα έχουν επίσης παρόμοιο format π.χ (<https://www.frysi.com/product/1> , <https://www.frysi.com/product/2> ).

Κατά την αρχική λειτουργία του import.io, προκειμένου να παραμετροποιηθούν οι crawlers, υπάρχει μια λειτουργία εκπαίδευσης τους, κατά τη διάρκεια της οποίας ο χρήστης εισάγει urls από διαφορετικά προϊόντα στο σύστημα ώστε το

σύστημα να δημιουργήσει ένα πρότυπο για τα urls των σελίδων που χρειάζεται να προσπελάσει κατά την κανονική λειτουργία του. Εν συνεχεία ο χρήστης επιλέγει, μέσα στις σελίδες που προτείνει για τη δημιουργία του προτύπου, τα χαρακτηριστικά των οποίων τα δεδομένα τον ενδιαφέρει να αντληθούν. Με αυτό τον τρόπο το σύστημα εκπαιδεύεται και ως προς την τοποθεσία των δεδομένων μέσα στη σελίδα που χρειάζεται να αντλήσει. Αφού εκπαιδευτεί το σύστημα με βάση την παραμετροποίηση του χρήστη, ξεκινάει η προσπέλαση των δεδομένων, η οποία ανάλογα με το πλήθος των ιστοσελίδων του ιστοτόπου, είναι δυνατόν να διαρκέσει από κάποιες ώρες, μέχρι και μέρες. Στην περίπτωση της παρούσας μεθοδολογίας το import.io, παραμετροποιήθηκε ώστε να αντλήσει όλα τα δεδομένα βαθμολογιών και σχολίων των ξενοδοχείων. Η παραμετροποίηση απαιτεί αρκετή εξοικείωση με την εφαρμογή αλλά το αποτέλεσμα της άντλησης δεδομένων είναι ικανοποιητικό. Τα επιλεγμένα δεδομένα προς άντληση μετατρέπονται σε πίνακες δεδομένων που τους ορίζει ο χρήστης και είναι δυνατόν να εξαχθούν σε μορφές εύκολα επεξεργάσιμες π.χ. spreadsheets.

Το dataminer αποτελεί μια επέκταση του προγράμματος περιήγησης Google Chrome για την επιλεκτική άντληση (scrap) δεδομένων από ιστοσελίδες και εξαγωγή τους σε spreadsheets του Microsoft Excel. Η διαφορά του με το import.io είναι ότι απαιτεί, παρά την εύκολη παραμετροποίηση του, προκειμένου να αντλεί δεδομένα, την συνεχή συμμετοχή του χρήστη κατά την άντληση των δεδομένων σε αντίθεση με το import.io που μετά την απαραίτητη παραμετροποίηση προχωρά στην άντληση δεδομένων αυτόνομα χρησιμοποιώντας τη νόρμα που του έχει ορίσει ο χρήστης εκ των προτέρων. Για να γίνει χρήση της εφαρμογής, ο χρήστης χρειάζεται να ορίσει σε κάθε ιστοσελίδα ξεχωριστά τα χαρακτηριστικά για τα οποία θέλει να αντλήσει δεδομένα. Αυτή η διαδικασία απαιτεί αρκετούς περισσότερους πόρους σε επίπεδο χρόνου και προσπάθειας, αλλά είναι αποδοτική σε σελίδες που τα χαρακτηριστικά δεν βρίσκονται πάντα στο ίδιο σημείο, οπότε δεν μπορεί να δημιουργηθεί ένα πρότυπο για συνεχή άντληση δεδομένων, όπως στην περίπτωση του import.io. Η dataminer χρησιμοποιήθηκε για την άντληση των στατικών χαρακτηριστικών των ξενοδοχείων, τα οποία διέφεραν για τα

διαφορετικά ξενοδοχεία, με αποτέλεσμα να βρίσκονται σε διαφορετικά σημεία της σελίδας. Τα δεδομένα, όπως και στην περίπτωση του import.io, είναι δυνατό να αποθηκευτούν σε spreadsheets.

Τα δεδομένα που αντλούνται με χρήση των προαναφερθεισών εφαρμογών, αποθηκεύονται σε μορφή συνεχόμενου κειμένου σε spreadsheets.

Παρακάτω παρατίθεται εικόνα (Σχήμα 3-3) από ένα spreadsheet με δεδομένα βαθμολογιών, απόψεων για ξενοδοχείο, αποτέλεσμα της άντλησης. Κάθε γραμμή αποτελεί μια διαφορετική αξιολόγηση χρήστη για ένα ξενοδοχείο. Στο πάνω μέρος της εικόνας εμφανίζεται μια αξιολόγηση σε όλη της έκταση, στη μορφή που αντλείται με χρήση του import.io.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	5.0 / 5 Outstanding "nice location" Cleanliness 5.0 Service 5.0 Comfort 5.0 Condition 5.0 Neighbourhood 5.0 hotel is on the quiet side of the old port yet in a short distance to the various restaurants and the shopping area . is about 10min walking distance from the bus stop if you are coming in the town with airport bus. Genuine Hotels.com guest review 3 night trip in Oct 2013 A Traveller IE																	
1	5.0 / 5 Outstanding	"nice location"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	hotel is on the quiet side of the old port yet in a short distance to the various restaurants and the shopping area . is about 10min walking distance from the bus stop if you are coming in the town with airport bus. Genuine Hotels.com guest review 3 night trip in Oct 2013 A Traveller IE										
2	5.0 / 5 Outstanding	"an excellent hotel."	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 4.0	Have just returned from a week's stay at this charming hotel. It is situated in a beautiful area with lovely views and is very close to the beach. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a relaxing holiday.										
3	5.0 / 5 Outstanding	"Great hotel!!!"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	Perfect breakfast with local cuisine . Beautiful rooms and a very polite staff. Genuinely enjoyed our stay here and would recommend it to anyone looking for a great holiday.										
4	5.0 / 5 Outstanding	"Wonderful hotel contributed to a wonderful vacatio"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	This hotel exceeded my expectations and was a great place to stay. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
5	5.0 / 5 Outstanding	"Fabulous Getaway"	Cleanliness 5.0	Service 4.0	Comfort 4.0	Condition 5.0	Neighbourhood 5.0	Everything was fab. Room was a little too 'business like', but comfy beds, great breakfast, no charge for drinks. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
6	5.0 / 5 Outstanding	"Oasis in city"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 4.0	Neighbourhood 5.0	Enjoyed the hotel very much, staff helpful, great breakfast, no charge for drinks. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
7	5.0 / 5 Outstanding	"Great hotel"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	Excellent hotel with great staff and facilities. Would certainly stay there again. Genuine Hotels.com guest review 3 night trip in Oct 2013 A Traveller IE										
8	5.0 / 5 Outstanding	"Superb value"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	Manto studios has it all at amazing value. Fantastic location (metres from bus stop) and very close to the beach. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
9	5.0 / 5 Outstanding	"Faultless!"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	Beautiful sunset views overlooking the old venetian harbour, extremely helpful and friendly staff. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
10	5.0 / 5 Outstanding	"Excellent 5 star quality Hotel on the waterfront."	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	Excellent coffee and breakfastFree bicycleThe perfect spot Genuine Hotels.com guest review 3 night trip in Oct 2013 A Traveller IE										
11	5.0 / 5 Outstanding	"Nice hotel on water close to action but not loud"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	We could not fault this Hotel. It was beautiful and the staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
12	4.0 / 5 Excellent	"Nice hotel on water close to action but not loud"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 4.0	Neighbourhood 5.0	This hotel was recommended by Greek friends who had a great stay. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
13	5.0 / 5 Outstanding	"What a hotel should be"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	Great hotel, newly refurbished, it's what you expect a hotel should be. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
14	5.0 / 5 Outstanding	"Excellent hotel brilliant location"	Cleanliness 5.0	Service 5.0	Comfort 4.0	Condition 4.0	Neighbourhood 5.0	The hotel is a perfect match of value and money. The location is excellent and the staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
15	5.0 / 5 Outstanding	"Fanatastic Service"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	Had a fabulous time staying at this hotel. It was very clean and the staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
16	5.0 / 5 Outstanding	"Great location, the best sunsets from your balcony"	Cleanliness 5.0	Service 5.0	Comfort 5.0	Condition 5.0	Neighbourhood 5.0	the best hotel of the Old Town most probably. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										
17	5.0 / 5 Outstanding	"Great hotel, great location"	Cleanliness 5.0	Service 5.0	Comfort 4.0	Condition 5.0	Neighbourhood 5.0	In old town but at the end of the strip so very quiet. Room was not too big but perfect for a couple. The staff were very helpful and the food was excellent. I would definitely recommend this hotel to anyone looking for a great holiday.										

Σχ. 3-3 Spreadsheet με αποτελέσματα άντλησης

Προκειμένου να γίνουν επεξεργασίμα, εφαρμόζονται συναρτήσεις κειμένου και συμβολοσειρών του Excel για τελική διαμόρφωση τους, σε μορφή πινάκων. Οι συναρτήσεις που χρησιμοποιούνται είναι οι εξής: LEFT, RIGHT, MID, LEN, FIND, SEARCH, REPLACE.

Στην παρακάτω εικόνα (Σχήμα 3-4) παρουσιάζονται οι συναρτήσεις που χρησιμοποιήθηκαν για την επεξεργασία των αρχικών δεδομένων, προκειμένου να κατηγοριοποιηθούν οι τιμές των χαρακτηριστικών και να μπορέσουμε εν συνεχεία να προχωρήσουμε σε περαιτέρω χρήση τους για την μεθοδολογική

ανάλυση, για την περίπτωση μιας κριτικής, από αυτές που εμφανίζονται στην παραπάνω εικόνα.

Στο κελί Α1 παρουσιάζεται μια κριτική, με τη μορφή που αντλήθηκε από τον ιστότοπο. Εν συνεχεία παρατίθενται οι συναρτήσεις που χρησιμοποιήθηκαν για να εξαχθεί η τιμή για κάθε ένα από τα κριτήρια καθώς και το σχόλιο του χρήστη, τα ονόματα των κριτηρίων, καθώς και οι τιμές. Οι συγκεκριμένες συναρτήσεις χρησιμοποιήθηκαν για καθεμία από τις κριτικές των ξενοδοχείων ώστε να προκύψουν δομημένοι πίνακες με τιμές για τα κριτήρια.

Με αντίστοιχο τρόπο κατηγοριοποιήθηκαν και τα υπόλοιπα δεδομένα που αντλήθηκαν με τη χρήση των `import.io`.

[illegible]

Σχ. 3-4 Συναρτήσεις για επεξεργασία των αρχικών δεδομένων

Στην επόμενη εικόνα παρουσιάζεται η μορφή που έχουν τα δεδομένα κατόπιν της προαναφερθείσας επεξεργασίας.

Review No	Hotel Name		Overall Review Score	Cleanliness	Service	Comfort	Condition	Neighbourhood	Review	Reviewer Origin
1	Cavo Spada Luxury Resort		5	4	5	5	4	3	The staff in this hotel is fantastic	IE
2	Cavo Spada Luxury Resort		5	5	5	5	5	4	Nothing much to say ... everything was perfect	IE
3	Cavo Spada Luxury Resort		5	5	5	5	4	4	We experienced an 'ultra inclusive' holiday	GB
4	Cavo Spada Luxury Resort		5	5	5	5	5	5	Our stay at Cavo Spada was very relaxing	GB
5	Cavo Spada Luxury Resort		4	5	4	5	5	4	Arrived at hotel to be told they were closed	GB
6	Cavo Spada Luxury Resort		4	5	4	5	4	5	We had a lovely stay at this hotel	GB
7	Cavo Spada Luxury Resort		5	5	5	5	5	5	My wife and I are celebrating our 25th wedding anniversary	US
8	Cavo Spada Luxury Resort		5	5	5	5	5	3	We were 2 families total 8 people	GB
9	Cavo Spada Luxury Resort		5	5	5	5	5	5	Our experience was A1 - the staff were excellent	GB
10	Cavo Spada Luxury Resort		4	5	4	5	4	4	We had a 5 night all inclusive holiday	GB
11	Cavo Spada Luxury Resort		5	4	4	4	4	4	5 Lovely spacious hotel. Always clean	GB
12	Cavo Spada Luxury Resort		4	4	4	4	4	3	Enjoyed our stay here although a bit noisy	GB
13	Cavo Spada Luxury Resort		5	5	5	5	5	5	The Cavo Spada is a beautiful hotel	GB
14	Cavo Spada Luxury Resort		4	5	4	4	4	4	Having stayed for a week at Cavo Spada	GB

**Σχ. 3-5 Μορφή δεδομένων κατόπιν της επεξεργασίας**

Τα δεδομένα που αντλούνται με τη χρήση του dataminer σε spreadsheet ενσωματώνονται υπό τη μορφή συνεχόμενων κελιών σε στήλη. Ουσιαστικά κάθε διαφορετικό χαρακτηριστικό εμφανίζεται σε επόμενη γραμμή της ίδιας στήλης. Στην παρακάτω εικόνα (Σχήμα 3-6) εμφανίζονται μερικά από τα χαρακτηριστικά ενός ξενοδοχείου στην αρχική μορφή άντλησης.

Activities			
Fishing			
Hiking			
Divling			
Snorkelling			
Massage			
Biogole rental			
Living Area			
Seating Area			
Sofa			
Desk			
Media & Technology			
Telephone			
TV			
Radio			
Flat-screen TV			
Food & Drink			
Restaurant			
Bar			
Breakfast in the room			
Snack bar			
Minibar			
Internet			

Σχ. 3-6 Χαρακτηριστικά ξενοδοχείου στην αρχική μορφή άντλησης

Τα συγκεκριμένα δεδομένα κατηγοριοποιούνται για κάθε ξενοδοχείο με χρήση των συναρτήσεων: if, iserror, vlookup, και προκύπτει η μορφή που απεικονίζει στην παρακάτω εικόνα(3-7).

D140		=IF(ISERROR(VLOOKUP(D\$1,'Kalyves Beach'!\$A:\$A,1,0)),0, 1)																					
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
1	Hotel Name	Stars	Outdoor pool (seasonal)	Outdoor pool	Massage	TV	Restaurant	Bar	Free! WiFi is available in all areas and is free of charge.	Free! private parking is possible on site (reservation is not needed).	WiFi is available in all areas and charge is applicable.	Free! Free private parking is possible on site (reservation is not needed).	Free! private parking is possible on site (reservation is not needed).	Free! Free public parking is possible on site (reservation is not needed).	Room service	Non-smoking rooms	Lift	English	German	Russian	French	Spanish	
140	Kalyves Beach	4Stars	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	
141	AKALI hotel	4Stars	1	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0	
142	Bozzali	4Stars	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	
143	Hotel Captain Vasili	4Stars	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	
144	Grand Bay Beach Res	4Stars	0	0	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	0	
145	Kydon Hotel	4Stars	0	0	0	0	0	0	1	1	0	0	1	0	1	1	1	1	1	0	0	0	
146	Louis Creta Princess	4Stars	0	1	0	1	0	1	0	0	1	0	0	1	0	0	1	1	0	1	0	0	
147	Fatma Boutique Hot	4Stars	0	0	1	0	0	1	1	0	0	0	0	0	0	1	1	0	1	0	0	0	
148	Paradise Hotel Luxu	4Stars	0	1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	0	0	0	0	
149	Samaria Hotel	4Stars	0	1	0	0	0	1	1	0	0	0	0	0	1	1	1	1	1	0	0	0	
150	Santa Marina Beach	4Stars	1	0	1	1	1	1	0	1	0	0	1	0	1	1	1	1	1	0	0	0	

Σχ. 3-7 Κατηγοριοποιημένα Χαρακτηριστικά ανά ξενοδοχείο

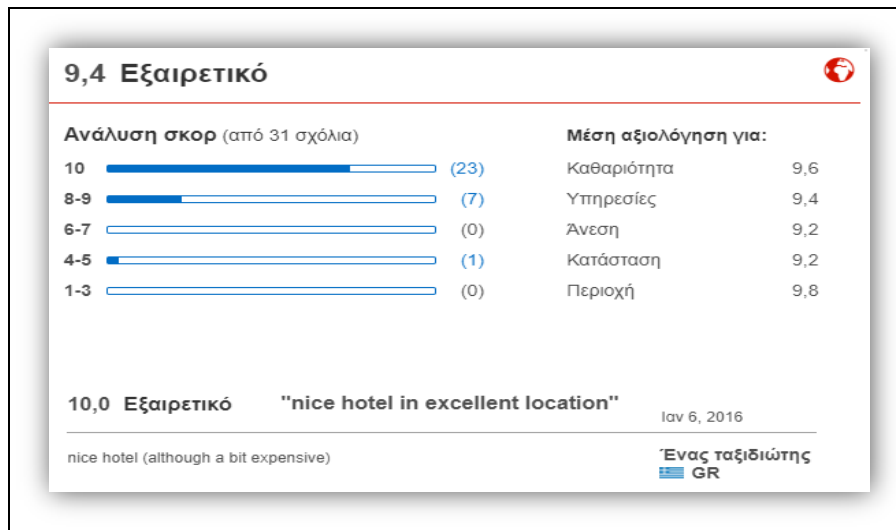
Η μορφή των δεδομένων στην περίπτωση των βαθμολογιών (ratings) απαιτείται να είναι συγκεκριμένη, προκειμένου να είναι αυτά συμβατά με τις αναγκαίες συνθήκες που διέπουν τα είδη των αναλύσεων που υλοποιούνται στα πλαίσια της μεθοδολογίας. Πιο συγκεκριμένα έστω ότι κάθε χρήστης  $U_g \in U$ , όπου  $g=1,2,3,\dots,n$ , με το  $n$  να είναι ο συνολικός αριθμός χρηστών που αξιολογεί τα προϊόντα  $P_i \in P_R$ , όπου  $P_R$  το σύνολο αναφοράς, για κάθε εναλλακτική  $P_i$ ,  $i=1,2,3,\dots,m$ , όπου  $i$  ο αριθμός των εναλλακτικών, ο χρήστης  $U$  θα πρέπει να υποβάλει μια βαθμολογία  $r_{c,i,g}$ , για κάθε κριτήριο  $c_j$ ,  $j=1,2,3,\dots,k$ , όπου  $k$  είναι ο συνολικός αριθμός κριτηρίων, ακολουθώντας μια προκαθορισμένη κλίμακα μέτρησης όσον αφορά τις τιμές κριτηρίων. Επιπρόσθετα είναι απαραίτητο ο χρήστης να παρέχει μια συνολική βαθμολογία προτίμησης  $O_{i,g}$  (overall score). Αρκετές από τις υπάρχουσες διαδικτυακές πλατφόρμες υπολογίζουν αυτόματα τη συνολική βαθμολογία προτίμησης σαν μέσο όρο των βαθμολογιών των επιμέρους κριτηρίων, κάτι που έχει σαν αποτέλεσμα τα δεδομένα ratings που παρέχουν να είναι μη συμβατά με τις μεθόδους ανάλυσης της παρούσας μεθοδολογίας.

Παρακάτω στους πίνακες 3-1, 3-2, παρουσιάζονται τα Δυναμικά και Στατικά Δεδομένα και ο τρόπος δόμησης τους σε πίνακα για την περαιτέρω ανάλυση τους ενώ στα σχήματα 3-8, 3-9, παρουσιάζονται πραγματικά screenshots από τη μορφή που έχουν τα δεδομένα στο διαδίκτυο πριν αντληθούν.

Πίνακας 3-1 Δυναμικά Δεδομένα

Χρήστης	Εναλλακτική	Συνολική Βαθμολογία	Βαθμολογία 1 <sup>ου</sup> Κριτηρίου	Βαθμολογία 2 <sup>ου</sup> Κριτηρίου	Βαθμολογία j <sup>ου</sup> Κριτηρίου	Σχόλιο
$U_1$	$P_1$	$O_{11}$	$r_{111}$	$r_{211}$	$r_{j11}$	review <sub>11</sub>
$U_2$	$P_1$	$O_{12}$	$r_{112}$	$r_{212}$	$r_{j12}$	review <sub>12</sub>
$U_3$	$P_2$	$O_{23}$	$r_{123}$	$r_{223}$	$r_{j23}$	review <sub>23</sub>
$U_4$	$P_2$	$O_{24}$	$r_{124}$	$r_{224}$	$r_{j24}$	review <sub>24</sub>
$U_5$	$P_2$	$O_{25}$	$r_{125}$	$r_{225}$	$r_{j25}$	review <sub>25</sub>
$U_6$	$P_3$	$O_{36}$	$r_{136}$	$r_{236}$	$r_{j36}$	review <sub>36</sub>
$U_g$	$P_i$	$O_{ig}$	$r_{iig}$	$r_{2ig}$	$r_{jig}$	review <sub>ig</sub>





Σχ. 3-8 Παράδειγμα κριτικής και βαθμολογιών πελάτη από διαδικτυακή πλατφόρμα

Πίνακας 3-2 Στατικά δεδομένα

Εναλλακτική	Χαρακτηριστικό 1	Χαρακτηριστικό 2	Χαρακτηριστικό Υ
P <sub>1</sub>	Ναι	ΟΧΙ	ΟΧΙ
P <sub>2</sub>	ΟΧΙ	Ναι	ΟΧΙ
P <sub>3</sub>	ΟΧΙ	Ναι	Ναι
P <sub>4</sub>	ΟΧΙ	Ναι	Ναι
P <sub>5</sub>	ΟΧΙ	ΟΧΙ	Ναι
P <sub>6</sub>	Ναι	ΟΧΙ	Ναι
P <sub>i</sub>	Ναι	ΟΧΙ	Ναι

Τιμή (για 1 νύχτα)

Ελάχιστο (€) 0 Μέγιστο (€) 500+

Κατηγορία αστέρων

1★ 2★ 3★ 4★ 5★

Αξιολόγηση επισκεπτών

2+ 4+ 6+ 8+ 10

Παροχές

- ☐ Internet
- ☐ Άμεση πρόσβαση σε πιστές σκι
- ☐ Γυμναστήριο
- ☐ Διαθέσιμος χώρος στάθμευσης
- ☐ Δωμάτια που επικοινωνούν εσωτερικά
- ☐ Δωρεάν πρωινό
- ☐ Επιτρέπονται τα κατοικίδια
- ☐ Επιχειρηματικές παροχές
- ☐ Εστιατόριο
- ☐ Κουζίνα
- ☐ Κρεβάτια μωρού
- ☐ Μεταφορά από/προς το αεροδρόμιο
- ☐ Μη καπνιστών
- ☐ Μπανιέρα στο δωμάτιο
- ☐ Μπαρ
- ☐ Παροχές συνεδριάσεων
- ☐ Πισίνα
- ☐ Δωρεάν WiFi

Σχ. 3-9 Παράδειγμα On/Off κριτηρίων από διαδικτυακή πλατφόρμα

II. Τα δεδομένα που εισάγονται από το χρήστη διακρίνονται σε δύο είδη: α. επιλογές του χρήστη στα κατ' αποκοπή κριτήρια, β. επιλογές του χρήστη ως προς την κατάταξη των κριτηρίων.

α. Τα δεδομένα σχετικά με τις επιλογές του χρήστη στα κατ' αποκοπή κριτήρια παίρνουν τις εξής τιμές: No choice, Yes, No. Αφορούν τα στατικά χαρακτηριστικά των προϊόντων/υπηρεσιών που διερευνά ο χρήστης, και με βάση τις συγκεκριμένες επιλογές του χρήστη, πραγματοποιείται φιλτράρισμα των εναλλακτικών.

β. Οι επιλογές του χρήστη σχετικά με την κατάταξη των κριτηρίων, αποτελούν τις βασικές εισόδους των αλγορίθμων που καταλήγουν στη διαμόρφωση του προτιμησιακού προφίλ του χρήστη. Τα δεδομένα που προκύπτουν από τις συγκεκριμένες επιλογές έχουν τη μορφή σειράς κατάταξης. Σχετικά με την προτιμησιακή σχέση είναι απαραίτητο να περιγραφούν τα παρακάτω.

Έστω ότι  $A$  είναι το σύνολο όλων των εναλλακτικών, τότε το  $\leq$  αποτελεί μια προτιμησιακή σχέση στο  $A$  αν υπάρχει δυαδική σχέση στο  $A$  τέτοια ώστε  $b \leq a$  αν και μόνο αν το  $a$  είναι το λιγότερο τόσο προτιμητέο όσο το  $b$ . Σε αυτή τη περίπτωση θεωρείται ότι το  $a$  είναι ασθενώς προτιμητέο από το  $b$ . Αν ισχύει  $b \leq a$  αλλά όχι  $a \leq b$  τότε θεωρείται ότι ο χρήστης έχει απόλυτη προτίμηση στο  $a$  σε σχέση με το  $b$  και αυτό γράφεται  $b < a$ . Αν ισχύει  $a \leq b$  και  $b \leq a$  τότε ο χρήστης θεωρείται αδιάφορος για τις εναλλακτικές  $a$  και  $b$ .

Μια προτιμησιακή σχέση πρέπει να υπόκειται στις εξής παραδοχές:

1. Η σχέση να είναι αντιμεταθετική:  $a \leq a$ .
2. Η σχέση να είναι μεταβατική: Αν  $a \leq b$  και  $b \leq c$  τότε  $a \leq c$ .
3. Η σχέση να είναι πλήρης: Για όλα τα  $a, b$  στο  $A$ , να έχουμε  $a \leq b$  ή  $b \leq a$  ή και τα δύο.
4. Η σχέση να είναι συνεχής.



### 3.3 Εφαρμογή MUSA

Προκειμένου να υπολογιστούν τα προτιμησιακά μοντέλα των χρηστών, που έχουν επιλέξει και βαθμολογήσει συγκεκριμένες εναλλακτικές, στα πλαίσια της παρούσας μεθοδολογίας χρησιμοποιείται η μέθοδος πολυκριτήριας ανάλυσης ικανοποίησης πελατών MUSA (Grigoroudis and Siskos, 2002),(Grigoroudis and Siskos, 2010).

#### Χρήση MUSA στο παρόν μεθοδολογικό πλαίσιο

Στο παρόν μεθοδολογικό πλαίσιο πραγματοποιείται ανάλυση ικανοποίησης για το σύνολο των πελατών καθενός από τα προϊόντα, με εισόδους βαθμολογίες  $r_{c,i,g}$ , (ratings) που έχουν αντληθεί από διαδικτυακή πλατφόρμα, και αφορούν καθορισμένα κριτήρια  $c_j$ , και συνολική αξιολόγηση  $O_{ig}$ , εναλλακτικών  $P_i$  που έχουν αξιολογηθεί από χρήστες  $U_g$  κατόπιν της επιλογής τους και χρήσης τους. Οι βαθμολογίες ακολουθούν προκαθορισμένη κλίμακα μέτρησης, η οποία χρησιμοποιείται εξ αρχής στη διαδικτυακή πλατφόρμα. Τα δεδομένα εισόδου της MUSA έχουν την παρακάτω μορφή (Πίνακας 3-3).

Πίνακας 3-3 Δεδομένα εισόδου MUSA

Χρήστης	Εναλλακτική	Συνολική Βαθμολογία	Βαθμολογία 1 <sup>ου</sup> Κριτηρίου	Βαθμολογία 2 <sup>ου</sup> Κριτηρίου	Βαθμολογία j Κριτηρίου
$U_1$	$P_1$	$O_{11}$	$r_{111}$	$r_{211}$	$r_{j11}$
$U_2$	$P_1$	$O_{12}$	$r_{112}$	$r_{212}$	$r_{j12}$
$U_3$	$P_2$	$O_{23}$	$r_{123}$	$r_{223}$	$r_{j23}$
$U_4$	$P_2$	$O_{24}$	$r_{124}$	$r_{224}$	$r_{j24}$
$U_5$	$P_2$	$O_{25}$	$r_{125}$	$r_{225}$	$r_{j25}$
$U_6$	$P_3$	$O_{36}$	$r_{136}$	$r_{236}$	$r_{j36}$
$U_g$	$P_i$	$O_{ig}$	$r_{iig}$	$r_{2ig}$	$r_{jig}$

Πραγματοποιείται ανάλυση ικανοποίησης πελατών για κάθε μια από της εναλλακτικές  $P_i$ , και ενώ αντλούνται όλα τα αναφερθέντα στην προηγούμενη παράγραφο αποτελέσματα, αυτά που χρησιμοποιούνται για το παρόν μεθοδολογικό πλαίσιο είναι τα εξής:

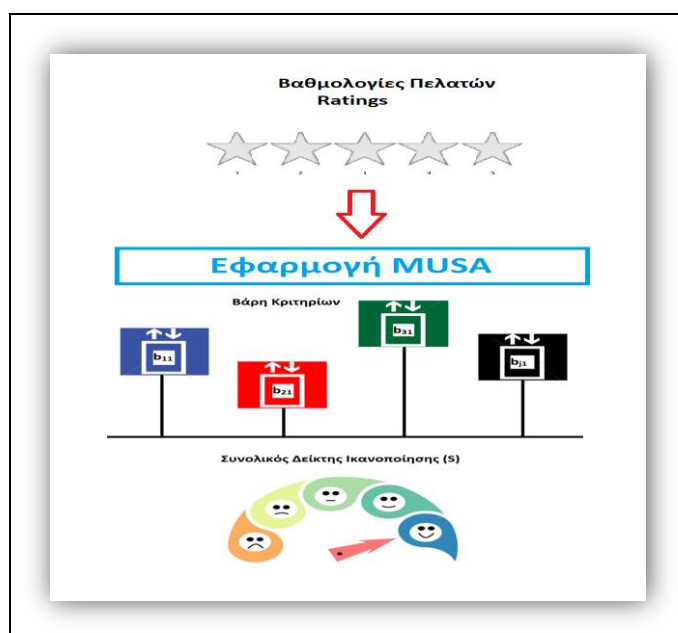
1. Τα βάρη των κριτηρίων  $b_{ji}$ .
2. Ο δείκτης συνολικής ικανοποίησης  $S_i$ .

Ο πίνακας αποτελεσμάτων της MUSA έχει την εξής μορφή (Πίνακας 3-4):

Πίνακας 3-4 Αποτελέσματα MUSA

Εναλλακτική	Δείκτης Συνολικής Ικανοποίησης	Βάρος 1 <sup>ου</sup> Κριτηρίου	Βάρος 2 <sup>ου</sup> Κριτηρίου	Βάρος j Κριτηρίου
$P_1$	$S_1$	$b_{11}$	$b_{21}$	$b_{j1}$
$P_2$	$S_2$	$b_{12}$	$b_{22}$	$b_{j2}$
$P_3$	$S_3$	$b_{13}$	$b_{23}$	$b_{j3}$
$P_4$	$S_4$	$b_{14}$	$b_{24}$	$b_{j4}$
$P_5$	$S_5$	$b_{15}$	$b_{25}$	$b_{j5}$
$P_6$	$S_6$	$b_{16}$	$b_{26}$	$b_{j6}$
$P_i$	$S_i$	$b_{i1}$	$b_{i2}$	$b_{ji}$

Επίσης παρακάτω παραθέτουμε ένα απλοποιημένο σχηματικό διάγραμμα για τον τρόπο με τον οποίο μέσω της εφαρμογής MUSA, μεταβαίνουμε από τις βαθμολογίες (ratings) που είχαν αντληθεί, στα βάρη κριτηρίων και το συνολικό δείκτη ικανοποίησης που είναι οι έξοδοι της MUSA και χρησιμοποιούνται από την μεθοδολογία μας για τα περαιτέρω βήματα της σύστασης.



Σχ. 3-10 Σχηματικό διάγραμμα εισόδων-εξόδων MUSA στην μεθοδολογία

Με την εφαρμογή της MUSA στα παραπάνω δεδομένα και την εξαγωγή των περιγεγραμμένων αποτελεσμάτων, ουσιαστικά έχει επιτευχθεί η δημιουργία του προτιμησιακού μοντέλου των χρηστών που έχουν επιλέξει συγκεκριμένες εναλλακτικές. Συνεπώς έχουμε μια αντιστοίχιση κάθε εναλλακτικής με ένα συγκεκριμένο προτιμησιακό μοντέλο, αυτό που προκύπτει από τις βαθμολογίες των χρηστών τους.

### 3.4 Εφαρμογή Ανάλυσης Συναισθήματος

Ανάλυση συναισθήματος (Sentiment Analysis) ή εξόρυξη απόψεων (Opinion Mining) χαρακτηρίζεται η υπολογιστική μελέτη απόψεων, συναισθημάτων και αισθημάτων που εκφράζονται μέσω κειμένου.

Για την περίπτωση συναισθηματικής ανάλυσης των απόψεων μοναδικών χρηστών για μοναδικά αντικείμενα σε ένα έγγραφο απόψεων, όπως στην περίπτωση των απόψεων πελατών για προϊόντα, είθισται να χρησιμοποιείται Ταξινόμηση συναισθήματος σε επίπεδο εγγράφου (Document Level Sentiment Classification). Οι συγκεκριμένες απόψεις ανήκουν στην κατηγορία των άμεσων απόψεων.

Για την ανάλυση συναισθήματος στο παρόν μεθοδολογικό πλαίσιο γίνεται χρήση του λογισμικού ανάλυσης συναισθήματος Semantria (<https://semantria.readme.io/>). Πιο συγκεκριμένα γίνεται χρήση της εφαρμογής-προσθήκης του Semantria για το Excel. Το Semantria δέχεται σαν είσοδο ένα κείμενο και δίνει σαν έξοδο ένα αριθμητικό αποτέλεσμα συναισθήματος (Sentiment Score) του συγκεκριμένου κειμένου, καθώς και τον προσανατολισμό συναισθήματος (sentiment polarity). Το σημαντικό πλεονέκτημα του συγκεκριμένου λογισμικού είναι ότι υποστηρίζει 16 διαφορετικές γλώσσες. Με την εφαρμογή του Semantria σε κείμενο, πέραν των βασικών προαπαιτούμενων αποτελεσμάτων sentiment score και polarity, εξάγει αυτόματα αποτελέσματα ανάλυσης για συγκεκριμένες οντότητες που υπάρχουν μέσα στο κείμενο, όπως ονόματα πόλεων, ανθρώπων, εταιρειών, @mentions,

#hashtags. Με αυτό τον τρόπο είναι δυνατό να πραγματοποιηθεί ανάλυση συναισθήματος και σε μεγαλύτερη λεπτομέρεια, πέραν της ανάλυσης του συναισθήματος του συνολικού κειμένου. Παρόλα αυτά η παρούσα μεθοδολογία απαιτεί το συνολικό συναίσθημα κάθε άποψης για συγκεκριμένο αντικείμενο.

Προκειμένου να υπολογιστεί το αριθμητικό αποτέλεσμα συναισθήματος, το Semantria χρησιμοποιεί μια λογαριθμική κλίμακα για το συναίσθημα. Σε αντίθεση με τη γραμμική κλίμακα όπου η επόμενη τιμή θα είναι πάντα μεγαλύτερη από τη προηγούμενη κατά μια συγκεκριμένη σταθερή τιμή π.χ. 1,2,3..., η λογαριθμική κλίμακα έχει μια τελείως διαφορετική εξάπλωση, όπου κάθε επόμενη τιμή έχει μια εκθετική αύξηση. Για δεκαδικό λογάριθμο η κλίμακα θα είναι 1,10,100,1000.κ.α. Το αριθμητικό αποτέλεσμα συναισθήματος για ένα κείμενο με χρήση του Semantria έχει ένα εύρος τιμών από -2 έως +2, όπου για τιμή -2 το σχόλιο θεωρείται πολύ αρνητικό, για -1 αρνητικό, για 1 θετικό και για 2 πολύ θετικό. Πιο συγκεκριμένα θεωρείται αρνητικό για τιμές μικρότερες του -0.05 ( $-0.05$ ) και θετικό για τιμές μεγαλύτερες από 0.22 ( $>0.22$ ). Για τις ενδιάμεσες τιμές θεωρείται ουδέτερο  $-0.05 < S < 0.22$ .

Τα συγκεκριμένα αριθμητικά αποτελέσματα συναισθήματος στηρίζονται σε ένα προ-διαμορφωμένο λεξικό φράσεων που είναι αναλυτικό και χρησιμοποιείται ευρέως σε πλήθος διαφορετικών τομέων. Οι φράσεις που ενσωματώνονται στο λεξικό και έχουν συγκεκριμένο score μπορούν να αποτελούνται από μια έως τρεις λέξεις. Στη περίπτωση που βρίσκονται υπο-φράσεις κερδίζει αυτή που περιλαμβάνει τις περισσότερες λέξεις. (<https://semantria.readme.io/docs/quick-start-with-python>)

Προκειμένου να προκύψει το αριθμητικό αποτέλεσμα συναισθήματος (Sentiment Score) λαμβάνονται αρκετές παράμετροι υπόψη. Κάποιες από αυτές τις παραμέτρους είναι οι εξής:

1. Όροι και η συχνότητα τους μέσα στο κείμενο. Σε κάποιες περιπτώσεις λαμβάνεται υπόψη και η θέση των συγκεκριμένων όρων μέσα στο κείμενο.
2. Προσφωνήσεις σαν μέρη του λόγου. Έχει αποδειχθεί ότι τα επίθετα αποτελούν σημαντικούς δείκτες υποκειμενικότητας και άποψης και για αυτό χρησιμοποιούνται σαν σημαντικά χαρακτηριστικά.

3. Λέξεις και φράσεις που χρησιμοποιούνται για έκφραση άποψης. Τέτοιου είδους λέξεις χρησιμοποιούνται συνήθως για να εκφραστεί ένα συναίσθημα.
4. Όροι άρνησης. Οι συγκεκριμένοι όροι είναι πολύ σημαντικοί καθώς η χρήση τους είναι δυνατόν να αλλάξει ολόκληρο το νόημα ενός κειμένου.

### Χρήση Sentiment Analysis

Στο παρόν μεθοδολογικό πλαίσιο τα δεδομένα που αποτελούν είσοδο της εφαρμογής ανάλυσης συναισθήματος είναι δομημένα σε παρόμοια μορφή με αυτή του πίνακα (3-5).

Πίνακας 3-5 Είσοδοι για την εφαρμογή ανάλυσης συναισθήματος

Χρήστης	Εναλλακτική	Σχόλιο
$U_1$	$P_1$	$review_{11}$
$U_2$	$P_1$	$review_{12}$
$U_3$	$P_2$	$review_{23}$
$U_4$	$P_2$	$review_{24}$
$U_5$	$P_2$	$review_{25}$
$U_6$	$P_3$	$review_{36}$
$U_g$	$P_i$	$review_{ig}$

Η έξοδος του συστήματος όπως έχει περιγραφεί παραπάνω είναι το sentiment score  $Se_{ig}$  που κυμαίνεται από -2 έως 2 και είναι ενδεικτικό του συναισθήματος του κάθε χρήστη  $U_q$  για την εναλλακτική  $P_i$ . Παρακάτω στον πίνακα 3-6, εμφανίζονται οι έξοδοι που προκύπτουν από την εφαρμογή της ανάλυσης συναισθήματος, προκειμένου να τις χρησιμοποιήσουμε για τα επόμενα στάδια της σύστασης.

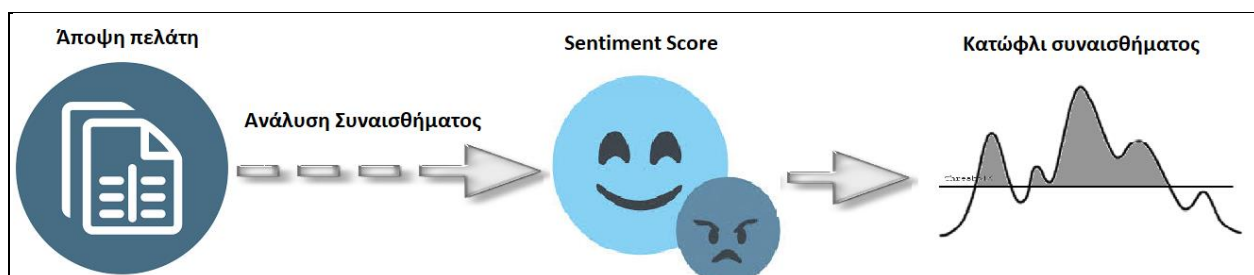
Πίνακας 3-6 Έξοδοι από την ανάλυση συναισθήματος

Χρήστης	Εναλλακτική	Sentiment Score	Average Sentiment Score
$U_1$	$P_1$	$Se_{11}$	$\overline{Se}_1$
$U_2$	$P_1$	$Se_{12}$	
$U_3$	$P_2$	$Se_{23}$	$\overline{Se}_2$
$U_4$	$P_2$	$Se_{24}$	

$U_5$	$P_2$	$Se_{25}$	
$U_6$	$P_3$	$Se_{36}$	$\bar{Se}_3$
$U_g$	$P_i$	$Se_{ig}$	$\bar{Se}_i$

Από τα αποτελέσματα συναισθηματικής ανάλυσης γίνεται υπολογισμός της μέσης τιμής του Sentiment Score για κάθε διαφορετική εναλλακτική και το αποτέλεσμα χρησιμοποιείται σαν κατώφλι σύστασης της συγκεκριμένης εναλλακτικής στο χρήστη του συστήματος.

Δηλαδή για κάθε Εναλλακτική  $P_i$  υπολογίζεται το αντίστοιχο μέσο  $\bar{Se}_i$  που αποτελεί τη μέση τιμή όλων των Sentiment Scores των σχολίων των χρηστών κάθε εναλλακτικής. Στη περίπτωση που το μέσο συναίσθημα που προκύπτει για μια εναλλακτική είναι θετικό, το σύστημα επιτρέπει στην εναλλακτική να αποτελέσει μέρος του συνόλου των εναλλακτικών, που πιθανόν να αποτελέσουν τελική σύσταση στο χρήστη, στην περίπτωση ταύτισης προτιμησιακών μοντέλων, όπως θα περιγραφεί εκτενώς στην επόμενη παράγραφο. Σε αντίθετη περίπτωση η συγκεκριμένη εναλλακτική θα κοπεί από το κατώφλι και δεν θα μπορεί να γίνει σύσταση της στο χρήστη. Παρακάτω στο σχήμα 3-11 παρουσιάζεται απλοποιημένο σχηματικό διάγραμμα της διαδικασίας ανάλυσης συναισθήματος.



Σχ. 3-11 Σχηματικό διάγραμμα εισόδων-εξόδων ανάλυσης συναισθήματος στην μεθοδολογία

### 3.5 Εφαρμογή Μοντελοποίησης Προτίμησης

Στη συγκεκριμένη φάση υλοποίησης της μεθοδολογίας γίνεται χρήση περιορισμένης πληροφορίας προκειμένου να προκύψουν αναλυτικά μοντέλα

προτιμήσεων για τον εκάστοτε χρήστη του συστήματος. Στόχος είναι να γίνει εφικτή η ταύτιση του με το προτιμησιακό μοντέλο των χρηστών που έχουν επιλέξει μια συγκεκριμένη εναλλακτική, σύμφωνα με τα αποτελέσματα της ανάλυσης ικανοποίησης πελατών και να προκύψει σύσταση συγκεκριμένων εναλλακτικών σε αυτόν.

Προκειμένου να υπολογιστεί το προτιμησιακό μοντέλο του χρήστη και να αποφευχθεί το cold start, είναι απαραίτητη η εισαγωγή ελάχιστης πληροφορίας από τον ίδιο. Για να είναι φιλικό το σύστημα προς το χρήστη, θα πρέπει η πληροφορία που του ζητείται να είναι όσο το δυνατόν πιο περιορισμένη σε μέγεθος και ταυτόχρονα κατανοητή. Η μέθοδος που χρησιμοποιείται και καλύπτει τους προαναφερθέντες περιορισμούς είναι μια παραλλαγή της μεθόδου WAP (Spyridakos et al., 2015; Tsotsolas et al., 2016b). Η συγκεκριμένη μέθοδος στηρίζεται σε παραλλαγές της μεθόδου SIMOS (Roy, 1990), (José Figueira and Roy, 2002; Simos, 1990a, 1990b) και καταλήγει σε υπολογισμό ευρών τιμών των βαρυτήτων των προεπιλεγμένων κριτηρίων.

#### Χρήση της WAP στο παρόν μεθοδολογικό πλαίσιο

Η μέθοδος WAP χρησιμοποιείται όπως περιγράφηκε παραπάνω με μια μικρή παραλλαγή στην παρούσα μεθοδολογία, προκειμένου να γίνει το σύστημα ακόμα πιο φιλικό προς το χρήστη χωρίς ταυτόχρονα να αναιρείται η εγκυρότητα και ποιότητα των αποτελεσμάτων της. Πιο συγκεκριμένα διαφοροποιείται το βήμα 3 υλοποίησης της WAP που περιγράφηκε παραπάνω, και αφορά τον υπολογισμό και οριοθέτηση των τιμών  $[zmin_r, zmax_r]$ .

Στην κλασική WAP οι τιμές των  $zmin$  και  $zmax$  προσδιορίζονται απευθείας από το χρήστη για κάθε ζεύγος διαδοχικών κριτηρίων. Στην παρούσα προσέγγιση ο χρήστης καλείται να επιλέξει μια λεκτική έκφραση που προσδιορίζει, για κάθε ζεύγος διαδοχικών κριτηρίων, την μεταξύ τους απόσταση. Για να επιτευχθεί το συγκεκριμένο αποτέλεσμα έχει ουσιαστικά προηγηθεί αντιστοίχιση διαφορετικών ευρών τιμών  $[zmin_r, zmax_r]$  σε λέξεις ή φράσεις που κατατοπίζουν ευκολότερα ένα απλό χρήστη για την ουσία της συγκεκριμένης επιλογής του. Προκειμένου να καταλήξουμε στην αντιστοίχιση ευρών τιμών των

z με λεκτικές περιγραφές διαφοράς σημαντικότητας, δημιουργήσαμε ερωτηματολόγιο που απευθύνουμε σε 50 χρήστες, οι οποίοι κλήθηκαν να δείξουν παραστατικά με χρήση των ράβδων κύλισης, πώς νοούν τις διαφορές σημαντικότητας για τις εξής διαφορές προτίμησης: Πολύ Ασθενής Προτίμηση, Ασθενής Προτίμηση, Προτίμηση, Ισχυρή Προτίμηση, Πολύ Ισχυρή Προτίμηση και απεικονίζουν την προτιμησιακή διαφορά που έχουν τα διαδοχικά κριτήρια μεταξύ τους, στην κατάταξη του χρήστη.

Επιπρόσθετα έχει πραγματοποιηθεί και μια τροποποίηση στο βήμα 2 της WAP, καθώς στην παρούσα προσέγγιση, δεν δίνεται η δυνατότητα στο χρήστη να ορίσει κατά την κατάταξη των κριτηρίων, δύο κριτήρια να είναι ίσα. Όλα τα κριτήρια θα πρέπει να έχουν μια διαφορετική σειρά κατάταξης. Στη περίπτωση που ο χρήστης τα θεωρεί ίσα, του δίνεται η δυνατότητα, στο βήμα 3, να τα χαρακτηρίσει ίσα με χρήση των λεκτικών συγκρίσεων που περιγράφηκαν στην προηγούμενη παράγραφο.

Η είσοδος στη WAP είναι δεδομένα που δίνει ο χρήστης και είναι τα εξής:

1. Κατάταξη των κριτηρίων.
2. Διαφορά σημαντικότητας μεταξύ διαδοχικών κριτηρίων με επιλογή λεκτικής περιγραφής.

Έξοδο της συγκεκριμένης φάσης της μεθοδολογίας αποτελούν τα εύρη βαρών για κάθε κριτήριο για ένα συγκεκριμένο χρήστη καθώς και η ευρωστία  $\mu$  και ο μέσος δείκτης ευστάθειας ASI. Πιο συγκεκριμένα αν  $Cr_x$ ,  $x=1... \#Criteria$  τα κριτήρια που έχουν καταταχθεί από το χρήστη  $U_f$ , κατόπιν εφαρμογής της WAP προκύπτουν τα εξής αποτελέσματα:  $MinWeightCr_x$ ,  $MidWeightCr_x$ ,  $MaxWeightCr_x$ . Στο πίνακα 3-7 παρουσιάζονται οι έξοδοι που προκύπτουν από την εφαρμογή της WAP, και στο σχήμα 3-12 αντίστοιχα ένα απλοποιημένο σχηματικό διάγραμμα της λειτουργικότητας της.

Πίνακας 3-7 Έξοδοι εφαρμογής WAP

Χρήστης	Ελάχιστο βάρος x Κριτηρίου	Μεσαίο Βάρος x Κριτηρίου	Μέγιστο Βάρος x Κριτηρίου	ASI	$\mu$
$U_f$	$MinWeightCr_x$	$MidWeightCr_x$	$MaxWeightCr_x$	$ASI_f$	$\mu_f$





Σχ. 3-12 Σχηματικό διάγραμμα εισόδων-εξόδων WAP στη μεθοδολογία

### 3.6 Τελική Σύσταση

Η τελική σύσταση εναλλακτικής/ων προς το χρήστη προκύπτει από το συνδυασμό των αποτελεσμάτων της εφαρμογής των προαναφερθέντων μεθοδολογιών στα δεδομένα. Παρακάτω αναφέρονται συνοπτικά τα αποτελέσματα που έχουν προκύψει από κάθε φάση της μεθοδολογίας και εν συνεχεία ο τρόπος συνδυασμού τους για την παρουσίαση της τελικής σύστασης.

#### Αποτελέσματα άντλησης και επεξεργασίας

Από την φάση άντλησης και επεξεργασίας των δεδομένων προκύπτουν τα αποτελέσματα, όπως παρουσιάζονται στον πίνακα 3-8:

Πίνακας 3-8 Αποτελέσματα φάσης άντλησης και επεξεργασίας δεδομένων

Χρήστης	Εναλλακτική	Συνολική Βαθμολογία	Βαθμολογία $j^{ou}$ Κριτηρίου	Σχόλιο
$U_g$	$P_i$	$O_{ig}$	$\Gamma_{jig}$	$review_{ig}$

Πιο περιγραφικά τα αποτελέσματα περιλαμβάνουν τα εξής στοιχεία:

1.  $U_g$ , το όνομα χρήστη που είχε κάνει αξιολόγηση.
2.  $P_i$ , τα προϊόντα/εναλλακτικές που έχει αξιολογήσει ο χρήστης  $U_g$ .
3.  $O_{ig}$ , η συνολική βαθμολογία προτίμησης (overall score) που έχει δώσει ο χρήστης  $U_g$  για την εναλλακτική  $P_i$ .

4.  $r_{jig}$ , η βαθμολογία, που έχει δώσει ο χρήστης  $U_g$  για κάθε κριτήριο  $c_j$ ,
5.  $review_{ig}$ , το σχόλιο του χρήστη  $U_g$  για την εναλλακτική  $P_i$ .

### Αποτελέσματα MUSA

Από τη φάση εφαρμογής της Πολυκριτήριας Ανάλυσης Ικανοποίησης Πελατών προκύπτουν στα δεδομένα βαθμολογιών (Ratings) της πρώτης φάσης προκύπτουν τα παρακάτω αποτελέσματα, στη μορφή που παρουσιάζονται στον πίνακα 3-9:

Πίνακας 3-9 Αποτελέσματα MUSA

Εναλλακτική	Δείκτης Συνολικής Ικανοποίησης	Βάρος $j$ Κριτηρίου
$P_i$	$S_i$	$b_{ji}$

- όπου:
1.  $P_i$ , Τα προϊόντα/εναλλακτικές.
  2.  $S_i$ , Ο δείκτης συνολικής Ικανοποίησης για την εναλλακτική  $P_i$ .
  3.  $b_{ji}$ , Το βάρος του  $j$  κριτηρίου, με  $\sum b_j=1$  για κάθε εναλλακτική  $P_i$ .

### Αποτελέσματα Ανάλυσης συναισθήματος

Η φάση Ανάλυσης Συναισθήματος στα σχόλια (reviews) των χρηστών οδηγεί στα ακόλουθα αποτελέσματα, όπως εμφανίζονται στον πίνακα 3-10:

Πίνακας 3-10 Αποτελέσματα ανάλυσης συναισθήματος

Χρήστης	Εναλλακτική	Sentiment Score	Average Sentiment Score
$U_g$	$P_i$	$Se_{ig}$	$\bar{Se}_i$

Όπου :

1.  $U_g$ , το όνομα του χρήστη που είχε κάνει αξιολόγηση
2.  $P_i$ , τα προϊόντα/εναλλακτικές
3.  $Se_{ig}$ , το sentiment score που κυμαίνεται από -2 έως 2

4.  $\bar{S}_i$ , το μέσο sentiment score για κάθε εναλλακτική  $P_i$

#### Αποτελέσματα μοντελοποίησης προτίμησης

Η φάση μοντελοποίησης προτίμησης WAP, κατόπιν των επιλογών/εισόδων του χρήστη  $U_f$  στο σύστημα συστάσεων καταλήγει στην ακόλουθη μορφή αποτελεσμάτων, όπως εμφανίζονται στον πίνακα 3-11:

Πίνακας 3-11 Αποτελέσματα μοντελοποίησης προτίμησης

Χρήστης	Ελάχιστο βάρος $x$ Κριτηρίου	Μεσαίο Βάρος $x$ Κριτηρίου	Μέγιστο Βάρος $x$ Κριτηρίου	ASI	$\mu$
$U_f$	$\text{MinWeightCr}_x$	$\text{MidWeightCr}_x$	$\text{MaxWeightCr}_x$	$\text{ASI}_f$	$\mu_f$

Όπου:

1.  $U_f$ , ο νέος χρήστης του συστήματος.
2.  $\text{MinWeightCr}_x$ , η ελάχιστη τιμή του βάρους του κριτηρίου  $x$ .
3.  $\text{MidWeightCr}_x$ , η μεσαία τιμή του βάρους του κριτηρίου  $x$ .
4.  $\text{MaxWeightCr}_x$ , ο μέγιστη τιμή του βάρους του κριτηρίου  $x$ .
5.  $\text{ASI}_f$ , ο μέσος δείκτης ευστάθειας (Average Stability Index).
6.  $\mu_f$ , ο δείκτης ευρωστίας.

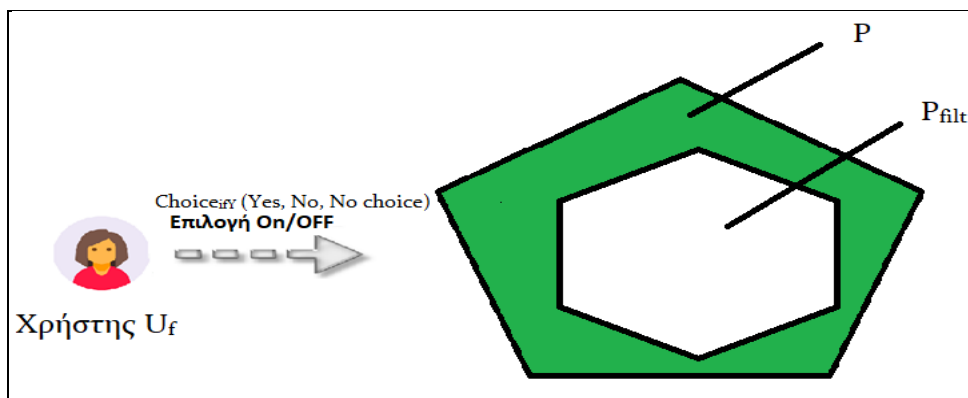
#### Διαδικασία Σύστασης

Παρακάτω παρατίθενται τα βήματα που πραγματοποιούνται για την τελική διαμόρφωση της σύστασης. Για την περιγραφή χρησιμοποιούνται τα σύμβολα που έχουν αναφερθεί προηγουμένως κατά την ανάλυση των φάσεων ξεχωριστά.

1. Ο χρήστης  $U_f$  επιλέγει  $\text{Choice}_{ifY}$  (Yes, No, No choice) από τα πιθανά στατικά χαρακτηριστικά των εναλλακτικών  $P_i$ , αυτά που θεωρεί αναγκαία να υπάρχουν στις τελικές επιλογές που θα του συσταθούν. Το αποτέλεσμα των επιλογών του χρήστη είναι η δημιουργία ενός υποσυνόλου των εναλλακτικών  $P$ , που το ονομάζουμε  $P_{\text{filt}}$  και ισχύει  $P_{\text{filt}} \subseteq P$ . Στον πίνακα 3-12 παρουσιάζεται η μορφή των αποτελεσμάτων, ενώ στο σχήμα 3-13 οπτικοποιούμε τα σύνολα των εναλλακτικών  $P_{\text{filt}}$ ,  $P$ .

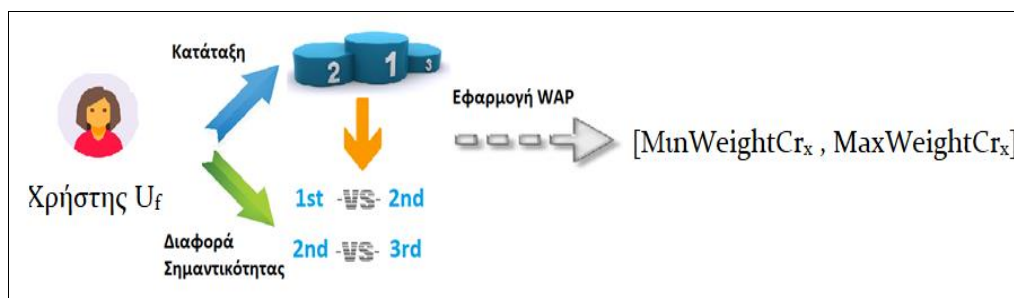
Πίνακας 3-12 Επιλογές On/Off από το χρήστη

Χρήστης	Εναλλακτική	Χαρακτηριστικό Y
$U_f$	$P_i$	Choice <sub>ifY</sub>



Σχ. 3-13 Σύνολα εναλλακτικών  $P$ ,  $P_{filt}$  μετά την επιλογή On/Off από το χρήστη

- Ο χρήστης  $U_f$  καλείται να κατατάξει τα προκαθορισμένα ποιοτικά κριτήρια και εν συνεχεία να επιλέξει την απόσταση μεταξύ των διαδοχικών κριτηρίων με επιλογή της λεκτικής περιγραφής που κατά την άποψη του δείχνει της διαφορά των διαδοχικών κριτηρίων. Οι επιλογές του χρήστη  $U_f$  αποτελούν τις εισόδους για τη μέθοδο WAP. Ως αποτέλεσμα μας προκύπτουν τα εύρη των βαρών για καθένα από τα προκαθορισμένα ποιοτικά κριτήρια. Δηλαδή για κάθε Κριτήριο  $x$  που έχει κατατάξει ο χρήστης προκύπτει ένα εύρος  $[MinWeightCr_x, MaxWeightCr_x]$ . Τα εύρη βαρών του συνόλου των ποιοτικών κριτηρίων καθορίζουν το προτιμησιακό προφίλ του χρήστη  $U_f$ . Στο σχήμα 3-14 παρουσιάζεται παραστατικά η διαδικασία της WAP, με τις εισόδους από το χρήστη και τα αποτελέσματα.



Σχ. 3-14 Διαδικασία και αποτελέσματα WAP στη μεθοδολογία

3. Γίνεται σύγκριση για καθένα από τα  $j$  ποιοτικά κριτήρια  $x=1,2,...,j$ , αν το βάρος  $b_{xi}$  του  $x$  κριτηρίου της εναλλακτικής  $P_i$ , ανήκει στο εύρος  $[MinWeightCr_x, MaxWeightCr_x]$ . Να σημειωθεί ότι το  $b_{xi}$  είναι το βάρος που έχει προκύψει από τη MUSA στη δεύτερη Φάση. Ορίζουμε μεταβλητή  $Comp_{xi}$  του  $x$  κριτηρίου και της εναλλακτικής  $P_i$ , η οποία στη περίπτωση που το  $b_{xi}$  ανήκει στο διάστημα  $[MinWeightCr_x, MaxWeightCr_x]$ , παίρνει την τιμή 1 αλλιώς 0. Εν συνεχεία αθροίζουμε το γινόμενο  $b_{xi} * Comp_{xi}$  και το αποτέλεσμα το αποθηκεύουμε στην μεταβλητή  $WeSum_i$ , η οποία αντιστοιχίζεται σε κάθε μια εναλλακτική  $P_i$ .

α. Αν  $b_{xi} \in [MinWeightCr_x, MaxWeightCr_x]$  (3-1)

τότε:  $Comp_{xi}=1$  αλλιώς  $Comp_{xi}=0$  (3-2)

β.  $WeSum_i = \sum (b_{xi} * Comp_{xi})$  (3-3)

ή

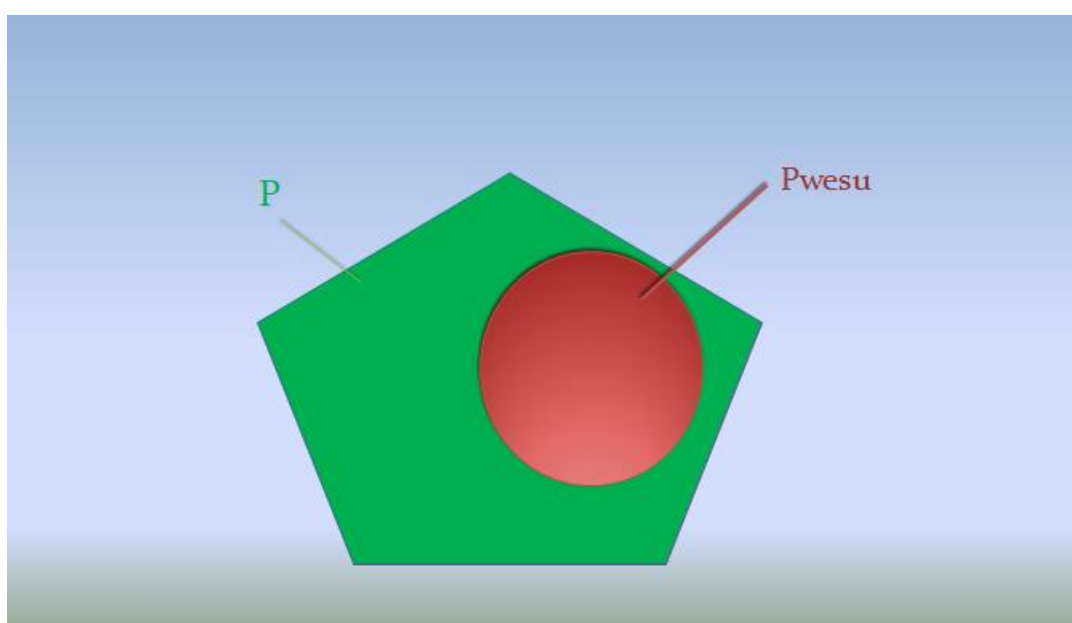
$WeSum_i = \sum (Comp_{xi})$  (3-4)

για το μοντέλο σύστασης με single sum, για κάθε κριτήριο  $x$  και εναλλακτική  $P_i$ .

Στο συγκεκριμένο σημείο γίνεται σύγκριση της ομοιότητας του προτιμησιακού μοντέλου του χρήστη  $U_f$  με τα προτιμησιακά μοντέλα των χρηστών  $U_g$  που έχουν δοκιμάσει και αξιολογήσει τις εναλλακτικές  $P_i$ .

4. Στα δεδομένα της μεταβλητής  $WeSum_i$ , που ουσιαστικά έχουν τη μορφή στήλης, εφαρμόζεται κατάταξη (ranking), ώστε οι εναλλακτικές που αντιστοιχούν στις μεγαλύτερες τιμές  $WeSum_i$  να βρεθούν πρώτες στην

κατάταξη, και να δημιουργηθεί λίστα προτεραιότητας εναλλακτικών, προς σύσταση στο χρήστη. Επιπρόσθετα, στην περίπτωση που η σειρά προτεραιότητας των εναλλακτικών δείχνει να ταυτίζεται, προτείνεται από τη μεθοδολογία η περαιτέρω κατάταξη των συγκεκριμένων εναλλακτικών, με βάση τη μέση συνολική αξιολόγηση για κάθε εναλλακτική. Το σύνολο εναλλακτικών που προκύπτει ονομάζεται  $P_{WESU}$  και αποτελεί υποσύνολο του  $P$ ,  $P_{WESU} \subseteq P$ . Στο σχήμα 3-15 παρουσιάζονται τα σύνολα  $P$ ,  $P_{WESU}$ .

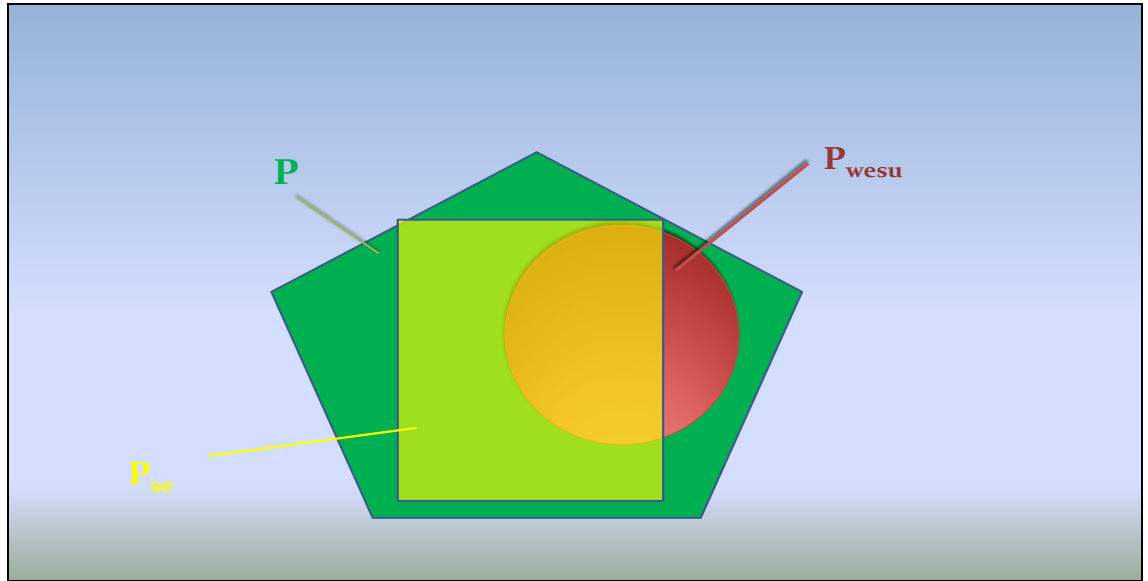


Σχ. 3-15 Σύνολα εναλλακτικών  $P$ ,  $P_{WESU}$  μετά την εφαρμογή WAP

5. Εφαρμόζεται Ανάλυση Συναισθήματος στο υποσύνολο  $P_{WESU}$ . Στόχος της εφαρμογής της συγκεκριμένης ανάλυσης στο παρόν στάδιο της μεθοδολογίας, είναι να εξασφαλισθεί, πέραν της ταύτισης του προτιμησιακού μοντέλου, δηλαδή του τρόπου λήψης απόφασης, ότι, οι εναλλακτικές που θα συσταθούν στο χρήστη, έχουν ικανοποιήσει τους χρήστες που τις έχουν επιλέξει στο παρελθόν. Συνεπώς με βάση τα αποτελέσματα ανάλυσης συναισθήματος στα σχόλια χρηστών, αφαιρούνται από το υποσύνολο  $P_{WESU}$  οι εναλλακτικές που δεν έχουν προκαλέσει θετικά συναισθήματα στους χρήστες. Συνεπώς η ανάλυση συναισθήματος χρησιμοποιείται στην παρούσα μεθοδολογία ως ένα

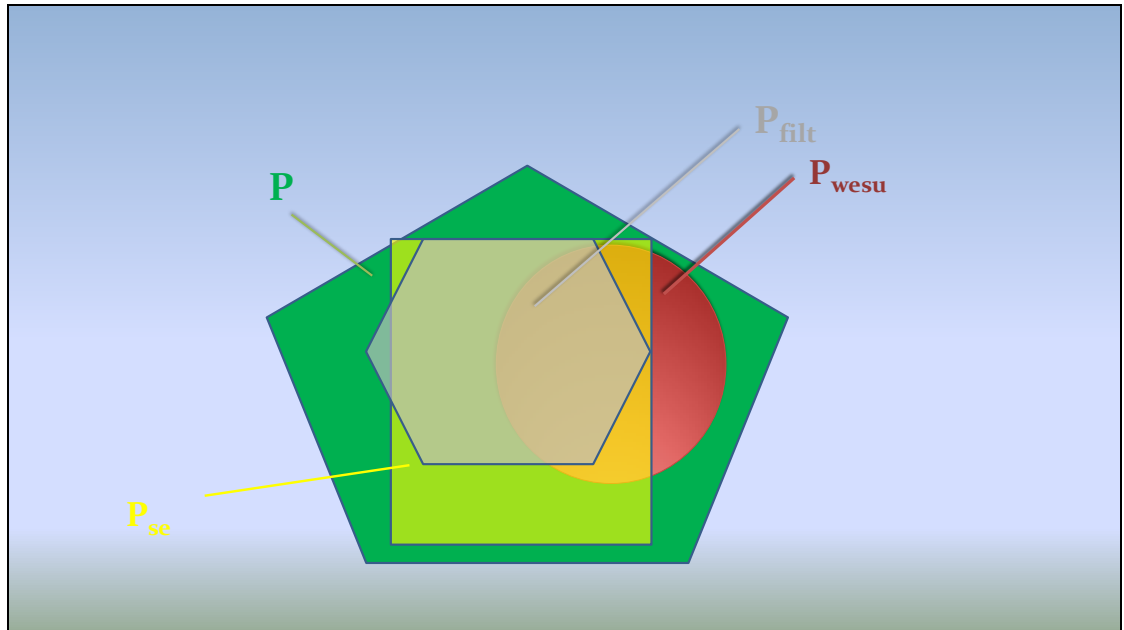
κατώφλι ικανοποίησης χρηστών, προκειμένου να βελτιστοποιηθεί η ποιότητα των συστάσεων που θα γίνουν προς το χρήστη. Το νέο υποσύνολο εναλλακτικών ονομάζεται  $P_{wesuse}$ ,

$P_{wesuse} \subseteq P_{wesu} \subseteq P$ . Στο σχήμα 3-16 παρουσιάζονται τα σύνολα  $P$ ,  $P_{wesu}$ ,  $P_{se}$ , ενώ είναι ορατό ότι το μέρος του κυκλικού συνόλου  $P_{wesu}$ , που βρίσκεται εντός του  $P_{se}$ , αποτελεί το  $P_{wesuse}$ .

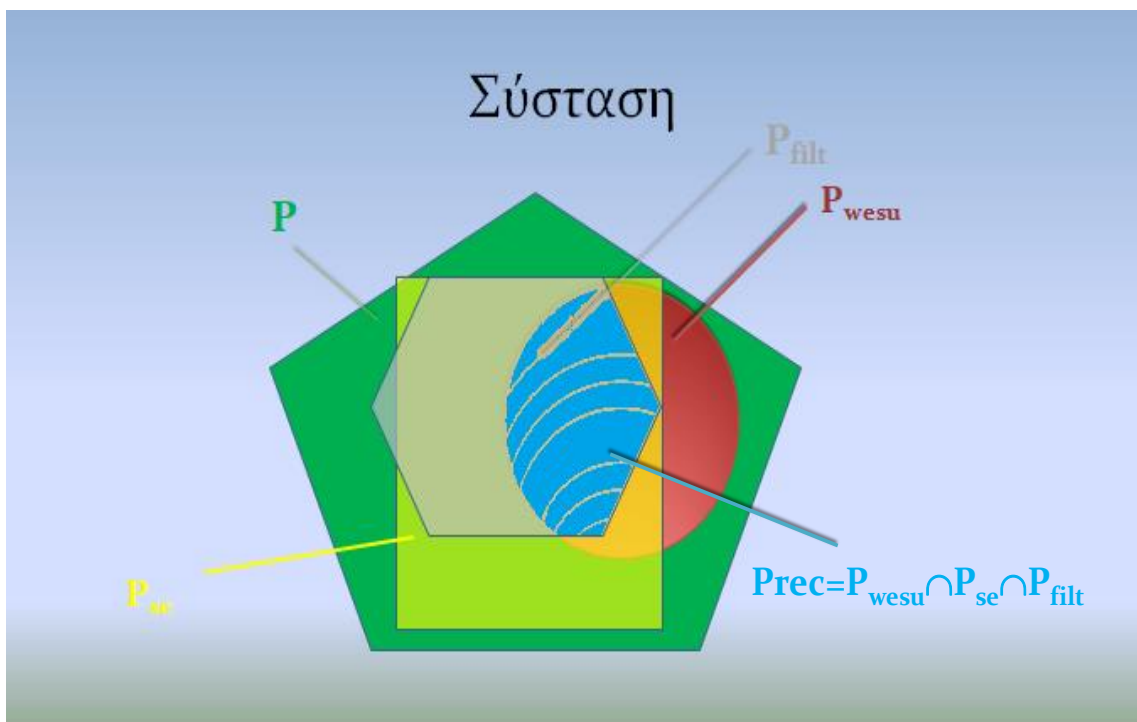


Σχ. 3-16 Σύνολα και υποσύνολα εναλλακτικών  $P$ ,  $P_{wesu}$ ,  $P_{se}$

6. Η τελική σύσταση προς το χρήστη  $U_f$ , προκύπτει από την τομή των υποσυνόλων των εναλλακτικών  $P_{wesuse}$  και  $P_{filt}$  (βλ. σχήμα 3-17). Το υποσύνολο των εναλλακτικών που αποτελεί την τελική σύσταση ονομάζεται  $P_{rec}$  και ισχύει ότι:  $P_{rec} = P_{wesuse} \cap P_{filt}$  (Σχήμα 3-18)



Σχ. 3-17 Σύνολα και υποσύνολα εναλλακτικών  $P$ ,  $P_{wesu}$ ,  $P_{se}$ ,  $P_{filt}$



Σχ. 3-18 Υποσύνολο συστάσεων  $P_{prec}$  και σύνολα εναλλακτικών  $P$ ,  $P_{wesu}$ ,  $P_{weseuse}$ ,  $P_{filt}$

Είναι απαραίτητο να σημειωθεί ότι στο σύστημα παρουσιάζονται 4 διαφορετικά μοντέλα συστάσεων. Παρακάτω παρατίθεται για καθένα από τα μοντέλα ο τρόπος υπολογισμού της τελικής σύστασης.



A. Σύσταση με βάση το προτιμησιακό μοντέλο και απλό άθροισμα (Recommendation Based on Preference Model with Single Sum)

$$P_{rec} = P_{wesu} \text{ όπου } WeSum_i = \sum(Comp_{xi}) \quad (3-5)$$

B. Σύσταση με βάση το προτιμησιακό μοντέλο και το σταθμισμένο άθροισμα (Recommendation Based on Preference model and Weighted Sum)

$$P_{rec} = P_{wesu} \text{ όπου } WeSum_i = \sum(b_{xi} * Comp_{xi}) \quad (3-6)$$

Γ. Σύσταση με βάση το Προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα και το κατώφλι συναισθήματος (Recommendation Based on Preference model, Weighted Sum and Sentiment Threshold)

$$P_{rec} = P_{wesuse} \quad (3-7)$$

Δ. Σύσταση με βάση το Προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα, το κατώφλι συναισθήματος και το μοντέλο φιλτραρίσματος (Recommendation Based on Preference Model, Weighted Sum, Sentiment Threshold and Filtering Model)

$$P_{rec} = P_{wesuse} \cap P_{filt} \quad (3-8)$$

### 3.7 Μηχανισμός Ανατροφοδότησης

Ο μηχανισμός ανατροφοδότησης δημιουργείται για δύο κύριους λόγους:

1. Δημιουργία βάσης γνώσης της λειτουργίας του συστήματος για περαιτέρω ανάλυση.
2. Βελτιστοποίηση λειτουργικότητας για επόμενη χρήση του συστήματος από το χρήστη.

Για κάθε νέα εγγραφή ή είσοδο χρήστη στο σύστημα σύστασης αποθηκεύονται στη βάση δεδομένα με τη μορφή νέας εγγραφής στη βάση.

Τα στοιχεία που καταχωρούνται για κάθε νέα εγγραφή είναι τα εξής:

- ⇒ Στοιχεία Χρήστη (Όνομα, Επώνυμο, Username, Login type).
- ⇒ Επιλογές του χρήστη στα On/Off κριτήρια.
- ⇒ Επιλογές του Χρήστη στην κατάταξη των κριτηρίων.

- ⇒ Επιλογές του χρήστη για τη διαφορά σημαντικότητας των διαδοχικών κριτηρίων.
- ⇒ Τα αποτελέσματα της WAP.
- ⇒ Τα αποτελέσματα των συστάσεων για το χρήστη.
- ⇒ Η ανατροφοδότηση του χρήστη προς το σύστημα.

Στις παρακάτω εικόνες (Σχήματα 3.19, 3.20) παρουσιάζονται πιο αναλυτικά τα δεδομένα που αποτελούν κάθε εγγραφή.

### Εγγραφή στη βάση δεδομένων

Στοιχεία Χρήστη			
Όνομα	Επώνυμο	Username	Login Type

Επιλογές του χρήστη στα ON/Off κριτήρια			
Χαρακτηριστικό #1	Χαρακτηριστικό #2	Χαρακτηριστικό #N	# of Filtered Hotels
π.χ No choice	π.χ Yes	π.χ No	π.χ 106

Επιλογές του Χρήστη στην κατάταξη κριτηρίων				
Criterion #1	Criterion #2	Criterion #3	Criterion #4	Criterion #5
π.χ 2 <sup>nd</sup>	π.χ. 1 <sup>st</sup>	π.χ 4 <sup>th</sup>	π.χ 3 <sup>rd</sup>	π.χ. 5 <sup>th</sup>

Επιλογές του Χρήστη για διαφορά σημαντικότητας διαδοχικών κριτηρίων			
1 <sup>st</sup> vs 2 <sup>nd</sup>	2 <sup>nd</sup> vs 3 <sup>rd</sup>	3 <sup>rd</sup> vs 4 <sup>th</sup>	4 <sup>th</sup> vs 5 <sup>th</sup>
π.χ Weak preference	π.χ Very Strong preference	π.χ Preference	π.χ Strong Preference

Σχ. 3-19 Εγγραφές σχετικές με τις επιλογές του χρήστη συστήματος συστάσεων

Αποτελέσματα WAP								
$C_{r,min}$ Weight	$C_{r,mid}$ Weight	$C_{r,max}$ Weight	$C_{r,min}$ Weight	$C_{r,mid}$ Weight	$C_{r,max}$ Weight	$C_{r,min}$ Weight	$C_{r,mid}$ Weight	$C_{r,max}$ Weight
$\pi_X: 0.33$	$\pi_X: 0.37$	$\pi_X: 0.40$	$\pi_X: 0.34$	$\pi_X: 0.36$	$\pi_X: 0.38$	$\pi_X: 0.3$	$\pi_X: 0.35$	$\pi_X: 0.35$

Αποτελέσματα συστάσεων	
1	Recommendation 1
2	Recommendation 2
3	Recommendation 3
m	Recommendation m

Αναπροϋφότιση Χρήστη		
Επιλογή Προϊόντος	Ευχαριστήριους/η με αίτημα/αίτηση	Σημείο
π.χ. Αλλεργία/α	π.χ. Ναι/Όχι	Π.χ. Good

### Σχ. 3-20 Αποτελέσματα συστάσεων και ανατροφοδότηση

Στην παρακάτω εικόνα (Σχήμα 3.21) εμφανίζεται ένα στιγμιότυπο εγγραφής στη βάση δεδομένων για χρήση του συστήματος.

Στοιχεία Χρήστη				Επιλογές του Χρήστη στην κατάσταση των κριτηρίων					Επιλογές του χρήστη στα On/Off κριτήρια																													
Name	Surname	Username	Login type	Cleanliness	Service	Comfort	Condition	Neighbourhood	Star	Outdoor pool	Indoor pool	Massage	T.V.	Restaurant	Bar	Fitness	HOT	Heating	Air Conditioning	Family room	Facilities for disabled guests	Apartment furnished well	Safety deposit box	Free Parking	Room service	Non-smoking throughout	Non-smoking rooms	Lift	English	German	Russian	French	Spanish	Poles	#Free hotels			
Yannis	Zis	giannis	Admin	2nd	4th	1st	5th	3rd	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	Yes	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	No Choice	18
Επιλογές του χρήστη για τη διαφορά σημαντικότητας των διαδοχικών κριτηρίων									Τα αποτελέσματα της WAP										Τα αποτελέσματα των αντιστάσεων για το χρήστη										Η ανατροφοδότηση του χρήστη προς το σύστημα									
1st vs 2nd	2nd vs 3rd	3rd vs 4th	4th vs 5th	Cleanliness Min Weight	Cleanliness Mid Weight	Cleanliness Max Weight	Service Min Weight	Service Mid Weight	Service Max Weight	Comfort Min Weight	Comfort Mid Weight	Comfort Max Weight	Condition Min Weight	Condition Mid Weight	Condition Max Weight	Neighborhood Min Weight	Neighborhood Mid Weight	Neighborhood Max Weight	HOTELS Recommendations										User Hotel Choice	Pleased with app	Comment							
Weak Preference	Weak Preference	Strong Preference	Strong Preference	0.2181	0.2429	0.2719	0.0527	0.0779	0.0998	0.4419	0.4371	0.5518	0.017	0.0255	0.0353	0.1301	0.1556	0.1857	Areti	Mythos Plataniar	Blue Dome Hotel	Sono Omega Platania	Beach Apartment	Grand Bay Beach Resort	Blue Dome Hotel	Yes	Good job. Pretty easy and fast on the recommendation											

**Σχ. 3-21 Στιγμιότυπο εγγραφής στη βάση δεδομένων**

Η μηχανισμός ανατροφοδότησης αποτελεί ένα προαιρετικό βήμα για το χρήστη  $U_f$  του συστήματος, όπου του ζητείται να εισάγει την τελική του απόφαση σχετικά με την επιλογή εναλλακτικής. Επίσης ο χρήστης καλείται να βαθμολογήσει το σύστημα συστάσεων καθώς και να εκφράσει τις σχετικές παρατηρήσεις του. Η βαθμολογία είναι σε κλίμακα Ναι/Όχι και δίνεται έμφαση στην ανάλυση του σχολίου του χρήστη, με ανάλυση κειμένου και συναισθήματος, ώστε να αναλυθούν σε βάθος οποιεσδήποτε παρατηρήσεις του,

και να εξαχθεί το συναίσθημα του σε σχέση με το σύστημα. Με βάση την ανάλυση των σχολίων αναμένεται να γίνουν βελτιώσεις στο σύστημα.

Κατόπιν της εισαγωγής του username του, του προσφέρεται άμεση ενημέρωση για όλες τις επιλογές του και αποφάσεις του σε προηγούμενες επισκέψεις. Με τη δημιουργία του μηχανισμού ανατροφοδότησης επιτυγχάνονται τα εξής αποτελέσματα:

1. Γίνεται εφικτή η αξιολόγηση της ποιότητας των συστάσεων που προσφέρονται σε κάθε χρήστη, με μετρήσιμα δεδομένα, καθώς και η επιρροή που ασκεί το σύστημα στην τελική απόφαση του.
2. Επιτυγχάνεται στοχοθετημένη ανάλυση των ζητημάτων που αφορούν τη βελτίωση λειτουργίας του συστήματος, με βάση τις απόψεις των χρηστών οι οποίοι σχολιάζουν την εμπειρία τους από τη χρήση του συστήματος.
3. Βελτιώνεται η εμπειρία χρήσης του συστήματος από το χρήστη.
4. Αντλείται πληροφορία βαθμολογιών και κριτικών των νέων χρηστών για τις εναλλακτικές και εμπλουτίζεται η βάση δεδομένων του συστήματος.

### 3.8 Συνοπτική Μαθηματική παρουσίαση της μεθοδολογίας

#### Α. Αποτελέσματα άντλησης και επεξεργασίας

Από την φάση άντλησης και επεξεργασίας των δεδομένων προκύπτουν τα εξής αποτελέσματα :

**Αξιολόγηση χρήστη** =  $\{ U_g, P_i, O_{ig}, r_{jig}, review_{ig} \}$  (3-9)

όπου:  $U_g$ , το όνομα χρήστη.

$P_i$ , τα προϊόντα/εναλλακτικές που έχει αξιολογήσει ο χρήστης  $U_g$ .

$O_{ig}$ , η συνολική βαθμολογία προτίμησης (overall score) που έχει δώσει ο χρήστης  $U_g$  για την εναλλακτική  $P_i$ .

$r_{jig}$ , η βαθμολογία, που έχει δώσει ο χρήστης  $U_g$  για το κριτήριο  $c_j$ ,

$review_{ig}$ , το σχόλιο του χρήστη  $U_g$  για την εναλλακτική  $P_i$ .

### B. Εφαρμογή MUSA

**Είσοδοι MUSA**= {  $U_g$  ,  $P_i$  ,  $O_{ig}$  ,  $r_{jig}$  } (3-10)

όπου:  $U_g$ , το όνομα χρήστη.

$P_i$  , τα προϊόντα/εναλλακτικές που έχει αξιολογήσει ο χρήστης  $U_g$ .

$O_{ig}$ , (overall score) η συνολική βαθμολογία προτίμησης που έχει δώσει ο χρήστης  $U_g$  για την εναλλακτική  $P_i$  .

$r_{jig}$  , η βαθμολογία, που έχει δώσει ο χρήστης  $U_g$  για το κριτήριο  $c_j$ ,

Η βασική εξίσωση της ποιοτικής ανάλυσης παλινδρόμησης έχει την εξής μορφή:

$$\tilde{Y}^* = \sum_{i=1}^n b_i X_i^* - \sigma^+ + \sigma^- \quad (3-5)$$

όπου  $\tilde{Y}^*$  είναι η εκτίμηση της συλλογικής συνάρτησης αξιών  $Y^*$ ,  $\sigma^+$  και  $\sigma^-$  είναι αντίστοιχα το σφάλμα υπερεκτίμησης και υποεκτίμησης. Η  $\tilde{Y}^*$  αντιστοιχεί στο  $O_{ig}$ , (overall score) που είναι η συνολική βαθμολογία που έχει δώσει ο χρήστης και η  $X_i^*$  στο  $r_{jig}$  που είναι η βαθμολογία που έχει δώσει σε κάθε κριτήριο.

Το μέγεθος του προηγούμενου μαθηματικού προγράμματος ελαττώνεται με τη χρήση νέων μεταβλητών, οι οποίες εκφράζουν τα διαδοχικά βήματα αύξησης των συναρτήσεων  $Y^*$  και  $X_i^*$

$$\begin{cases} Z_m = y^{*m+1} - y^{*m}, \text{για } m = 1, 2, \dots, a-1 \\ w_{ik} = b_i x_i^{*k+1} - b_i x_i^{*k} \text{ για } k = 1, 2, \dots, a-1 \text{ και } i = 1, 2, \dots, n \end{cases} \quad (3-6)$$

Οι αρχικές μεταβλητές απόφασης του γραμμικού προγράμματος γράφονται:

$$\begin{cases} y^{*m} = \sum_{i=1}^{m-1} z_i \text{ για } m = 2, 3, \dots, a \\ b_i x_i = \sum_{i=1}^{m-1} z_i \text{ για } k = 2, 3, \dots, a \text{ και } i = 1, 2, \dots, n \end{cases} \quad (3-7)$$

οπότε εισάγοντας τις νέες μεταβλητές  $z_m$  και  $w_{ik}$  και χρησιμοποιώντας τις εξισώσεις η εξίσωση παλινδρόμησης γράφεται ως εξής:

$$\sum_m z_m = \sum_i \sum_k w_{ik} - \sigma^+ + \sigma^- \quad (3-8)$$

Πιο συγκεκριμένα, έστω ότι ο πελάτης  $j$  έχει εκφράσει την ικανοποίησή του με βάση τις καθορισμένες ποιοτικές κλίμακες  $Y$  και  $X_i$  , δηλαδή:

$$\left\{ \begin{array}{l} \text{Ολική Ικανοποίηση } \bar{y}^{*j} = y^{tj} \text{ και } \bar{y}^{*j} \in Y = \{y^1, y^{2j}, \dots, y^{tj}, y^a\} \\ \text{Μερική Ικανοποίηση } \bar{x}^{*j} = x^{tj} \text{ και } \bar{x}^{*j} \in Xi = \{x^1, x^{2j}, \dots, x^{tj}, x^a\} \end{array} \right. \quad (3-9)$$

Τότε θα ισχύει:

$$\sum_{m=1}^{t_j-1} z_m = \sum_{i=1}^n \sum_{k=1}^{t_{ji}-1} w_{ik} - \sigma^+ + \sigma^- \quad \forall j \quad (3-10)$$

Άρα η τελική μορφή του γραμμικού προγράμματος έχει ως εξής:

$$\left\{ \begin{array}{l} [\min] F = \sum_{j=1}^M \sigma_j^+ + \sigma_j^- \\ \text{subject to} \\ \sum_{i=1}^n \sum_{k=1}^{x_i^j-1} w_{ik} - \sum_{m=1}^{y^j-1} z_m - \sigma_j^+ + \sigma_j^- \text{ for } j=1, 2, \dots, M \\ \sum_{m=1}^{\alpha-1} z_m = 100 \\ \sum_{i=1}^n \sum_{k=1}^{\alpha_i-1} w_{ik} = 100 \\ z_m, w_{ik}, \sigma_j^+, \sigma_j^- \geq 0 \quad \forall m, i, k, j \end{array} \right. \quad (3-11)$$

όπου M ο συνολικός αριθμός των πελατών.

Οι αρχικές μεταβλητές του προβλήματος υπολογίζονται με βάση τη βέλτιστη λύση του προηγούμενου γραμμικού προγράμματος, αφού εύκολα αποδεικνύεται ότι:

$$\left\{ \begin{array}{l} y^{*m} = \sum_{t=1}^{m-1} z_t \quad \text{για } m=2, 3, \dots, \alpha \\ b_i = \frac{\sum_{t=1}^{\alpha_i-1} w_{it}}{100} \quad \text{για } i=1, 2, \dots, n \\ x_i^{*k} = 100 \frac{\sum_{t=1}^{k-1} w_{it}}{\sum_{t=1}^{\alpha_i-1} w_{it}} \quad \text{για } i=1, 2, \dots, n \text{ και } k=2, 3, \dots, \alpha_i \end{array} \right. \quad (3-12)$$

Οι μέσοι δείκτες ολικής ικανοποίησης  $S$  και μερικής ικανοποίησης  $S_i$  ορίζονται με βάση τις σχέσεις:

$$\left\{ \begin{array}{l} S = \frac{1}{100} \sum_{m=1}^a p^m y^{*m} \\ S_i = \frac{1}{100} \sum_{k=1}^{a_i} p_i^k x_i^{*k}, i = 1, 2, \dots, n \end{array} \right. \quad (3-13)$$

**Έξοδοι MUSA= { P<sub>i</sub> , S<sub>i</sub> , b<sub>ji</sub> } (3-20)**

όπου: P<sub>i</sub>, τα προϊόντα/εναλλακτικές.

S<sub>i</sub>, ο δείκτης συνολικής Ικανοποίησης για την εναλλακτική P<sub>i</sub>.

b<sub>ji</sub>, το βάρος του j κριτηρίου, με  $\sum b_j=1$  για κάθε εναλλακτική P<sub>i</sub>.

Τα βάρη των κριτηρίων ικανοποίησης υποδηλώνουν το σχετικό βαθμό σπουδαιότητας που δίνει το σύνολο των πελατών στις αξίες των διαστάσεων ικανοποίησης που έχουν καθοριστεί. Το γεγονός αυτό υποδηλώνει ότι η απόφαση για να θεωρηθεί κάποιο κριτήριο ως «σημαντικό», σε ένα βαθμό, εξαρτάται και από το πλήθος των κριτηρίων που χρησιμοποιούνται. Δεν θα πρέπει να λησμονείται η φυσική ερμηνεία των συντελεστών βαρύτητας, ότι τα βάρη είναι βαθμοί παραχώρησης (trade-offs) μεταξύ των αξιών στα κριτήρια.

Το b<sub>ji</sub> δείχνει το προτιμησιακό μοντέλο του χρήστη που έχει επιλέξει την εναλλακτική P<sub>i</sub>.

#### Διαδικασία Σύστασης

##### Γ. Φιλτράρισμα εναλλακτικών με βάση τις επιλογές του χρήστη στα On/Off κριτήρια

Ο χρήστης U<sub>f</sub> του συστήματος σύστασης επιλέγει Choice<sub>ifY</sub> {Yes, No, No choice} από τα πιθανά στατικά χαρακτηριστικά των εναλλακτικών P<sub>i</sub>,

**Επιλογές On/Off χρήστη=( U<sub>f</sub>, P<sub>i</sub>, Choice<sub>ifY</sub>) (3-14)**

Δημιουργία υποσυνόλου P<sub>filt</sub>. Ισχύει  $P_{filt} \subseteq P$ .

Όπου: Choice<sub>ifY</sub> η επιλογή του χρήστη U<sub>f</sub> για τα On/Off χαρακτηριστικά των εναλλακτικών P<sub>i</sub>.

##### Δ. Εφαρμογή της μεθόδου WAP με βάση τις προτιμήσεις του χρήστη

Ο χρήστης U<sub>f</sub> προβαίνει στις ακόλουθες ενέργειες ανάλογα με τις προτιμήσεις του.

1. Κατάταξη των κριτηρίων.  $c_x, x=1... \#Criteria$  τα κριτήρια που έχουν καταταχθεί από το χρήστη  $U_f$ .
2. Διαφορά σημαντικότητας μεταξύ διαδοχικών κριτηρίων με επιλογή λεκτικής περιγραφής.

Για δύο διαδοχικά κριτήρια  $c_r, c_{r+1}$  το εύρος τιμών  $[zmin_r, zmax_r]$  είναι τέτοιο ώστε να ισχύει  $zmin_r \leq z_r \leq zmax_r$  και  $c_r = z_r^* c_{r+1}$ . Στη συγκεκριμένη προσέγγιση ο αποφασίζοντας ρωτάται να κατατάξει τα  $n$  κριτήρια σε  $m$  κλάσεις ( $m \leq n$ ) και να προσδιορίσει  $2(m+1)$  τιμές για τα  $zmin_r$  και  $zmax_r$  όπου ( $r=1,2,...,m-1$ ). Το τελικό  $z$  μπορεί να υπολογιστεί από το γινόμενο των  $z_r$  δεικτών:

Δηλαδή:  $z = z_1 z_2 \dots z_{m-1} = (c_1/c_2)(c_2/c_3) \dots (c_{m-1}/c_m) = c_1/c_m$  Στην παρούσα προσέγγιση προβαίνουμε σε παραλλαγή της WAP ως προς τον τρόπο υπολογισμού των  $z$ . Έχει προηγηθεί αντιστοίχιση διαφορετικών ευρών τιμών  $[zmin_r, zmax_r]$  σε λέξεις ή φράσεις που κατατοπίζουν ευκολότερα ένα απλό χρήστη για την επιλογή του. Δημιουργούμε ερωτηματολόγιο που απευθύνουμε σε 50 χρήστες, οι οποίοι καλούνται να δείξουν παραστατικά με χρήση των ράβδων κύλισης, πώς νοούν τις διαφορές σημαντικότητας για τις εξής διαφορές προτίμησης: Πολύ Ασθενής Προτίμηση, Ασθενής Προτίμηση, Προτίμηση, Ισχυρή Προτίμηση, Πολύ Ισχυρή Προτίμηση και απεικονίζουν την προτιμησησική διαφορά που έχουν τα διαδοχικά κριτήρια μεταξύ τους, στην κατάταξη του χρήστη. Με βάση τα αποτελέσματα για  $zmin, zmax$ , για όλους τους ερωτηθέντες, καταλήγουμε σε αντιστοίχιση των τιμών  $zmin, zmax$  σε λεκτικές περιγραφές της διαφοράς σημαντικότητας.

Επίσης για τα ελάχιστα και μέγιστα  $z$  ισχύει:

$$\prod_{i=1}^{m-1} zmin_i \leq z \leq \prod_{i=1}^{m-1} zmax_i \quad (3-15)$$

Min  $c_i$  & Max  $c_i$ , for  $i = 1, 2, \dots, n$

$$c_i - c_{i+1} = 0, \quad (3-16)$$

αν  $g_{i+1}$  ακολουθείται από το  $g_i$ , και το  $g_{i+1}$  ανήκει στην ίδια τάξη σημαντικότητας ( $j$ ) όπως  $g_i$ ,

$$c_i - c_{i+1} \geq zmin_j, c_i - c_{i+1} \leq zmax_j \quad (3-17)$$

Αν  $g_r$  ακολουθείται από το  $g_{i+1}$ ,  $g_i$  ανήκει στην  $j$  και το  $g_{i+1}$  ανήκει στη τάξη  $j+1$ .



$$c_1 + c_2 + \dots + c_n = 1 \quad (3-18)$$

$$c_1 \geq 0, c_2 \geq 0, \dots, c_n \geq 0$$

$$w_1 \geq 0, w_2 \geq 0, \dots, w_k \geq 0$$

Αποτέλεσμα μοντελοποίησης προτίμησης

**Προτιμησιακό μοντέλο χρήστη: {  $U_f$ ,  $\text{MinWeightCr}_x$ ,  $\text{MidWeightCr}_x$ ,  $\text{MaxWeightCr}_x$ ,  $\text{ASI}_f$ ,  $\mu_f$  } (3-19)**

Όπου:  $U_f$ , ο νέος χρήστης του συστήματος συστάσεων.

$\text{MinWeightCr}_x$ , η ελάχιστη τιμή του βάρους του κριτηρίου x.

$\text{MidWeightCr}_x$ , η μεσαία τιμή του βάρους του κριτηρίου x.

$\text{MaxWeightCr}_x$ , ο μέγιστη τιμή του βάρους του κριτηρίου x.

$\text{ASI}_f$ , ο μέσος δείκτης ευστάθειας (Average Stability Index).

$\mu_f$ , ο δείκτης ευρωστίας.

E. Σύγκριση της ομοιότητας του προτιμησιακού μοντέλου του χρήστη  $U_f$  με τα προτιμησιακά μοντέλα των χρηστών  $U_g$

**α. Αν  $b_{xi} \in [\text{MinWeightCr}_x, \text{MaxWeightCr}_x]$  τότε:  $\text{Comp}_{xi}=1$  αλλιώς  $\text{Comp}_{xi}=0$  (3-20)**

**β.  $\text{WeSum}_i = \sum (b_{xi} * \text{Comp}_{xi})$**

ή

**$\text{WeSum}_i = \sum (\text{Comp}_{xi})$**  για το μοντέλο σύστασης με single sum για κάθε κριτήριο x και εναλλακτική  $P_i$ .

Με βάση τα δεδομένα της μεταβλητής  $\text{WeSum}_i$ , εφαρμόζεται κατάταξη (ranking) των εναλλακτικών.

Επιπρόσθετα, στην περίπτωση που η σειρά προτεραιότητας των εναλλακτικών δείχνει να ταυτίζεται, προτείνεται από τη μεθοδολογία η περαιτέρω κατάταξη των συγκεκριμένων εναλλακτικών, με βάση τη μέση συνολική αξιολόγηση για κάθε εναλλακτική,  $O_{i,g}$ , έχει δώσει ο χρήστης  $U_g$  για την εναλλακτική  $P_i$ , όπως έχει προαναφερθεί.

Προκύπτει σύνολο εναλλακτικών  $P_{WESU}$  που αποτελεί υποσύνολο του  $P$ ,  $P_{WESU} \subseteq P$  (3-21)

#### Z. Εφαρμογή Ανάλυσης Συναισθήματος

Εφαρμόζεται Ανάλυση Συναισθήματος στο υποσύνολο  $P_{WESU}$ .

**Είσοδοι Ανάλυσης συναισθήματος:**  $\{ U_g, P_i, review_{ig} \}$  (3-229)

Εφαρμόζεται ανάλυση του κειμένου με χρήση του λογισμικού Semantria.

**Αποτελέσματα Ανάλυσης συναισθήματος:**  $\{ U_g, P_i, Se_{ig}, \overline{Se}_i \}$  (3-30)

Όπου :

1.  $U_g$  το όνομα του χρήστη που έχει κάνει review
2.  $P_i$  τα προϊόντα/εναλλακτικές
3.  $Se_{ig}$  το sentiment score που κυμαίνεται από -2 έως 2
4.  $\overline{Se}_i$  το μέσο sentiment score για κάθε εναλλακτική  $P_i$

Η ανάλυση συναισθήματος χρησιμοποιείται στην παρούσα μεθοδολογία σαν κατώφλι ικανοποίησης προκειμένου να βελτιστοποιηθεί η ποιότητα των συστάσεων που θα γίνουν προς το χρήστη.

Προκύπτει υποσύνολο εναλλακτικών  $P_{wesuse}$ ,  $P_{wesuse} \subseteq P_{wesu} \subseteq P$  (3-23)

#### ΣΤ. Τελική Σύσταση προς το χρήστη $U_f$

Η τελική σύσταση προς το χρήστη  $U_f$ , προκύπτει από την τομή των υποσυνόλων των εναλλακτικών  $P_{wesuse}$  και  $P_{filt}$ .

Το υποσύνολο των εναλλακτικών που αποτελεί την τελική σύσταση ονομάζεται  $P_{rec}$  και ισχύει ότι:  $P_{rec} = P_{wesuse} \cap P_{filt}$  (3-24)

Οι κύριες μετρικές αξιολόγησης του συστήματος συστάσεων είναι οι παρακάτω:

$$MAE = \frac{\sum_{i=1}^N |p_i - r_i|}{N} \quad (3-33) \quad MSE = \frac{\sum_{i=1}^N (p_i - r_i)^2}{N} \quad (3-34) \quad RMSE = \sqrt{MSE} \quad (3-35)$$

όπου  $p_i$  είναι η προβλεπόμενη βαθμολογία,  $r_i$  είναι η πραγματική βαθμολογία και  $N$  είναι ο συνολικός αριθμός προβλέψεων.

Επίσης η μετρική ποσοστιαίου σφάλματος χρησιμοποιείται και είναι αρκετά παραστατική της καλής λειτουργίας.

$$PE = \frac{e_i - t_i}{e_i} * 100 \quad (3-36)$$

Όπου  $e_i$  είναι η πειραματική βαθμολογία και  $t_i$  είναι η θεωρητική βαθμολογία.

## 4. Σχεδίαση και Ανάπτυξη συστήματος

- 4.1 Δεδομένα
- 4.2 Εφαρμογή MUSA
- 4.3 Εφαρμογή Sentiment analysis
- 4.4 Εφαρμογή WAP
- 4.5 Τελική Σύσταση

### 4. Σχεδίαση και Ανάπτυξη συστήματος

Στην παρούσα έρευνα στόχος είναι να καταλήξουμε σε προσωποποιημένες συστάσεις προϊόντων/υπηρεσιών σε πελάτες, ανάλογα με το προτιμησιακό προφίλ τους, χρησιμοποιώντας μεγάλου όγκου δεδομένα, που αντλούμε από το διαδίκτυο. Για επαλήθευση του περιγεγραμμένου μεθοδολογικού πλαισίου, προβαίνουμε σε επεξεργασία δεδομένων με διαφορετικές μεθόδους, εν τέλει συνδυάζοντας τα αποτελέσματα των μεθόδων χρησιμοποιώντας σαν προϊόντα, ξενοδοχεία και σαν δεδομένα, κριτικές και βαθμολογίες πελατών, από γνωστή διαδικτυακή πλατφόρμα.

Σε συνέχεια της προηγούμενης παραγράφου όπου αποδεικνύεται το τεράστιο ενδιαφέρον για καλύτερες συστάσεις στο τομέα του τουρισμού και με δεδομένο ότι το μεγαλύτερο μέρος της διδακτορικής έρευνας πραγματοποιήθηκε στα Χανιά, Κρήτης, έναν από τους δημοφιλέστερους τουριστικούς προορισμούς παγκοσμίως, αποφασίστηκε η μελέτη περίπτωση της διατριβής να αναπτυχθεί με δεδομένα για τα ξενοδοχεία που βρίσκονται στον νομό Χανίων. Η επιλογή αυτή έγινε και με δεδομένο ότι ήταν δυνατή η επιτόπια επαλήθευση των αποτελεσμάτων των μεθόδων, καθώς υπήρχε η δυνατότητα επίσκεψης στα ξενοδοχεία ή εύκολης εύρεσης ατόμων που τα έχουν επισκεφθεί στο παρελθόν.

#### 4.1 Δεδομένα

Τα δεδομένα που χρησιμοποιήθηκαν, αντλήθηκαν με μεθόδους εξόρυξης δεδομένων, από τη διαδικτυακή πλατφόρμα κρατήσεων ξενοδοχείων, [www.hotels.com](http://www.hotels.com). Η συγκεκριμένη πλατφόρμα χρησιμοποιείται σε παγκόσμιο επίπεδο για κρατήσεις σε τουριστικά καταλύματα και χρησιμοποιεί πρότυπα πιστοποίησης της εγκυρότητας των κριτικών και αξιολογήσεων από τους χρήστες της. Πέραν των κριτικών των πελατών που έχουν τη μορφή κειμένου, η πλατφόρμα ζητάει από τους πελάτες να αξιολογήσουν την εμπειρία τους βαθμολογώντας το ξενοδοχείο με βάση 5 κριτήρια και ένα συνολικό σκορ. Τα κριτήρια αυτά είναι τα εξής: Καθαριότητα (Cleanliness), Εξυπηρέτηση (Service), Άνεση (Comfort), Κατάσταση (Condition), Γειτονιά (Neighbourhood) και συνολικό σκορ (Overall Score). Τα προαναφερθέντα κριτήρια έχουν βαθμολογηθεί σε 5-βάθμια κλίμακα, με το 1 να είναι η χειρότερη και 5 η μεγαλύτερη βαθμολογία.

- ✓ Η Καθαριότητα αναφέρεται στα πρότυπα υγιεινής του ξενοδοχείου, στις εγκαταστάσεις καθαριότητας και στην ποιότητα του προσωπικού.
- ✓ Η Εξυπηρέτηση περιλαμβάνει την παροχή στους επισκέπτες πληροφοριών, ειδικού εξοπλισμού και προμηθειών.
- ✓ Η Άνεση χρησιμοποιείται για να περιγράψει τις ανέσεις του ξενοδοχείου ή των δωματίων και είναι αρκετά σχετική με την πολυτέλεια που προσφέρεται.
- ✓ Η Κατάσταση αναφέρεται στην ποιότητα των υποδομών, δηλαδή στο επίπεδο παλαιότητας και συντήρησης.
- ✓ Η Γειτονιά είναι το κριτήριο που δείχνει την ποιότητα της θέσης στην οποία βρίσκεται το ξενοδοχείο.

Μερικές επίσης παράμετροι που χρησιμοποιήθηκαν σαν κριτήρια προκειμένου να καταλήξουμε σε αυτήν την πλατφόρμα, ήταν οι παρακάτω:

- α. Η Εγκυρότητα, ο μεγάλος Όγκος και η ποιότητα των δεδομένων.
- β. Η συμβατότητα της ιστοσελίδας με τα λογισμικά που χρησιμοποιήθηκαν για την άντληση των δεδομένων.

γ. Η συμβατότητα της δομής των κριτηρίων αξιολόγησης με τα προαπαιτούμενα, για την εφαρμογή των μεθόδων MUSA, ανάλυσης συναισθήματος και WAP.

#### Δεδομένα που αντλήθηκαν

Το σύνολο των αξιολογήσεων και κριτικών που αντλήθηκαν προέρχονταν από 4618 πελάτες και αφορούσαν 166 ξενοδοχεία του νομού Χανίων. Πρέπει να σημειωθεί ότι στη βάση δεδομένων μας δεν συμπεριλήφθηκαν ξενοδοχεία για τα οποία υπήρχαν κριτικές από λιγότερους από 5 πελάτες, καθώς δεν κρίνεται επαρκές δείγμα για να κριθεί ένα ξενοδοχείο.

Παρακάτω παρατίθεται ο πίνακας 4-1 που δείχνει αναλυτικά, τον αριθμό των κριτικών, σε κάθε κατηγορία Ξενοδοχείων, καθώς και ο αριθμός των ξενοδοχείων ανά κατηγορία.

**Πίνακας 4-1 Αριθμός κριτικών και ξενοδοχείων ανά κατηγορία ξενοδοχείων που αντλήθηκαν**

Κατηγορία Ξενοδοχείου	#Ξενοδοχείων	#Κριτικών
1 Stars	16	258
2 Stars	49	1100
3 Stars	61	1927
4 Stars	25	722
5 Stars	15	611
<b>Σύνολο</b>	<b>166</b>	<b>4618</b>

Για καθένα από τα 166 ξενοδοχεία αντλήθηκαν τα ακόλουθα δεδομένα που παρουσιάζονται στο πίνακα 4-2:

**Πίνακας 4-2 Στατικά δεδομένα που αντλήθηκαν για τα ξενοδοχεία**

	Δεδομένα	Εύρος τιμών
1	Αριθμός Αστεριών (Stars)	[1-5]
2	Εξωτερική Πισίνα (Outdoor pool)	[Ναι, Όχι]
3	Εσωτερική Πισίνα (Indoor pool)	[Ναι, Όχι]
4	Μασάζ (Massage)	[Ναι, Όχι]
5	Τηλεόραση (TV)	[Ναι, Όχι]

6	Εστιατόριο (Restaurant)	[Ναι, Όχι]
7	Μπαρ (Bar)	[Ναι, Όχι]
8	Δωρεάν ασύρματο δίκτυο (Free WIFI)	[Ναι, Όχι]
9	Θέρμανση (Heating)	[Ναι, Όχι]
10	Κλιματισμός (Air Conditioning)	[Ναι, Όχι]
11	Οικογενειακά Δωμάτια (Family rooms)	[Ναι, Όχι]
12	Εγκαταστάσεις για άτομα με ειδικές ανάγκες (Facilities for disabled guests)	[Ναι, Όχι]
13	Λεωφορείο από Αεροδρόμιο (Airport shuttle)	[Ναι, Όχι]
14	Χρηματοκιβώτιο (Safety deposit box)	[Ναι, Όχι]
15	Δωρεάν Πάρκινγκ (Free Parking)	[Ναι, Όχι]
16	Υπηρεσία Δωματίου (Room service)	[Ναι, Όχι]
17	Μη καπνιζόντων (Non-smoking throughout)	[Ναι, Όχι]
18	Δωμάτια Μη καπνιζόντων (Non-smoking rooms)	[Ναι, Όχι]
19	Ανελκυστήρας (Lift)	[Ναι, Όχι]
20	Αγγλικά (English)	[Ναι, Όχι]
21	Γερμανικά (German)	[Ναι, Όχι]
22	Ρώσικά (Russian)	[Ναι, Όχι]
23	Γαλλικά (French)	[Ναι, Όχι]
24	Ισπανικά (Spanish)	[Ναι, Όχι]
25	Τιμή (Price)	[0-1500] €
26	Απόσταση ως το κέντρο της πόλης (Distance to city center)	[0-50]miles
27	Απόσταση ως το αεροδρόμιο (Distance to the airport)	[0-50]miles

Τα παραπάνω δεδομένα έχουν καταχωρηθεί στη βάση του hotels.com, από τα ίδια τα ξενοδοχεία.

Τα παραπάνω δεδομένα θα χρησιμοποιηθούν εν συνεχεία στο σύστημα σύστασης μας, σαν αρχικές επιλογές και κατ' αποκοπή κριτήρια για το χρήστη (On/Off criteria), προκειμένου να φιλτραριστούν οι διαθέσιμες επιλογές ξενοδοχείων, πριν την τελική σύσταση.

Η επιλογή των παραπάνω χαρακτηριστικών για χρήση τους σαν κατ' αποκοπή κριτήρια έγινε κατόπιν σχετικής έρευνας για τα κριτήρια που χρησιμοποιούνται στις πιο δημοφιλείς πλατφόρμες κρατήσεων τουριστικών καταλυμάτων,

Booking.com, Expedia.com, Tripadvisor. Επιπρόσθετα αντλήθηκαν πληροφορίες από ερευνητική βιβλιογραφία (Fuchs and Zanker, 2012),(Berka and Plöbzig, 2004), καθώς και διαδικτυακές πηγές πληροφοριών, Quora.com. Για καθεμία από τις 4618 κριτικές αντλήθηκαν επίσης τα δεδομένα που εμφανίζονται στο πίνακα 4-3:

Πίνακας 4-3 Δεδομένα βαθμολογιών και κριτικών πελατών που αντλήθηκαν

	Δεδομένα	Εύρη Τιμών
1	Όνομα Ξενοδοχείου	Κείμενο
2	Overall Reviewer Score	[1-5]
3	Cleanliness	[1-5]
4	Service	[1-5]
5	Comfort	[1-5]
6	Condition	[1-5]
7	Neighbourhood	[1-5]
8	Review	Κείμενο
9	Εθνικότητα του Reviewer	Κείμενο

Τα δεδομένα του παραπάνω πίνακα έχουν παραχθεί από χρήστες της πλατφόρμας που έχουν κλείσει τη διαμονή τους με χρήση της, και κατόπιν της διαμονής τους, έχουν δημοσιεύσει την κριτική τους και τις βαθμολογίες τους.

#### Επεξεργασία δεδομένων για περαιτέρω ανάλυση

Στο παρακάτω πλαίσιο εμφανίζεται μια κριτική ξενοδοχείου (τυχαίο δείγμα), με τη μορφή που καταχωρήθηκε στο spreadsheet, μετά την άντληση της από τη διαδικτυακή πλατφόρμα.

4.0 / 5 Excellent Cleanliness 3.0 Service 4.0 Comfort 3.0 Condition 4.0 Neighbourhood 5.0 Hotel was excellent, service was great. Located in a very quiet spot and ideal for families. It maybe a good idea to rent a car as the local bus is very few and fair between but the hotel provides info and the bus stop is about a ten min walk. Close to the beach and our room came with a great see view, the pool was excellent as well. Genuine Hotels.com guest review 2 night trip in Jul 2013 A Traveller GB

Όλες οι κριτικές που αντλήθηκαν είχαν ακριβώς την ίδια μορφή (format).



Αφότου ολοκληρώθηκε η άντληση δεδομένων, πραγματοποιήθηκε επεξεργασία των δεδομένων. Στην παράγραφο 3.2, περιγράφεται αναλυτικά η διαδικασία άντλησης και επεξεργασίας (Βλέπε Σχήματα: 3.3, 3.4, 3.5, 3.6, 3.7).

Το αποτέλεσμα της επεξεργασίας ήταν τα δεδομένα να ταξινομηθούν σε κελιά και να είναι εύκολα προσπελάσιμα για την εφαρμογή των αλγορίθμων. Στο πίνακα 4-4 εμφανίζεται ένα στιγμιότυπο από τη μορφή της κριτικής κατόπιν επεξεργασίας σε δομημένο πίνακα.

Πίνακας 4-4 Μορφή της κριτικής κατόπιν επεξεργασίας

Overall Reviewer Score	Cleanliness	Service	Comfort	Condition	Neighbourhood	Review	Reviewer Origin
4	3.0	4.0	3.0	4.0	5.0	Hotel was excellent, service was great. Located in a very quiet spot and ideal for familiesClose to the beach and our room came with a great see view, the pool was excellent as well. Genuine Hotels.com guest review 2 night trip in Jul 2013 A Traveller GB	GB

## 4.2 Εφαρμογή MUSA

Επόμενος στόχος ήταν να διαμορφωθούν τα δεδομένα αξιολόγησης με κατάλληλο τρόπο, ώστε να εφαρμοστεί η μέθοδος MUSA. Μέσω της εφαρμογής της μεθόδου MUSA στις αξιολογήσεις των ξενοδοχείων, προκύπτουν οι βαρύτητες των κριτηρίων για καθένα από τα ξενοδοχεία. Ουσιαστικά λοιπόν μέσω της ανάλυσης ικανοποίησης των πελατών, που προκύπτει από όλες τις βαθμολογίες, δημιουργείται ένα συγκεκριμένο προφίλ για καθένα από τα ξενοδοχεία. Στην παράγραφο 3.3 περιγράφεται αναλυτικά η διαδικασία εφαρμογής της MUSA στο παρόν μεθοδολογικό πλαίσιο (Βλέπε Πίνακες: 3.3,3.4). Στόχος είναι μετά την εφαρμογή MUSA για κάθε ξενοδοχείο να προκύπτει το διάνυσμα βαρυτήτων του ως εξής:

[b<sub>1</sub>(Cleanliness), b<sub>2</sub>(Service), b<sub>3</sub>(Comfort), b<sub>4</sub>(Condition), b<sub>5</sub>(Neighborhood)]

Ισχύει ότι:  $b_1+b_2+b_3+b_4+b_5=100$  (4-1)

Το οποίο θα χρησιμοποιηθεί σε μεταγενέστερο χρόνο για σύγκριση με το προτιμησιακό προφίλ του χρήστη καθώς και για τη διαδικασία σύστασης.

Προκειμένου να τρέξει αξιόπιστα η μέθοδος MUSA, απαιτείται για κάθε προϊόν/υπηρεσία (στην περίπτωση μας ξενοδοχείο) να υπάρχουν τουλάχιστον 25 αξιολογήσεις. Με δεδομένο ότι σε αρκετές περιπτώσεις ξενοδοχείων, ο αριθμός των κριτικών δεν ξεπερνούσε τις 25, εφαρμόστηκε ομαδοποίηση των κριτικών των συγκεκριμένων ξενοδοχείων με βάση την απόσταση τους από το κέντρο της πόλης καθώς και την κατηγορία αστεριών στην οποία ανήκουν. Για τα υπόλοιπα εφαρμόστηκε η μέθοδος ανεξάρτητα.

Πίνακας 4-5 Αριθμός Ξενοδοχείων και κριτικών ανά ομάδα

1 Star	2 Stars	3 Stars	4 Stars	5 Stars
<ul style="list-style-type: none"> <li>• Group1 (0-1 miles) #hotels=2 #Reviews=29</li> <li>• Group2 (1-2.5 miles) #hotels=2 #Reviews=37</li> <li>• Group3 (3-15 miles) #hotels=5 #Reviews=66</li> <li>• Group4 (20 miles) #hotels=3 #Reviews=51</li> <li>• Group5 (23 miles) #hotels=2 #Reviews=23</li> <li>• Ungrouped Hotels #hotels=2 #totalReviews=61</li> </ul>	<ul style="list-style-type: none"> <li>• Group6 (0-1 miles) #hotels=7 #Reviews=117</li> <li>• Group7 (1-2.5 miles) #hotels=4 #Reviews=60</li> <li>• Group8 (3-5.8 miles) #hotels=10 #Reviews=295</li> <li>• Group9 (6-7 miles) #hotels=9 #Reviews=92</li> <li>• Group10 (10-32miles) #hotels=6 #Reviews=79</li> <li>• Ungrouped Hotels #hotels=12 #totalReviews=447</li> </ul>	<ul style="list-style-type: none"> <li>• Group11 (0-2.5 miles) #hotels=17 #Reviews=337</li> <li>• Group12 (3-6 miles) #hotels=9 #Reviews=148</li> <li>• Group13 (6.1-7 miles) #hotels=5 #Reviews=73</li> <li>• Group14 (7-15 miles) #hotels=6 #Reviews=92</li> <li>• Group15 (17-22miles) #hotels=6 #Reviews=110</li> <li>• Ungrouped Hotels #hotels=19 #totalReviews=1137</li> </ul>	<ul style="list-style-type: none"> <li>• Group16 (0-0.5 miles) #hotels=6 #Reviews=156</li> <li>• Group17 (3-10 miles) #hotels=4 #Reviews=57</li> <li>• Group18 (13-25 miles) #hotels=5 #Reviews=81</li> <li>• Ungrouped Hotels #hotels=10 #totalReviews=428</li> </ul>	<ul style="list-style-type: none"> <li>• Group19 (0-0.5 miles) #hotels=4 #Reviews=54</li> <li>• Group20 (5-20 miles) #hotels=6 #Reviews=222</li> <li>• Ungrouped Hotels #hotels=5 #totalReviews=335</li> </ul>

Στον παραπάνω πίνακα (Πίνακας 4-5) εμφανίζεται συνοπτικά ο αριθμός των ομάδων ξενοδοχείων και ο αντίστοιχος αριθμός αξιολογήσεων κάθε ομάδας που χρησιμοποιήθηκαν σαν είσοδοι για την ανάλυση ικανοποίησης με τη μέθοδο MUSA. Για τα ξενοδοχεία που δεν εντάχθηκαν σε κάποια ομάδα ο αλγόριθμος έτρεξε ξεχωριστά.

Μια ακόμη σημαντική παράμετρος ήταν να διερευνηθούν οι βέλτιστες τιμές για το Κατώφλι προτίμησης,  $\gamma$  (preference threshold) και το Βαθμό παραχώρησης,  $\epsilon$  (post optimality threshold), που εν συνεχεία χρησιμοποιούνται για όλα τα τρεξίματα MUSA στα υπάρχοντα δεδομένα. Προκειμένου να καταλήξουμε στις κατάλληλες τιμές εφαρμόσαμε MUSA για διαφορετικές τιμές των ( $\gamma$ ,  $\epsilon$ ), σε δείγμα δεδομένων. Κατόπιν συγκρίναμε τα αποτελέσματα των μεταβλητών-δεικτών :

- Ικανοποίηση (Satisfaction)
- Απαιτητικότητα (Demand)
- Δείκτης Προσαρμογής (Fitting Index)
- Ευστάθεια (Stability Index)
- Συνολική Πρόβλεψη (Overall Prediction)

Προσοχή πρέπει να δοθεί σε αυτό το σημείο στο γεγονός, ότι για κάποιες περιπτώσεις τιμών ( $\gamma$ ,  $\epsilon$ ), παρόλο που τα αποτελέσματα των προαναφερθεισών μεταβλητών ήταν πολύ υψηλά, οι βαρύτητες των κριτηρίων εμφανίζονται οι ίδιες. Συνεπώς είναι σκόπιμο να προτιμηθούν τιμές κατά τις οποίες οι δείκτες έχουν υψηλές τιμές και οι βαρύτητες των κριτηρίων παρουσιάζουν διακύμανση. Στην περίπτωση των παρόντων δεδομένων καταλήξαμε στις παρακάτω τιμές:

Κατώφλι προτίμησης,  $\gamma=2$ , Βαθμός παραχώρησης,  $\epsilon=7$

Εν συνεχεία εφαρμόστηκε η μέθοδος MUSA για όλες τις αξιολογήσεις, και τα αποτελέσματα αποθηκεύτηκαν σε spreadsheet. Η μορφή των αποτελεσμάτων για κάθε ξενοδοχείο είναι πανομοιότυπη με τον παρακάτω πίνακα 4-6 :

**Πίνακας 4-6 Αποτελέσματα MUSA για ένα ξενοδοχείο του δείγματος**

b1 Cleanliness	b2 Service	b3 Comfort	b4 Condition	b5 Neighborhood	Stability Index	Fitting Index	Global Satisfaction Index	Preference Threshold	Post Optimality Threshold
10,156	10,156	9,078	50,610	20,000	71,77%	96,98%	92,233%	2	7

### 4.3 Εφαρμογή Sentiment analysis

Για την εφαρμογή της συναισθηματικής ανάλυσης χρησιμοποιήθηκαν σαν δεδομένα οι κριτικές (κείμενο) των πελατών για τα ξενοδοχεία, οι οποίες έχουν αντληθεί μαζί με τις αξιολογήσεις, όπως έχει αναφερθεί παραπάνω. Στόχος της χρήσης της συναισθηματικής ανάλυσης στο σύστημα μας είναι να αποτελέσει ένα επιπρόσθετο φίλτρο ικανοποίησης για τη σύσταση. Στην περίπτωση μας το αποτέλεσμα της λειτουργεί σαν κατώφλι για τα ξενοδοχεία, ώστε να προχωρήσουν και προαχθούν σε βασική επιλογή σύστασης για το χρήστη. Στην παράγραφο 3.4 περιγράφεται αναλυτικά η εφαρμογή της ανάλυσης συναισθήματος στο παρόν μεθοδολογικό πλαίσιο (Βλέπε Πίνακες: 3.5,3.6).

Η συναισθηματική ανάλυση πραγματοποιήθηκε με χρήση της εφαρμογής *semantria*, και μετά την εφαρμογή της, για κάθε κριτική προέκυψε μια βαθμολογία συναισθήματος εγγράφου (document sentiment). Η βαθμολογία αυτή κυμαίνεται στα όρια  $[-2, 2]$ , και είναι ανάλογη το περιεχομένου του κειμένου. Όταν η βαθμολογία είναι  $(-2)$  η κριτική είναι πολύ αρνητική, όταν είναι  $(-1)$  θεωρείται αρνητική, ενώ για τιμή  $(+1)$  θεωρείται θετική και  $(+2)$  πολύ θετική. Για το εύρος τιμών από  $-0.05$  έως  $+0.22$  η κριτική θεωρείται ουδέτερη.

Προκειμένου να χρησιμοποιήσουμε τις βαθμολογίες συναισθήματος, για κάθε ξενοδοχείο, υπολογίσαμε τον μέσο όρο των βαθμολογιών που έχουν προκύψει από την ανάλυση όλων των διαφορετικών κριτικών που το αφορούν. Επιπρόσθετα για την περίπτωση των δεδομένων μας θεωρήσαμε σαν κατώφλι ικανοποίησης τη βαθμολογία  $+0.22$ , καθώς στόχος ήταν οι βαθμολογίες συναισθήματος για τα ξενοδοχεία που κάνουμε σύσταση στο χρήστη του συστήματος να είναι τουλάχιστον θετικές. Έτσι καταλήγουμε σε ένα πίνακα της μορφής του Πίνακα 4-7:

Πίνακας 4-7 Αποτελέσματα ανάλυσης συναισθήματος για 2 ξενοδοχεία του δείγματος

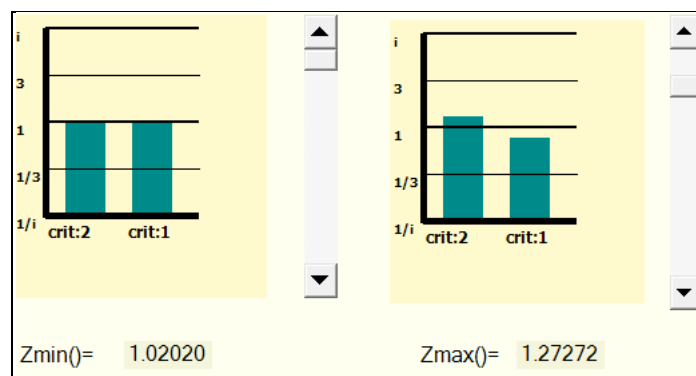
Hotel Name	Average Sentiment	Pass the threshold (Threshold=0.22) (0=No pass, 1=pass)

Hotel_1	0.091	0
Hotel_2	0.348	1

#### 4.4 Εφαρμογή WAP

Κατά τη λειτουργία της WAP, ο χρήστης καλείται να εισάγει στο σύστημα, μια αρχική κατάταξη των κριτηρίων που έχουν χρησιμοποιηθεί για την αξιολόγηση των ξενοδοχείων (Καθαριότητα, Εξυπηρέτηση, Άνεση, Κατάσταση, Γειτονιά), όπου ουσιαστικά εκφράζει ποια κριτήρια αποτελούν προτεραιότητα για αυτόν και με ποια σειρά. Εν συνεχεία καλείται μέσω επιλογής να δείξει πόσο σημαντικότερο κατ' αυτήν/αυτόν είναι το κάθε κριτήριο από το προηγούμενο, στην κατάταξη που έχει προηγουμένως προσδιορίσει. Στην παράγραφο 3.5 περιγράφεται αναλυτικά η διαδικασία εφαρμογής της μοντελοποίησης προτίμησης στο παρόν μεθοδολογικό πλαίσιο (Βλέπε Πίνακα: 3.7).

Οι μεταβλητές που χρησιμοποιούνται από την WAP προκειμένου να προκύψει η διαφορά σημαντικότητας μεταξύ των διαδοχικών κριτηρίων είναι οι  $Z_{min}, Z_{max}$ . Πιο συγκεκριμένα, κατόπιν της διαδικασίας κατάταξης, ο χρήστης καλείται να τροποποιήσει τις τιμές  $Z_{min}, Z_{max}$ , χρησιμοποιώντας ράβδους κύλισης (βλέπε σχήμα 4-1) προκειμένου οπτικά να καθορίσει τη διαφορά σημαντικότητας του ενός κριτηρίου από το επόμενο στην κατάταξη. Στο κάτω μέρος της εικόνας εμφανίζονται οι αντίστοιχες τιμές  $Z_{min}, Z_{max}$ , που αλλάζουν ανάλογα με τις επιλογές του χρήστη.



Σχ. 4-1 Ράβδοι κύλισης για υπολογισμό  $Z_{min}$ ,  $Z_{max}$  στη WAP

Κατά την υλοποίηση του συστήματος μας, στόχος ήταν να προσφέρουμε στο χρήστη ένα όσο το δυνατόν πιο απλοποιημένο, κατανοητό και φιλικό περιβάλλον για την διεξαγωγή της διαδικασίας σύστασης. Για αυτό το λόγο αποφασίστηκε, αντί να καλούμε το χρήστη για κάθε σύγκριση διαδοχικών κριτηρίων να επιλέξει τη διαφορά σημαντικότητας με χρήση των ράβδων κύλισης, να του δίνεται η δυνατότητα να επιλέξει μεταξύ λεκτικών τιμών που αντιστοιχίζονται σε συγκεκριμένες (καθορισμένες εκ των προτέρων) τιμές  $Z_{min}, Z_{max}$ .

Για να προβούμε στην προαναφερθείσα ενέργεια, αναπτύξαμε ερωτηματολόγιο το οποίο απευθύνθηκε σε 50 διαφορετικούς χρήστες. Από τους χρήστες ζητήθηκε να χρησιμοποιήσουν τις ράβδους κύλισης και παρατηρώντας τις οπτικά, να αναφέρουν πώς νοούνται για αυτούς οι εξής έννοιες όσον αφορά τη διαφορά σημαντικότητας: πολύ ασθενής προτίμηση (very weak preference), ασθενής προτίμηση (weak preference), απλή προτίμηση (preference), ισχυρή προτίμηση (strong preference), πολύ ισχυρή προτίμηση (very strong preference). Για κάθε χρήστη, αντλήθηκαν οι τιμές  $Z_{min}, Z_{max}$  και στο τέλος υπολογίστηκαν οι μέσοι όροι τους για κάθε διαφορετική λεκτική τιμή. Οι μέσοι όροι που προέκυψαν, χρησιμοποιήθηκαν στο σύστημα, ενώ ο χρήστης έχει τη δυνατότητα να επιλέξει μεταξύ των λεκτικών τιμών: very weak preference, weak preference, preference, strong preference, very strong preference, προκειμένου να δείξει τη διαφορά σημαντικότητας μεταξύ διαδοχικών κριτηρίων.

Στον παρακάτω πίνακα 4-8 εμφανίζονται οι μέσοι όροι των  $Z_{min}$ ,  $Z_{max}$ . Για τα αναλυτικά αποτελέσματα του ερωτηματολογίου βλέπε παράρτημα.

Πίνακας 4-8 Μέσοι όροι των  $Z_{min}$ ,  $Z_{max}$  με βάση τα αποτελέσματα του ερωτηματολογίου

Z/Pref.type	Very Weak Preference	Weak Preference	Preference	Strong Preference	Very Strong Preference
$Z_{min}$	1.0243	1.3873	1.8289	2.3161	2.8138
$Z_{max}$	1.3512	1.7978	2.3051	2.8055	3.4009

Με βάση τα παραπάνω δεδομένα για το χρήστη, γίνεται εφαρμογή της WAP και το αποτέλεσμα είναι, για κάθε κριτήριο ένα εύρος βαρυτήτων. Τα εύρη βαρυτήτων για όλα τα κριτήρια προσδιορίζουν το προτιμησιακό προφίλ του χρήστη.

Έτσι, μετά την εφαρμογή της WAP για ένα χρήστη, προκύπτει πίνακας της παρακάτω μορφής (πίνακας 4-9).

Πίνακας 4-9 Αποτελέσματα WAP για ένα τυχαίο χρήστη

Criteria	Min/Mid/Max Weights	Values
Cleanliness	Cleanliness Min Weight	0.0492
	Cleanliness Mid Weight	0.0621
	Cleanliness Max Weight	0.0761
Service	Service Min Weight	0.2823
	Service Mid Weight	0.3115
	Service Max Weight	0.3411
Comfort	Comfort Min Weight	0.1073
	Comfort Mid Weight	0.123
	Comfort Max Weight	0.139664
Condition	Condition Min Weight	0.242642
	Condition Mid Weight	0.258826
	Condition Max Weight	0.276233
Neighbourhood	Neighbourhood Min Weight	0.227907
	Neighbourhood Mid Weight	0.244398
	Neighbourhood Max Weight	0.262236

## 4.5 Τελική Σύσταση

Για την τελική σύσταση έχουν αναπτυχθεί 4 διαφορετικές μέθοδοι, που προοδευτικά καταλήγουν σε όλο και λιγότερα σε αριθμό και πιο ακριβή αποτελέσματα συστάσεων για το χρήστη. Στην παράγραφο 3.6 περιγράφεται αναλυτικά η διαδικασία της τελικής σύστασης για το παρόν μεθοδολογικό πλαίσιο.

### **A. Σύσταση με βάση το προτιμησιακό μοντέλο και απλό άθροισμα (Recommendation Based on Preference Model with Single Sum)**

Χρησιμοποιώ για τη σύσταση, την κατάταξη των ξενοδοχείων που έχει προκύψει με βάση τη σύγκριση των βαρών των κριτηρίων από MUSA για το κάθε ξενοδοχείο με τα εύρη βαρυτήτων από τη WAP. Για κάθε ξενοδοχείο ελέγχεται αν το βάρος του αντίστοιχου κριτηρίου βρίσκεται στο διάστημα ( $W_{cmin}$ - $W_{cmax}$ ) που αποτελεί το προτιμησιακό προφίλ του χρήστη για το συγκεκριμένο κριτήριο και σε περίπτωση σύγκλισης τοποθετείται 1 στο αντίστοιχο κελί αλλιώς 0. Για κάθε ξενοδοχείο αθροίζονται τα κελιά όλων των κριτηρίων και τελικά κατατάσσονται με φθίνουσα σειρά ανάλογα με το άθροισμα. Τα αποτελέσματα της προηγούμενης κατάταξης, τα κατατάσσουμε και σε δεύτερο βαθμό με χρήση της μέσης συνολικής αξιολόγησης για κάθε ξενοδοχείο ώστε να καταλήγουμε σε απόλυτα ακριβή κατάταξη.

### **B. Σύσταση με βάση το προτιμησιακό μοντέλο και το σταθμισμένο άθροισμα (Recommendation Based on Preference model and Weighted Sum)**

Χρησιμοποιώ για τη σύσταση, την κατάταξη των ξενοδοχείων που έχει προκύψει με βάση τη σύγκριση των βαρών των κριτηρίων από MUSA για το κάθε ξενοδοχείο με τα εύρη βαρυτήτων από τη WAP. Για κάθε ξενοδοχείο ελέγχεται αν το βάρος του αντίστοιχου κριτηρίου βρίσκεται στο διάστημα ( $W_{cmin}$ - $W_{cmax}$ ) που αποτελεί το προτιμησιακό προφίλ του χρήστη για το συγκεκριμένο κριτήριο και σε περίπτωση σύγκλισης τοποθετείται 1 στο



αντίστοιχο κελί, αλλιώς 0. Για κάθε ξενοδοχείο αθροίζονται τα κελιά όλων των κριτηρίων πολλαπλασιασμένα με το βάρος των αντίστοιχων κριτηρίων των ξενοδοχείων από MUSA, και τελικά κατατάσσονται με φθίνουσα σειρά ανάλογα με το άθροισμα. Τα αποτελέσματα της προηγούμενης κατάταξης, τα κατατάσσουμε και σε δεύτερο βαθμό με χρήση της μέσης συνολικής αξιολόγησης για κάθε ξενοδοχείο.

#### **Γ. Σύσταση με βάση το Προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα και το κατώφλι συναισθήματος (Recommendation Based on Preference model, Weighted Sum and Sentiment Threshold)**

Για κάθε ξενοδοχείο ελέγχεται αν το βάρος του αντίστοιχου κριτηρίου βρίσκεται στο διάστημα ( $W_{cmin}$ - $W_{cmax}$ ) που αποτελεί το προτιμησιακό προφίλ του χρήστη για το συγκεκριμένο κριτήριο και σε περίπτωση σύγκλισης τοποθετείται 1 στο αντίστοιχο κελί αλλιώς 0. Για κάθε ξενοδοχείο αθροίζονται τα κελιά όλων των κριτηρίων πολλαπλασιασμένα με το βάρος των αντίστοιχων κριτηρίων των ξενοδοχείων από MUSA, και τελικά κατατάσσονται με φθίνουσα σειρά ανάλογα με το άθροισμα. Τα αποτελέσματα της προηγούμενης κατάταξης, τα κατατάσσουμε και σε δεύτερο βαθμό με χρήση της μέσης συνολικής αξιολόγησης για κάθε ξενοδοχείο.

Στο συγκεκριμένο αποτέλεσμα προσθέτω συνθήκη για το κατώφλι ευχαρίστησης για όλα τα ξενοδοχεία, με βάση την ανάλυση συναισθήματος στην οποία έχω προβεί για κάθε ξενοδοχείο με βάση το κείμενο των κριτικών χρηστών. Αν το average sentiment threshold είναι 1 για κάποιο ξενοδοχείο, σημαίνει ότι το συναίσθημα των πελατών που έχει προκύψει από όλα τα reviews για το συγκεκριμένο ξενοδοχείο είναι θετικό και συνεπώς εξασφαλίζεται ότι υπάρχει ταύτιση του προτιμησιακού μοντέλου του χρήστη με το ξενοδοχείο, και τα reviews αποδεικνύουν ότι οι πελάτες του είναι ευχαριστημένοι. Σε αντίθετη περίπτωση το συναίσθημα για το ξενοδοχείο είναι αρνητικό ή ουδέτερο και δεν περνάει το κατώφλι προκειμένου να αποτελεί μέλος των τελικών εναλλακτικών σύστασης.

#### **Δ. Σύσταση με βάση το Προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα, το κατώφλι συναισθήματος και το μοντέλο φιλτραρίσματος (Recommendation Based on Preference Model, Weighted Sum, Sentiment Threshold and Filtering Model)**

Για κάθε ξενοδοχείο ελέγχεται αν το βάρος του αντίστοιχου κριτηρίου βρίσκεται στο διάστημα ( $W_{cmin}$ - $W_{cmax}$ ) που αποτελεί το προτιμησιακό προφίλ του χρήστη για το συγκεκριμένο κριτήριο και σε περίπτωση σύγκλισης τοποθετείται 1 στο αντίστοιχο κελί αλλιώς 0. Για κάθε ξενοδοχείο αθροίζονται τα κελιά όλων των κριτηρίων πολλαπλασιασμένα με το βάρη των αντίστοιχων κριτηρίων των ξενοδοχείων από MUSA, και τελικά κατατάσσονται με φθίνουσα σειρά ανάλογα με το άθροισμα. Τα αποτελέσματα της προηγούμενης κατάταξης, τα κατατάσσουμε και σε δεύτερο βαθμό με χρήση της μέσης συνολικής αξιολόγησης για κάθε ξενοδοχείο.

Στο συγκεκριμένο αποτέλεσμα προσθέτω συνθήκη για το κατώφλι ευχαρίστησης για όλα τα ξενοδοχεία, με βάση την ανάλυση συναισθήματος στην οποία έχω προβεί για κάθε ξενοδοχείο με βάση το κείμενο των κριτικών χρηστών. Αν το average sentiment threshold είναι 1 για κάποιο ξενοδοχείο, σημαίνει ότι το συναίσθημα των πελατών που έχει προκύψει από όλα τα reviews για το συγκεκριμένο ξενοδοχείο είναι θετικό και συνεπώς εξασφαλίζεται ότι υπάρχει ταύτιση του προτιμησιακού μοντέλου του χρήστη με το ξενοδοχείο, και τα reviews αποδεικνύουν ότι οι πελάτες του είναι ευχαριστημένοι. Σε αντίθετη περίπτωση το συναίσθημα για το ξενοδοχείο είναι αρνητικό ή ουδέτερο και δεν περνάει το κατώφλι προκειμένου να αποτελεί μέλος των τελικών εναλλακτικών σύστασης.

Επιπλέον οι επιλογές των ξενοδοχείων φιλτράρονται με βάση τις προεπιλογές του χρήστη στα ON/OFF κατ' αποκοπή κριτήρια, σχετικά με τα χαρακτηριστικά των ξενοδοχείων, που θεωρεί απαραίτητα για την τελική απόφαση του.

Οι συναρτήσεις του excel που χρησιμοποιήθηκαν κυρίως για τη σύνθεση των αναλυτικών συναρτήσεων ώστε να καταλήξουμε στα αποτελέσματα των συστάσεων είναι οι παρακάτω:

IF(),ROW(),ROWS(),INDEX(),SMALL(),AVERAGE(),COUNTIFS(),ISERROR(),  
MATCH(),OR(),SUM(), VLOOKUP(), COUNTA(), COUNTBLANK().

Στο παράρτημα αναγράφονται οι συναρτήσεις που χρησιμοποιήθηκαν για κάθε διαδικασία.

## 5. Παρουσίαση του Συστήματος για μελέτη περίπτωσης: Τουριστικά Καταλύματα

### 5. Παρουσίαση του συστήματος για μελέτη περίπτωσης: Τουριστικά Καταλύματα

Το περιβάλλον του συστήματος δημιουργήθηκε με γλώσσα προγραμματισμού Visual Basic (VB) σε περιβάλλον Microsoft Visual Studio 2013/2017. Πιο συγκεκριμένα με χρήση VB έχει αναπτυχθεί το περιβάλλον διεπαφής με το χρήστη, για την εισαγωγή δεδομένων στο σύστημα, ο αλγόριθμος WAP, καθώς και οι συναρτήσεις επικοινωνίας του συστήματος με τις βάσεις των δεδομένων και τα spreadsheet όπου έχουν αναπτυχθεί οι μεθοδολογίες συστάσεων.

Οι βάσεις δεδομένων, έχουν διαχωριστεί σε τρεις: 1. τη βάση που περιλαμβάνει τα αρχικά δεδομένα που έχουν αντληθεί από το διαδίκτυο (Αξιολογήσεις χρηστών + Χαρακτηριστικά ξενοδοχείων), τα αποτελέσματα εφαρμογής MUSA και sentiment analysis για τα ξενοδοχεία, 2. Τη βάση όπου αποθηκεύονται τα δεδομένα εισόδου του χρήστη στην WAP μέθοδο (Κατάταξη κριτηρίων και  $z_{min}$ ,  $z_{max}$ ), 3. Τη βάση με τα δεδομένα εισόδου του χρήστη, όλες τις επιλογές του σχετικά με τα χαρακτηριστικά των ξενοδοχείων που προτιμά, τα αποτελέσματα συστάσεων που προκύπτουν από το σύστημα, καθώς και τα δεδομένα ανατροφοδότησης του συστήματος.

Η εφαρμογή τόσο της MUSA όσο και της sentiment analysis πραγματοποιούνται με ανεξάρτητα λογισμικά. Στο παρόν σύστημα ενσωματώνονται τα αποτελέσματα εφαρμογής των μεθόδων στα δεδομένα, σε spreadsheet. Το σύστημα αντλεί και επεξεργάζεται τα δεδομένα των spreadsheet με χρήση Visual Basic.

### Σελίδα Εισόδου (Login Page)

Το σύστημα αρχικά ξεκινάει με τη φόρμα εισόδου στοιχείων του νέου ή ήδη υπάρχοντα χρήστη (Σχήμα 5-1). Ο χρήστης έχει τη δυνατότητα να εγγραφεί είτε σαν administrator είτε σαν guest, εισάγοντας τα προσωπικά του στοιχεία και ένα username.

Το περιβάλλον για τον administrator είναι πιο σύνθετο και περιλαμβάνει τα αποτελέσματα συστάσεων με βάση και τις 4 μεθοδολογίες που προαναφέρθηκαν καθώς και περαιτέρω στατιστικά στοιχεία από τη λειτουργία του συστήματος. Το περιβάλλον για τον χρήστη guest είναι απλό και στόχο έχει την εισαγωγή ελάχιστης πληροφορίας καθώς και παραγωγή της βέλτιστης σύστασης ξενοδοχείου/ων που να ικανοποιούν τις ανάγκες του.

Σχ. 5-1 Σελίδα Εισόδου του συστήματος συστάσεων

Στην περίπτωση που ο χρήστης δεν έχει προηγουμένως εγγραφεί στο σύστημα, του ζητείται να υποβάλλει το όνομα, το επώνυμο του καθώς και ένα username. Σε περίπτωση που προϋπάρχει το username με διαφορετικά δεδομένα ονόματος και επώνυμου, του ζητείται να υποβάλει ένα διαφορετικό username ή να αλλάξει

τα στοιχεία του. Επίσης ο χρήστης καλείται να υποβάλει το τύπο χρήστη που θέλει να έχει (administrator/guest). Στην περίπτωση που θέλει να εγγραφεί με δικαιώματα administrator, καλείται να υποβάλει ένα μυστικό κωδικό που μπορεί να παραλάβει από τον διαχειριστή του συστήματος.

Στην περίπτωση που ο χρήστης έχει προηγουμένως εγγραφεί στο σύστημα, του ζητείται μόνο να υποβάλει το username του και να πατήσει το κουμπί 'log in', προκειμένου να πλοηγηθεί στην επόμενη οθόνη. Αν το username δεν υπάρχει στη βάση δεδομένων, του προβάλλεται μήνυμα για λάθος username και καλείται να το επανυποβάλει στο σχετικό πλαίσιο. Αν το username προϋπάρχει στη βάση, τότε του προβάλλεται μήνυμα με τα ξενοδοχεία τα οποία έχει επιλέξει σαν τελική του απόφαση στις προηγούμενες επισκέψεις του στο σύστημα.

Παρακάτω εμφανίζεται η σχετική οθόνη (Σχήμα 5-2) για χρήστη ο οποίος έχει προηγουμένως εγγραφεί και χρησιμοποιήσει το σύστημα, έχοντας το τέλος εισάγει στην ανατροφοδότηση την επιλογή του ξενοδοχείου στην οποία κατέληξε κατόπιν της σύστασης.

Σχ. 5-2 Διαδικασία login στο σύστημα συστάσεων

Σελίδα ON/OFF filtering

Η πρώτη σελίδα μετάβασης μετά το επιτυχημένο login του χρήστη είναι η σελίδα με τις ON/OFF επιλογές (Σχήμα 5-3). Πιο συγκεκριμένα στη συγκεκριμένη οθόνη εμφανίζονται όλα τα κριτήρια αποκλεισμού επιλογών ξενοδοχείων, όπου γίνεται το αρχικό φιλτράρισμα.

Εδώ παρέχονται στο χρήστη 25 κριτήρια (23 με επιλογές: Yes, No, No choice και 2 κριτήρια με περισσότερες επιλογές: Stars: 1,2,3,4,5 Stars, No Choice και Price Range: 0-50,51-100,101-150,151-200,>200 euros, No choice). Στο πάνω μέρος της οθόνης, εμφανίζεται ο αριθμός των διαθέσιμων επιλογών ξενοδοχείων, ανάλογα με τις επιλογές του χρήστη.

Ο χρήστης ανάλογα με τις προτιμήσεις του είναι δυνατόν να επιλέξει μία ή περισσότερες επιλογές κριτηρίων, τα οποία θεωρεί προαπαιτούμενα για τη διαμονή του σε ένα ξενοδοχείο. Ουσιαστικά με κάθε επιλογή του αποκλείονται τα ξενοδοχεία που καλύπτουν τα συγκεκριμένα κριτήρια/χαρακτηριστικά. Επίσης του δίνεται η δυνατότητα να καθαρίσει τις επιλογές του και να επανεκκινήσει τη διαδικασία, επιλέγοντας το κουμπί 'clear choices'. Αφού τελειώσει με τη διαδικασία επιλογής, ο χρήστης καλείται να επιλέξει την επιλογή 'I finished with ON/OFF Criteria', προκειμένου να μεταβεί στην επόμενη οθόνη. Παρακάτω οι σχετικές οθόνες (Σχήματα 5-3, 5-4).

The screenshot shows a web application window titled 'Form1' with a tab 'On/OFF Input'. At the top, there is a text input for 'Number of Hotels' and two buttons: 'Clear Choices' and 'I finished with ON/OFF Criteria'. Below this, there are 25 dropdown menus arranged in a grid, representing different hotel criteria: Stars, Free WiFi, Family rooms, Non-smoking rooms, Price Range (Euros), Air Conditioning, Heating, Non-smoking Facilities for..., Restaurant, Bar, Massage, Indoor pool, Outdoor pool, Room service, Lift, English, German, Russian, French, and Spanish. A modal dialog box is open in the center, displaying the following text: 'Please choose a value for the ON/OFF criteria that concern you, using the dropdown lists. The number of Hotels after each choice is depicted in the top of the page. You can clear you choices using the 'Clear Choices' Button. When you finish with ON/OFF choices please select the 'I finished with On/OFF Criteria' Button.' with an 'OK' button at the bottom.

Σχ. 5-3 Σελίδα ON/OFF filtering

Σχ. 5-4 Επιλογές τιμών για τα κατ' αποκοπή κριτήρια

#### Σελίδα ON/OFF Results

Στη σελίδα ON/OFF Results (Σχήμα 5-5), εμφανίζονται τα ονόματα όλων των ξενοδοχείων που έχουν περάσει από το αρχικό φιλτράρισμα με βάσει τις επιλογές του χρήστη, καθώς και οι επιλογές του στα κριτήρια που έχει επιλέξει στην προηγούμενη σελίδα.

Με αυτόν τον τρόπο του δίνεται η δυνατότητα να έχει μια πρώτη εικόνα των ξενοδοχείων που καλύπτουν τις προτιμήσεις αποκλεισμού του. Επίσης του υπενθυμίζονται οι επιλογές του.

Η συγκεκριμένη μέθοδος φιλτραρίσματος επιλογών είναι η πιο ευρέως χρησιμοποιούμενη στις υπάρχουσες διαδικτυακές πλατφόρμες κρατήσεων ή/και αξιολογήσεων τουριστικών καταλυμάτων και ξενοδοχείων. Εν συνεχεία οι διαδικτυακές πλατφόρμες, παρουσιάζουν στο χρήστη το σύνολο των εναπομεινάντων επιλογών ξενοδοχείων χρησιμοποιώντας συνήθως μια μέθοδο κατάταξης από τις εξής:

1. Με φθίνουσα ή αύξουσα σειρά, τα πιο δημοφιλή, δηλαδή αυτές με τις περισσότερες κρατήσεις ή προβολές σελίδας.



2. Με φθίνουσα ή αύξουσα σειρά ανάλογα με την κατηγορία αστεριών στην οποία ανήκουν.
3. Με φθίνουσα ή αύξουσα σειρά ανάλογα με την βαθμολογία των επισκεπτών.
4. Με φθίνουσα ή αύξουσα σειρά με βάση την τιμή.

Στο παρόν σύστημα ο τρόπος τελικής κατάταξης καθορίζεται από το προτιμησιακό μοντέλο όπως έχει προαναφερθεί παραπάνω, συνεπώς δεν προβαίνουμε σε μια τέτοιου είδους κατάταξη.

Για τη μετάβαση στην επόμενη σελίδα ο χρήστης καλείται να επιλέξει το κουμπί 'Move to preference modeling'.

Form1

Start Page | On/OFF Input | **ON/OFF Results** | Preferences

**Filtering Results**

Best Western Candia Hotel	Omega Platani
Sun City	Sonio Beach A
Top Hotel	Theros Hotel
Epimenidis Hotel	Corissia Prince
Alkion	Oasis Guestho
Poseidon Apartments	Akasti Hotel
Mirtilos Studios & Apartments	Alkion Hotel
Olive Tree	Athina Beach I
Zorbas Beach Village Hotel	Balito
Nireas Hotel	Ledra Maleme
Nontas Hotel Apartments	Menia Beach I
Hotel Elotia	Minos Village
Elotis Suites	Popi Hotel Ap
Minerva Dore	Porto Kalamal
Anniko	Renieris Hotel
Marianthi Studios	Sea View Apar
Aloni Suites	Athina Beach I
Anthemis	Rodon Hotel
Areti Aparthotel	Zeus Village
Mythos Platanias Apartments	Perla Beach
Blue Dome Hotel	Amalthia Beac

Filtering Results are based on the Following Choices

**Yes**

No ChoiceOutdoor pool

**No**

Move to Preference Modeling

Σχ. 5-5 Σελίδα παρουσίασης αποτελεσμάτων εναλλακτικών μετά το On/ Off Filtering

### Σελίδα Preferences

Στη σελίδα Preferences (Σχήμα 5-6) ο χρήστης καλείται να κατατάξει τα πέντε κριτήρια: Καθαριότητα (Cleanliness), Εξυπηρέτηση (Service), Άνεση (Comfort), Κατάσταση (Condition), Γειτονιά (Neighborhood), ανάλογα με τις προτεραιότητες που έχει ο ίδιος όταν διαμένει σε ένα κατάλυμα. Δίνεται η δυνατότητα να επιλέξει για κάθε κριτήριο τη σειρά κατάταξης του σε σχέση με τα υπόλοιπα με το 1<sup>ο</sup> να είναι το πιο σημαντικό και το 5<sup>ο</sup> το λιγότερο σημαντικό για τον ίδιο.

Form1

Start Page On/OFF Input ON/OFF Results Preferences

Cleanliness Service Comfort Condition Neighbourhood

1st 1st 1st 1st 1st

2nd 2nd 2nd 2nd 2nd

3rd 3rd 3rd 3rd 3rd

4th 4th 4th 4th 4th

5th 5th 5th 5th 5th

---

Clear All Choices

By pressing one button for each criterion you actually rank it among the other five criteria. The ranking order shows your preference concerning the importance of these criteria when you decide for a hotel.

Please choose your ranking of Criteria and press 'Continue'. After that, please select a choice from the dropdown lists for each pair of criteria. You can always clear your choices by pressing 'Clear All Choices'.

OK

Recommend me

VS VS VS VS

By selecting a choice from the dropdown lists, next to every comparison, you express your own opinion concerning the difference of importance of consecutive criteria that have resulted from your ranking above. Very weak preference means that you consider the consecutive criteria almost the same important. Very strong preference means that you consider the first criterion much more important than the next one.

Σχ. 5-6 Σελίδα εισαγωγής επιλογών κατάταξης και διαφοράς σημαντικότητας κριτηρίων

Επιλέγοντας τη σειρά κατάταξης για ένα κριτήριο, ταυτόχρονα απενεργοποιείται η επιλογή ίδιας κατάταξης στα υπόλοιπα κριτήρια και εμφανίζεται ο αριθμός κατάταξης σε πλαίσιο κάτω από το κάθε κριτήριο (Σχήμα 5-7). Επίσης του δίνεται η δυνατότητα με την επιλογή 'Clear all Choices', να αρχικοποιήσει την κατάταξη του, και να προβεί εκ νέου σε διαφορετική κατάταξη για τα κριτήρια. Εν συνεχεία ο χρήστης καλείται να επιλέξει το κουμπί 'Continue'.

Form1

Start Page | On/OFF Input | ON/OFF Results | Preferences

**Cleanliness** **Service** **Comfort** **Condition** **Neighbourhood**

1st 1st 1st 1st 1st  
 2nd 2nd 2nd 2nd 2nd  
 3rd 3rd 3rd 3rd 3rd  
 4th 4th 4th 4th 4th  
 5th 5th 5th 5th 5th  
 3rd 1st 4th 2nd 5th

Clear All Choices Continue

VS VS VS VS

By pressing one button for each criterion you actually rank it among the other five criteria. The ranking order shows your preference concerning the importance of these criteria when you decide for a hotel.

By selecting a choice from the dropdown lists, next to every comparison, you express your own opinion concerning the difference of importance of consecutive criteria that have resulted from your ranking above. Very weak preference means that you consider the consecutive criteria almost the same important. Very strong preference means that you consider the first criterion much more important than the next one.

Recommend me

Σχ. 5-7 Κατάταξη κριτηρίων και μετάβαση στην επιλογή διαφοράς σημαντικότητας

Με την επιλογή του κουμπιού 'Continue', ο κέρσορας μεταφέρεται στο κάτω μέρος της οθόνης. Ταυτόχρονα στα μέχρι πρότινος άδεια πλαίσια, εμφανίζονται τα κριτήρια με τη σειρά κατάταξης που έχει ορίσει ο χρήστης. Πιο συγκεκριμένα εμφανίζονται ως εξής: 1<sup>st</sup> vs 2<sup>nd</sup> κριτήριο, 2<sup>nd</sup> vs 3<sup>rd</sup>, 3<sup>rd</sup> vs 4<sup>th</sup>, 4<sup>th</sup> vs 5<sup>th</sup>. Δίπλα σε καθεμιά από τις συγκρίσεις υπάρχει μια dropdown λίστα επιλογών. Ο χρήστης καλείται να επιλέξει για κάθε ζεύγος κριτηρίων, όπως έχουν προκύψει από την κατάταξη, της διαφορά σημαντικότητας των κριτηρίων κατά την προτίμηση τους.

Πολύ ασθενής προτίμηση (very weak preference) σημαίνει ότι ο χρήστης θεωρεί τα διαδοχικά κριτήρια σχεδόν το ίδιο σημαντικά, ενώ πολύ ισχυρή προτίμηση (Very Strong preference), ότι θεωρεί το πρώτο κριτήριο ως πολύ πιο σημαντικό από το επόμενο. Υπάρχουν 5 διαβαθμίσεις προτίμησης, και ο χρήστης μπορεί να επιλέξει 1 από τις 5 για κάθε ζεύγος. Οι διαβαθμίσεις είναι οι εξής:

Πολύ ασθενής προτίμηση (Very weak preference), Ασθενής προτίμηση (Weak preference), Προτίμηση (Preference), Ισχυρή προτίμηση (Strong preference), Πολύ ισχυρή προτίμηση (Very Strong preference).

Μετά τις επιλογές στις dropdown λίστες, ο χρήστης καλείται να επιλέξει το κουμπί 'Recommend Me'. (Σχήμα 5-8)

Form1

Start Page | On/OFF Input | ON/OFF Results | Preferences

**Cleanliness** **Service** **Comfort** **Condition** **Neighbourhood**

1st 1st 1st 1st 1st  
 2nd 2nd 2nd 2nd 2nd  
 3rd 3rd 3rd 3rd 3rd  
 4th 4th 4th 4th 4th  
 5th 5th 5th 5th 5th  
 3rd 1st 4th 2nd 5th

Clear All Choices Continue

Service VS Condition Preference  
 Condition Cleanliness  
 Cleanliness VS Comfort Strong Preference  
 Comfort Neighbourhood

By pressing one button for each criterion you actually rank it among the other five criteria. The ranking order shows your preference concerning the importance of these criteria when you decide for a hotel.

By selecting a choice from the dropdown lists, next to every comparison, you express your own opinion concerning the difference of importance of consecutive criteria that have resulted from your ranking above. Very weak preference means that you consider the consecutive criteria almost the same important. Very strong preference means that you consider the first criterion much more important than the next one.

Recommend me

Σχ. 5-8 Επιλογή διαφορών σημαντικότητας μεταξύ συνεχόμενων κριτηρίων

### Σελίδα Analytical Recommendations

Στην επόμενη σελίδα εμφανίζονται τα αποτελέσματα των 4 διαφορετικών μεθόδων σύστασης (Σχήματα 5-9, 5-10). Η σελίδα αυτή είναι ορατή μόνο σε χρήστες με δικαίωμα administrator.

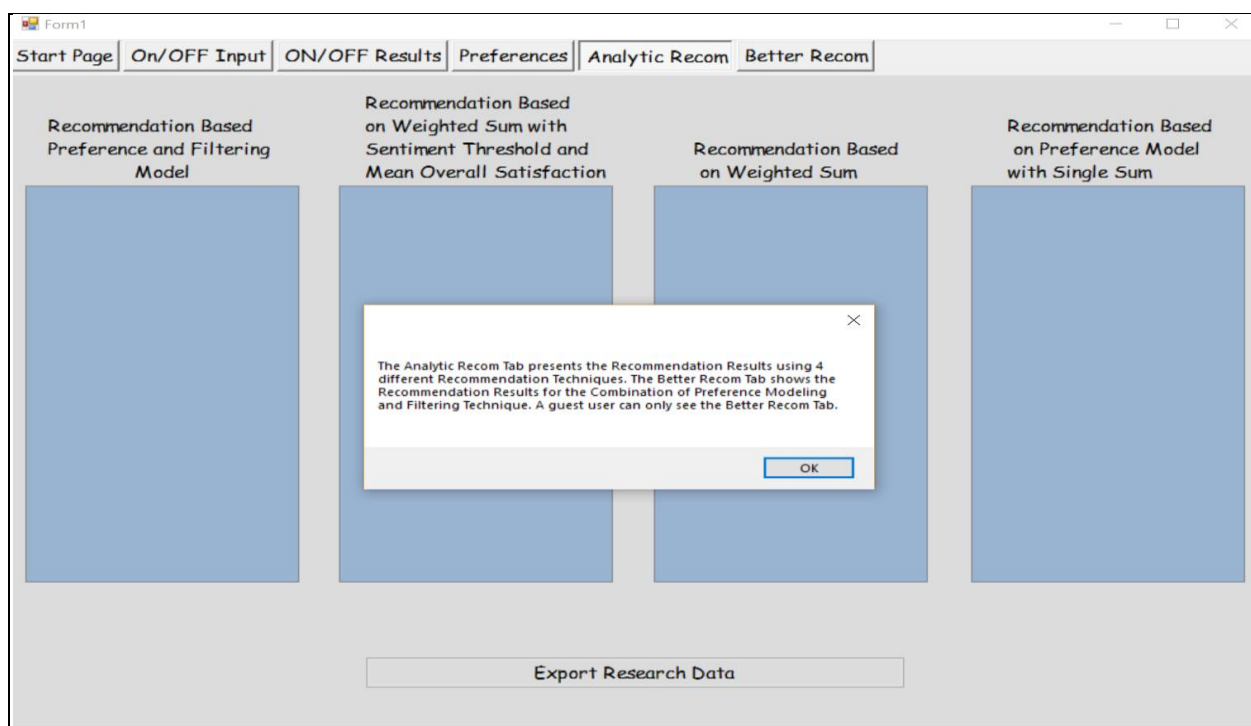
Στο πρώτο πλαίσιο από αριστερά εμφανίζονται οι Συστάσεις με βάση το Προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα, το κατώφλι συναισθήματος και το μοντέλο φιλτραρίσματος (Recommendation Based on Preference Model, Weighted Sum, Sentiment Threshold and Filtering Model).

Στο επόμενο πλαίσιο αμέσως δεξιά οι συστάσεις με βάση το Προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα και το κατώφλι συναισθήματος

(Recommendation Based on Preference model, Weighted Sum and Sentiment Threshold)

Στο τρίτο πλαίσιο βρίσκονται οι συστάσεις με βάση το προτιμησιακό μοντέλο και το σταθμισμένο άθροισμα (Recommendation Based on Preference model and Weighted Sum) και στο τέταρτο από αριστερά οι συστάσεις με βάση το προτιμησιακό μοντέλο με απλό άθροισμα (Recommendation Based on Preference Model with Single Sum).

Στις επόμενες δύο οθόνες παρουσιάζεται ο τρόπος εμφάνισης των αποτελεσμάτων συστάσεων (Σχήματα 5-9, 5-10).



Σχ. 5-9 Σελίδα παρουσίασης αποτελεσμάτων των 4 τύπων συστάσεων για χρήστη admin

Σχ. 5-10 Παρουσίαση συστάσεων

### Σελίδα Better Recom

Η σελίδα Better Recom είναι η σελίδα παρουσίασης των αποτελεσμάτων συστάσεων για τους χρήστες με δικαιώματα guest στο σύστημα (Σχήμα 5-11). Εμφανίζεται επίσης και στους χρήστες με δικαιώματα administrator.

Τα αποτελέσματα της σύστασης προκύπτουν από τη μέθοδο με βάση το Προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα, το κατώφλι συναισθήματος και το μοντέλο φιλτραρίσματος (Recommendation Based on Preference Model, Weighted Sum, Sentiment Threshold and Filtering Model). Στο περιβάλλον του χρήστη εμφανίζονται οι 5 πρώτες συστάσεις όπως προκύπτουν από την κατάταξη. Στην παρούσα οθόνη εμφανίζονται δύο συστάσεις, καθώς όπως είναι προφανές και από τις προηγούμενες οθόνες, τα αποτελέσματα συστάσεων ήταν 2 για τη συγκεκριμένη μέθοδο σύστασης.

Στο κάτω μέρος της σελίδας, δίνεται η δυνατότητα στο χρήστη, αναγράφοντας τα πρώτα γράμματα του ξενοδοχείου, να πάρει περισσότερες πληροφορίες σχετικά με τον αριθμό αστεριών του, τον αριθμό των κριτικών που υπάρχουν στη βάση δεδομένων και στις οποίες στηρίζεται η σύσταση, τη μέση συνολική

αξιολόγηση του, τη μέση τιμή του καθώς και την απόσταση από το κέντρο της πόλης και από το αεροδρόμιο.

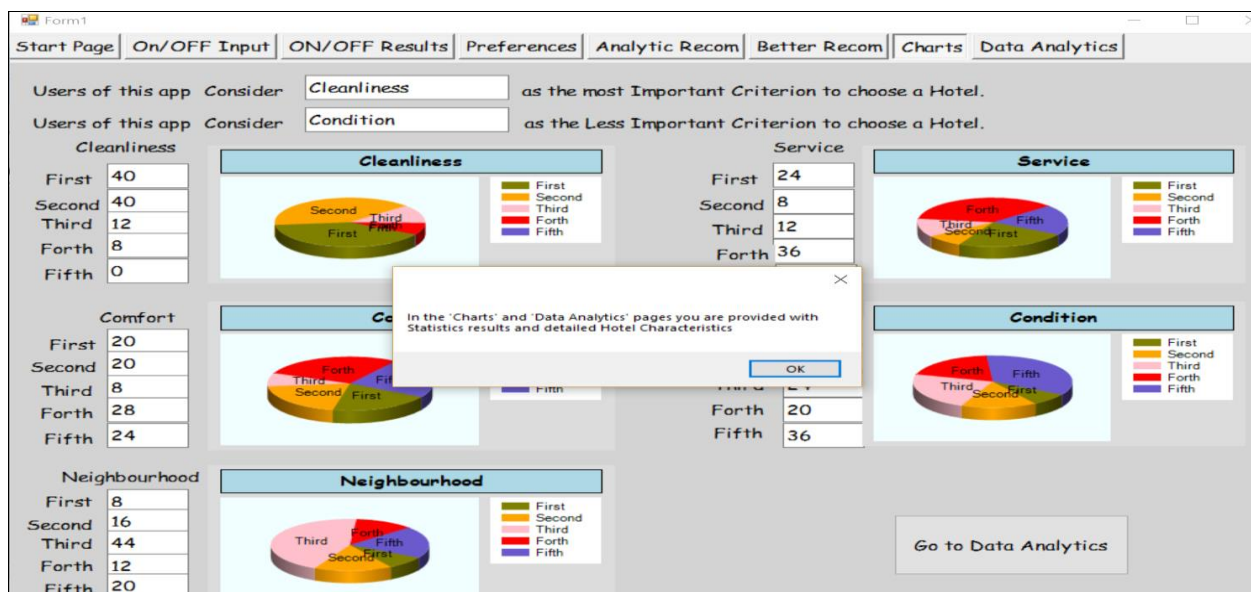
Εν συνεχεία ο χρήστης καλείται να επιλέξει το κουμπί 'Go to charts'.

Σχ. 5-11 Σελίδα παρουσίασης αποτελεσμάτων συστάσεων Prec και πληροφοριών εναλλακτικής

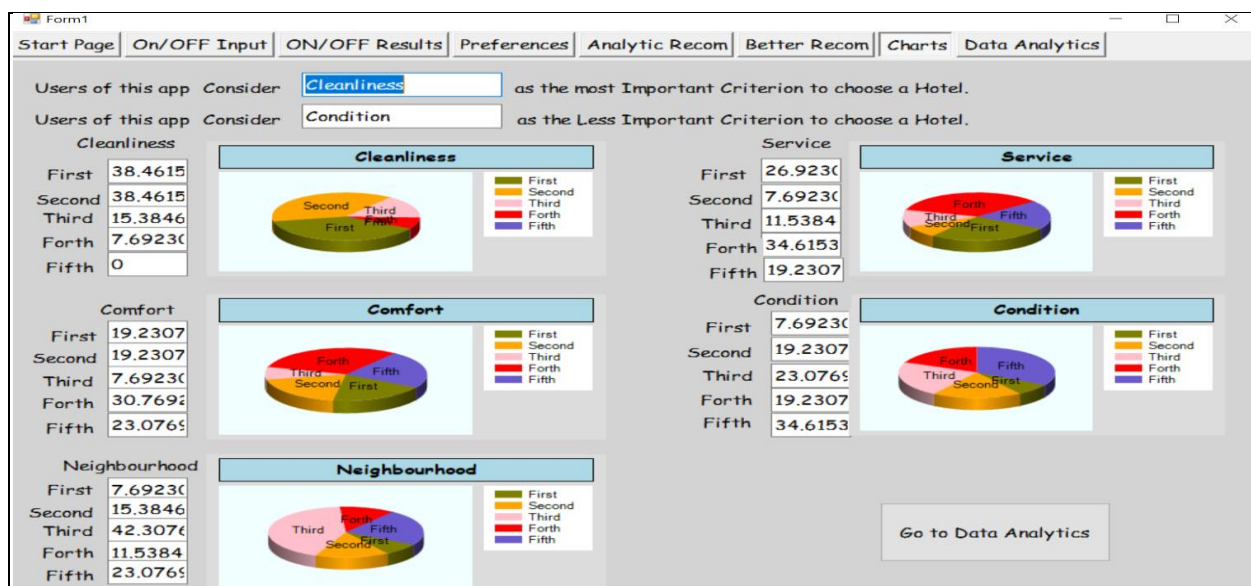
### Σελίδα Charts

Στη σελίδα Charts εμφανίζονται στατιστικά δεδομένα που προκύπτουν από τις επιλογές όλων των χρηστών του συστήματος. Πιο συγκεκριμένα παρουσιάζεται το κριτήριο που έχει επιλεγεί τις περισσότερες φορές από τους χρήστες σαν το πιο σημαντικό καθώς και αυτό που έχει επιλεγεί σαν το λιγότερο σημαντικό. Επίσης εμφανίζονται διαγράμματα όπου παρουσιάζεται ποσοστιαία, για κάθε κριτήριο η σειρά της κατάταξης του για όλες τις χρήσεις του συστήματος. Στην πρώτη οθόνη (Σχήμα 5-12), εμφανίζονται τα ποσοστά πριν την τελευταία χρήση του συστήματος. Στην δεύτερη οθόνη (Σχήμα 5-13) παρατηρούμε την αλλαγή που συντελείται στα ποσοστά, λόγω των επιλογών του τελευταίου χρήστη.





Σχ. 5-12 Σελίδα με charts



Σχ. 5-13 Σελίδα με charts με αποτελέσματα

### Σελίδα Data Analytics

Η σελίδα Data Analytics (Σχήμα 5-14) είναι προσβάσιμη σε χρήστες με δικαίωμα administrator. Εδώ δίνεται η δυνατότητα στο χρήστη να προσπελάσει όλα τα δεδομένα που είναι διαθέσιμα για ένα ξενοδοχείο, καθώς και τα αριθμητικά αποτελέσματα της ανάλυσης συναισθήματος και ανάλυσης ικανοποίησης πελατών MUSA, που είναι καταχωρημένα στη βάση δεδομένων.



The screenshot shows a web application window titled 'Form1'. It has a navigation bar with tabs: 'Start Page', 'On/OFF Input', 'ON/OFF Results', 'Preferences', 'Analytic Recom', 'Better Recom', 'Charts', and 'Data Analytics'. The main content area is titled 'Please choose a hotel from the List Below to show its characteristics and Analysis Results'. A dropdown menu shows 'Canea Mare Hotel'. Below this, there are six input fields for hotel characteristics: 'Stars' (3Stars), 'Number of Reviews' (50), 'Mean Overall Rating' (4), 'Average Price' (180), 'Distance from Center' (5.4), and 'Distance from Airport' (12). The 'Sentiment Analysis Result' section shows a value of 0.435 and a table of sentiment ranges: Negative Sentiment (From: -2 To: -0.05), Neutral Sentiment (From: -0.05 To: 0.22), and Positive Sentiment (From: 0.22 To: 2). The 'Multicriteria Satisfaction Analysis Results' section shows a 'Global Satisfaction index' of 91.817, a 'Preference threshold' of 2, and a 'Post optimality threshold' of 7. Below these are five weight fields: 'Cleanliness Weight' (10.128), 'Service Weight' (39.744), 'Comfort Weight' (20), 'Condition Weight' (10.128), and 'Neighbourhood Weight' (20). A 'Go to Feedback' button is at the bottom right.

Σχ. 5-14 Σελίδα με αναλυτικά δεδομένα για κάθε εναλλακτική

### Σελίδα Feedback

Στη σελίδα Feedback (Σχήμα 5-15) ο χρήστης του συστήματος καλείται να προβεί σε ανατροφοδότηση του συστήματος και δηλώσει ποιο ξενοδοχείο επέλεξε. Έτσι, δημιουργείται ιστορικό στη βάση δεδομένων για τον συγκεκριμένο χρήστη και σε κάθε νέα του επίσκεψη του προσφέρεται άμεση ενημέρωση για όλες τις επιλογές του και αποφάσεις του σε προηγούμενες επισκέψεις. Αυτό τον βοηθάει να αποφασίσει ακόμα πιο γρήγορα στην επόμενη επίσκεψη του. Επίσης του δίνεται η δυνατότητα να αξιολογήσει τη πλατφόρμα καθώς και να γράψει σχετικό σχόλιο/κριτική για αυτή. Στην παράγραφο 3.7 περιγράφεται αναλυτικά η διαδικασία ανατροφοδότησης στο παρόν μεθοδολογικό πλαίσιο.

Στο κάτω μέρος της σελίδας υπάρχει επιλογή κουμπιού για νέα είσοδο από χρήστη στο σύστημα ή για έξοδο από αυτό.

Form1

Start Page | On/OFF Input | ON/OFF Results | Preferences | Analytic Recom | Better Recom | Charts | Data Analytics | Feedback

Please Help us improve our Recommendations by filling the following choices

Choose from the list below the Hotel that you preferred at last

Balos Beach Hotel

You can start writing the hotel name and by pressing the anchor it will appear in the list

Was this App helpful to your decision for a Hotel? ☒ Yes ☐ No

Please write any comments you have about the App in the textbox below

Great app

Different User Login Exit App

Σχ. 5-15 Σελίδα για ανατροφοδότηση από το χρήστη

## 6. Αξιολόγηση αποτελεσμάτων

### 6. Αξιολόγηση αποτελεσμάτων

Το σύστημα συστάσεων που υλοποιήθηκε αποτελεί ένα υβριδικό σύστημα που συνδυάζει Multi-criteria rating-based preference elicitation με filtering. Με δεδομένη την συνεχώς αυξανόμενη τάση για ενσωμάτωση των κριτικών σαν αναπόσπαστο κομμάτι όλων των διαδικτυακών εφαρμογών για αγορά προϊόντων, η συγκεκριμένη ερευνητική εργασία αναμένεται να προσφέρει μια σημαντική λύση στην παραγωγή αποτελεσματικών συστάσεων που στηρίζονται αποκλειστικά και μόνο σε δεδομένα κριτικών σεβόμενη ταυτόχρονα τα προσωπικά δεδομένα του χρήστη.

Για την επαλήθευση της ποιότητας των συστάσεων που προκύπτουν από το σύστημα, εφαρμόστηκαν δύο διαφορετικές διαδικασίες: 1. Εφαρμογή του συστήματος από τουριστικό γραφείο για δύο συναπτά έτη για πελάτες που είχαν σαν προορισμό τα Χανιά (100), 2. Εφαρμογή του συστήματος με χρήστες 50 φυσικά πρόσωπα που είχαν στο παρελθόν επισκεφτεί καταλύματα στα Χανιά.

1. Κατά την εφαρμογή του συστήματος, σε ταξιδιωτικό γραφείο, ζητήθηκε σε σύνολο 100 πελατών του, που είχαν προορισμό τα Χανιά, να τρέξουν το σύστημα πριν αποφασίσουν για το ξενοδοχείο που θα διέμεναν κατά τις διακοπές τους.

Εν συνεχεία ζητήθηκε από τους πελάτες που τελικά επέλεξαν ένα από τα καταλύματα που αποτελούσε σύσταση για να επισκεφθούν, να αξιολογήσουν την εμπειρία τους βαθμολογώντας το κατάλυμα με βάση τα 5 κριτήρια και μια συνολική βαθμολογία, σε κλίμακα 1-5, όπου 1 τη χειρότερη και 5 η καλύτερη βαθμολογία.

Το όφελος της χρήσης της συγκεκριμένης διαδικασίας για τεκμηρίωση λειτουργίας του συστήματος ήταν ότι απευθυνθήκαμε σε πραγματικούς χρήστες

που ενδιαφέρονταν να διερευνήσουν τις επιλογές τους για κατάλυμα, ήταν δυνατό να γίνει σε μαζικό επίπεδο πελατών, και επίσης ότι ήταν δυνατή η διαδικασία ανατροφοδότησης σχετικά με την τελική επιλογή καταλύματος.

Ένα περιορισμός που προέκυψε στη συγκεκριμένη διαδικασία ήταν ότι το ταξιδιωτικό γραφείο συνεργαζόταν με περιορισμένο αριθμό ξενοδοχείων από το συγκεκριμένο προορισμό, συνεπώς από την τελική σύσταση αφαιρούνταν οι επιλογές που δεν αποτελούσαν συνεργάτες.

Από το σύνολο 166 καταλυμάτων στην βάση του συστήματος, το τουριστικό γραφείο συνεργαζόταν με τα 89 (54% των ξενοδοχείων).

Από τους 100 πελάτες, οι 73 αποφάσισαν να κάνουν κράτηση σε ξενοδοχείο που αποτελούσε σύσταση του συστήματος. Συνεπώς προκύπτει ένα ποσοστό 73% των συστάσεων του συστήματος που έτυχαν της αποδοχής και προτίμησης των χρηστών. Το συγκεκριμένο ποσοστό πελατών, κρίνεται πολύ υψηλό όσο αφορά την ακρίβεια των συστάσεων σε σχέση με το προτιμησιακό προφίλ των χρηστών.

Οι 73 πελάτες που επέλεξαν κάποια από τις συστημένες επιλογές ξενοδοχείου, μετά το τέλος των διακοπών τους, κλήθηκαν να υποβάλουν τις βαθμολογίες τους στα κριτήρια καθώς και τη συνολική αξιολόγηση τους σε σχέση με την εμπειρία τους.

Προκειμένου να αξιολογήσουμε τα αποτελέσματα των βαθμολογιών των πελατών, τα συγκρίναμε με τις μέσες τιμές των βαθμολογιών πελατών για τα ξενοδοχεία, στα αντίστοιχα κριτήρια, που προέρχονται από την αρχική άντληση μεγάλου όγκου δεδομένων. Τα αποτελέσματα των βαθμολογιών παρουσιάζονται στο παράρτημα.

Όσον αφορά τη συνολική βαθμολογία για τα καταλύματα που επισκέφθηκαν οι 73 πελάτες προέκυψαν οι εξής μετρικές σφάλματος (πίνακας 6-1):

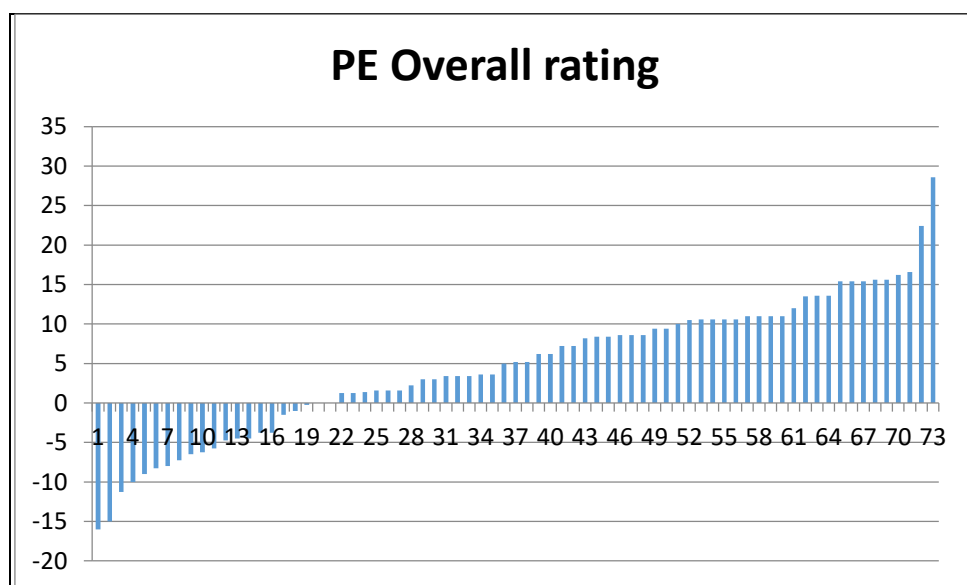
Πίνακας 6-1 Αποτελέσματα Μετρικών σφάλματος για συνολική βαθμολογία

MSE	RMSE	MAE	MAPE
0.077819	0.2789608	0.3531507	7.589726%

Τα αποτελέσματα των μετρικών σφάλματος (πίνακας 6-1) αποδεικνύουν ότι οι αξιολογήσεις των πελατών ήταν πάρα πολύ κοντά στις μέσες συνολικές

αξιολογήσεις των ξενοδοχείων, που προέκυψαν από τα δεδομένα που είχαν αντληθεί στο διαδίκτυο. Το συγκεκριμένο γεγονός υποδεικνύει ότι οι συστάσεις ήταν ακριβείς όσον αφορά την ταύτιση των προτιμησιακών μοντέλων των χρηστών με τα προφίλ των προϊόντων (ξενοδοχείων).

Στη γραφική 6-1 απεικονίζονται οι τιμές ποσοστιαίου σφάλματος για κάθε μια από τις αξιολογήσεις που δόθηκαν από τους χρήστες για τη συνολική βαθμολογία ξενοδοχείων.



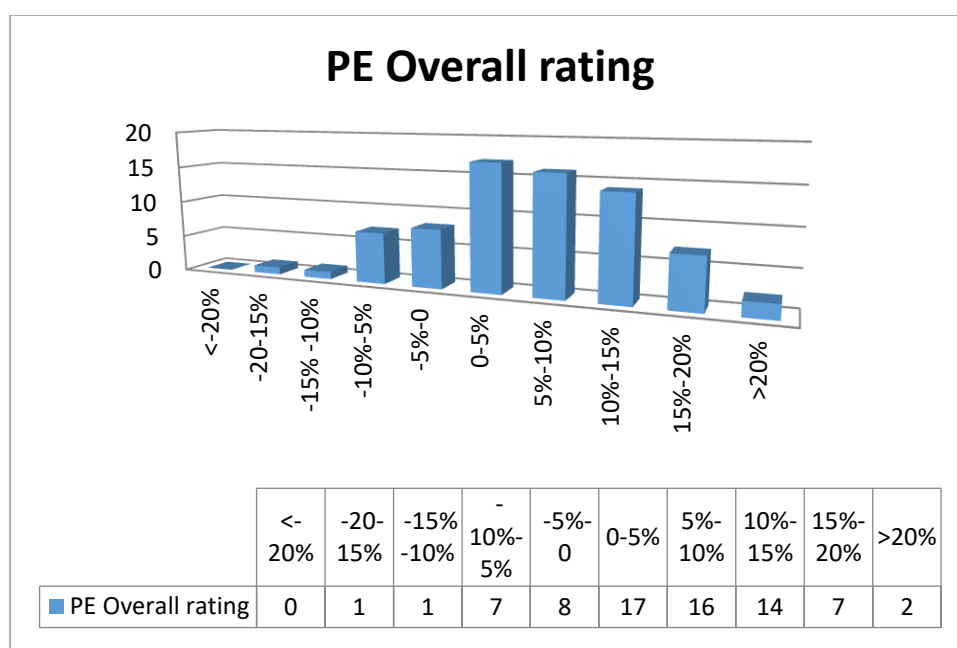
Γραφική 6-1 Ποσοστιαία σφάλματα μεταξύ βαθμολογιών χρήστη και ξενοδοχείου

Με βάση τις γραφικές παραστάσεις είναι εμφανές ότι μόνο 17 χρήστες από τους 73 (23%) βαθμολόγησαν χαμηλότερα από τη μέση συνολική βαθμολογία τα ξενοδοχείο που τους συστήθηκαν και αυτοί τα επισκέφθηκαν. Άρα το 77% των χρηστών που βαθμολόγησαν το ξενοδοχείο τους παραπάνω από το μέσο όρο, είναι ένα άκρως ενθαρρυντικό ποσοστό όσον αφορά την αποτελεσματικότητα των συστάσεων και την ταύτιση των προτιμησιακών μοντέλων ξενοδοχείων και χρηστών.

Στον παρακάτω πίνακα (πίνακας 6-2) παρουσιάζεται το πλήθος των βαθμολογιών ανά εύρος 5% ποσοστιαίου σφάλματος, για κάθε ένα από τα κριτήρια, καθώς και για τη συνολική βαθμολογία.

Πίνακας 6-2 Πλήθος βαθμολογιών ανά εύρος ποσοστιαίων σφαλμάτων

range	PE Overall rating	PE Cleanlines	PE Service	PE Comfort	PE Condition	PE Neighbourhood
<-20%	0	2	0	3	3	6
-20-15%	1	1	2	0	2	1
-15%-10%	1	4	3	7	6	1
-10%-5%	7	5	6	6	12	10
-5%-0	8	2	4	9	7	3
0-5%	17	26	11	11	11	14
5%-10%	16	18	27	7	11	14
10%-15%	14	11	11	13	9	7
15%-20%	7	2	7	12	11	14
>20%	2	2	2	5	1	3



Γραφική 6-2 Σύγκριση βαθμολογιών από ανατροφοδότηση χρηστών με βαθμολογίες ξενοδοχείων

Σύμφωνα με την γραφική παράσταση 6.2 , που δείχνει το πλήθος των βαθμολογιών για κάθε εύρος τιμών ποσοστιαίου σφάλματος, το 70% βαθμολόγησαν το ξενοδοχείο από 0 έως 20% παραπάνω από το μέσο όρο βαθμολογίας, που σε 5-βάθμια κλίμακα είναι 0 έως 1 μονάδα παραπάνω.

Επιπρόσθετα είναι χαρακτηριστικό ότι από το 23% των χρηστών που βαθμολόγησαν την εμπειρία τους με κατώτερο βαθμό από το μέσο όρο, το 88% βαθμολόγησαν μέχρι 10% χαμηλότερα από το μέσο όρο τα ξενοδοχεία τους. Αυτό σημαίνει ότι για την 5-βάθμια κλίμακα που χρησιμοποιούμε στο παρόν πλαίσιο, βαθμολόγησαν το ξενοδοχείο με 0.5 λιγότερο.

Όσον αφορά τα αποτελέσματα για τα λοιπά κριτήρια παρατηρούμε τα εξής:

#### Καθαριότητα

Το 78% των πελατών βαθμολόγησαν το ξενοδοχείο από 0 έως 20% παραπάνω από το μέσο όρο βαθμολογίας.

Από το 20% των πελατών που βαθμολόγησαν το ξενοδοχείο που επισκέφθηκαν, με χαμηλότερη βαθμολογία από τη μέση τιμή, το 79% το βαθμολόγησαν από 0-15% λιγότερο. (Γραφικές 9.2, 9.8 στο παράρτημα)

#### Εξυπηρέτηση

Το 77% των πελατών βαθμολόγησαν το ξενοδοχείο από 0 έως 20% παραπάνω από το μέσο όρο βαθμολογίας.

Από το 21% των πελατών που βαθμολόγησαν το ξενοδοχείο που επισκέφθηκαν, με χαμηλότερη βαθμολογία από τη μέση τιμή, το 67% το βαθμολόγησαν από 0-10% λιγότερο, ενώ κανένας δεν βαθμολόγησε το κριτήριο αυτό με χαμηλότερη βαθμολογία από 20% λιγότερο από τη μέση τιμή. (Γραφικές 9.3, 9.9 στο παράρτημα)

#### Άνεση

Το 59% των πελατών βαθμολόγησαν το ξενοδοχείο από 0 έως 20% παραπάνω από το μέσο όρο βαθμολογίας.

Από το 34% των πελατών που βαθμολόγησαν το ξενοδοχείο που επισκέφθηκαν, με χαμηλότερη βαθμολογία από τη μέση τιμή, το 88% το βαθμολόγησαν από 0-15% λιγότερο.

Για το συγκεκριμένο κριτήριο παρατηρούμε μια συσσώρευση των τιμών ποσοστιαίου σφάλματος μεταξύ -15 και +20%. (Γραφικές 9.4, 9.10 στο παράρτημα)

#### Κατάσταση

Το 57% των πελατών βαθμολόγησαν το ξενοδοχείο από 0 έως 20% παραπάνω από το μέσο όρο βαθμολογίας.

Από το 41% των πελατών που βαθμολόγησαν το ξενοδοχείο που επισκέφθηκαν, με χαμηλότερη βαθμολογία από τη μέση τιμή, το 83% το βαθμολόγησαν από 0-15% λιγότερο.

Για το συγκεκριμένο κριτήριο παρατηρούμε μια συσσώρευση των τιμών ποσοστιαίου σφάλματος μεταξύ -15 και +20%. (Γραφικές 9.5, 9.11 στο παράρτημα)

#### Γειτονιά

Το 67% των πελατών βαθμολόγησαν το ξενοδοχείο από 0 έως 20% παραπάνω από το μέσο όρο βαθμολογίας.

Από το 29% των πελατών που βαθμολόγησαν το ξενοδοχείο που επισκέφθηκαν, με χαμηλότερη βαθμολογία από τη μέση τιμή, το 67% το βαθμολόγησαν από 0-15% λιγότερο. (Γραφικές 9.6, 9.12 στο παράρτημα)

Μελετώντας τα ποσοστιαία σφάλματα για καθένα από τα κριτήρια παρατηρούμε ότι οι βαθμολογίες των πελατών κυμαίνονται σε πλαίσια ανώτερα των μέσων τιμών βαθμολογιών για όλα. Αποδεικνύεται λοιπόν ότι το μεγαλύτερο ποσοστό των πελατών που χρησιμοποίησαν τις συστάσεις, μείνανε παραπάνω ικανοποιημένοι από τους πελάτες με παρόμοιο προτιμησιακό μοντέλο.

Πιο συγκεκριμένα τα κριτήρια καθαριότητα και εξυπηρέτηση δείχνουν να ξεπέρασαν τις προσδοκίες των περισσότερων, όπως ακριβώς συνέβει και για τα υπόλοιπα τρία κριτήρια αλλά σε λίγο μικρότερο βαθμό.

2. Εφαρμογή του συστήματος με χρήστες 50 φυσικά πρόσωπα που είχαν στο παρελθόν επισκεφτεί καταλύματα στα Χανιά και είχαν σε μεγάλο βαθμό



γνώση περί της ποιότητας τους. Οι συγκεκριμένοι χρήστες εισήγαγαν τις προτιμήσεις τους στο σύστημα και προέκυψαν οι αντίστοιχες συστάσεις ξενοδοχείων για καθένα από αυτούς. Εν συνεχεία τους ζητήθηκε, με βάση τη γνώση τους περί των ξενοδοχείων της περιοχής των Χανίων, αν οι συστάσεις ανταποκρίνονταν στις προτιμήσεις τους. Τα αποτελέσματα έδειξαν αρκετά υψηλή και επιτυχή στόχευση των συστάσεων στις ανάγκες των χρηστών. Πιο συγκεκριμένα μέσα από τη λειτουργία ανατροφοδότησης επαληθεύτηκε ότι σε ένα ποσοστό 82%, οι χρήστες επιβεβαίωσαν την προτίμηση τους για τουλάχιστο ένα από τα ξενοδοχεία που τους προτάθηκαν κατά τη διαδικασία σύστασης. Το συγκεκριμένο δείγμα χρηστών αποτελεί δείγμα ευκολίας, αλλά με δεδομένη τη δυσκολία αντικειμενικής τεκμηρίωσης της καλής λειτουργίας ενός συστήματος συστάσεων, μπορεί να θεωρηθεί σαν μια παραπάνω ένδειξη αποτελεσματικής και στοχευμένης σύστασης.

## 7. Συμπεράσματα

7.1. Περιορισμοί

7.2. Προτάσεις για μελλοντική έρευνα

### 7. Συμπεράσματα

Στη παρούσα διδακτορική διατριβή παρουσιάστηκε μια νέα μεθοδολογία υβριδικού συστήματος συστάσεων, όπου συνδυάστηκαν επιτυχώς μέθοδοι πολυκριτήριας ανάλυσης, μέθοδοι ανάλυσης συναισθήματος και φιλτραρίσματος σε μεγάλο όγκο δεδομένων για την παραγωγή συστάσεων προϊόντων.

Το σύστημα συστάσεων αποτελεί μια τεχνολογία φιλτραρίσματος πληροφορίας που εφαρμόζεται σε περιβάλλον όπου αντικείμενα προτείνονται σε χρήστες, επειδή είναι πιθανό να τους ενδιαφέρουν και τυπικά είναι τόσο μεγάλο το πλήθος τόσο των χρηστών όσο και των προϊόντων που το πρόβλημα γίνεται αρκετά δύσκολο και ακριβό να επιλυθεί. Τυπικά το σύστημα συστάσεων χρησιμοποιεί λεπτομέρειες από το προφίλ του χρήστη, απόψεις και συνήθειες της κοινότητας χρηστών και συγκρίνοντας την πληροφορία με τα συγκεκριμένα χαρακτηριστικά παρουσιάζει συστάσεις.

Προκειμένου να χτιστεί ένα σύστημα σύστασης απαιτείται να υπάρχει σύνολο δεδομένων αντικειμένων και χρηστών και ιδανικά αλληλεπιδράσεις μεταξύ τους. Υπάρχουν πολλοί τομείς εφαρμογής τους και τυπικά οι χρήστες είναι πελάτες, τα αντικείμενα προϊόντα και οι αλληλεπιδράσεις αγορές προϊόντων.

Τα συστήματα συστάσεων προσθέτουν αξία τόσο στους χρήστες όσο και στους παρόχους υπηρεσιών. Μειώνουν τα κόστη για έρευνα και επιλογή αντικειμένων σε ένα διαδικτυακό περιβάλλον και βελτιώνουν τη διαδικασία λήψης απόφασης καθώς και την ποιότητα της απόφασης. Σε επίπεδο παρόχου υπηρεσιών βελτιώνουν τα έσοδα, αυξάνοντας τις πωλήσεις προϊόντων τους. Επιπλέον ανοίγουν νέες προοπτικές για την παροχή προσωποποιημένης πληροφορίας στο διαδίκτυο, βοηθούν στη μείωση του προβλήματος της υπερφόρτωσης

πληροφοριών και δίνουν πρόσβαση στους χρήστες σε προϊόντα ή υπηρεσίες που σε μια πρώτη ανάγνωση δεν τους είναι διαθέσιμα.

Στο μεθοδολογικό πλαίσιο που περιγράφηκε, γίνεται εφικτός ο προσδιορισμός των δυναμικών προφίλ των χρηστών μέσω της χρήσης μιας παραλλαγμένης μορφής της μεθόδου WAP και των αντικειμένων ενδιαφέροντος τους, με χρήση της μεθόδου MUSA. Τα προφίλ προτιμήσεων των χρηστών αντιστοιχίζονται σε «προφίλ» προϊόντων/ υπηρεσιών που εμπεριέχουν τα χαρακτηριστικά που αποτελούν προτεραιότητα για τους συγκεκριμένους χρήστες και εν συνεχεία προτείνονται σε αυτούς. Δίνεται μεγάλη έμφαση στη χρήση της μεθόδου ανάλυσης συναισθήματος σε κείμενο κριτικών πελατών, τα αποτελέσματα της οποίας λειτουργούν σαν κατώφλι για την παραγωγή συστάσεων ενώ με αντίστοιχο τρόπο λειτουργεί και η μέθοδος φιλτραρίσματος με χρήση κατ' αποκοπή κριτηρίων.

Με τη χρήση της παρούσας μεθοδολογίας σύστασης προϊόντων, δίνεται η δυνατότητα τόσο στον απλό χρήστη να εντοπίσει σε μικρό χρόνο, χωρίς παραβίαση των προσωπικών του δεδομένων και αποτελεσματικά τα προϊόντα που ανταποκρίνονται στις ανάγκες του (προτιμησιακό του προφίλ), όσο και στις επιχειρήσεις παρόχους υπηρεσιών/προϊόντων να καταφέρουν να αποκωδικοποιήσουν τις ανάγκες των πελάτων τους εξατομικευμένα, να εντοπίσουν τα χαρακτηριστικά των προϊόντων τους που χρειάζεται να επαναπροσδιοριστούν για να είναι πιο ελκυστικά προς τους πελάτες, δαπανώντας λιγότερο χρόνο και επιτυγχάνοντας καλύτερη κατανομή των πόρων τους.

Για την εφαρμογή της μεθοδολογίας αναπτύχθηκε σύστημα σε γλώσσα προγραμματισμού VB το οποίο παρουσιάστηκε στο 5 κεφάλαιο. Η μελέτη περίπτωσης αφορά τα Ξενοδοχεία του νομού Χανίων. Τα δεδομένα που χρησιμοποιήθηκαν, αντλήθηκαν με μεθόδους εξόρυξης δεδομένων, από τη διαδικτυακή πλατφόρμα κρατήσεων ξενοδοχείων, [www.hotels.com](http://www.hotels.com). Τα κριτήρια που χρησιμοποιήθηκαν ήταν τα εξής: Καθαριότητα (Cleanliness), Εξυπηρέτηση (Service), Άνεση (Comfort), Κατάσταση (Condition), Γειτονιά (Neighbourhood), συνολικό σκορ (Overall Score) και έχουν βαθμολογηθεί σε 5-βάθμια κλίμακα, με

το 1 να είναι η χειρότερη και 5 η μεγαλύτερη βαθμολογία. Το σύνολο των προϊόντων (ξενοδοχείων) που αποτελούν τη βάση δεδομένων είναι 166 και το σύνολο των κριτικών που αντλήθηκαν ήταν 4618. Επίσης αντλήθηκαν δεδομένα για 27 κριτήρια από τα ξενοδοχεία, που αποτέλεσαν τα κατ' αποκοπή κριτήρια.

Για την επαλήθευση της καλής λειτουργίας των συστάσεων που προκύπτουν από το σύστημα, εφαρμόστηκαν δύο διαφορετικές διαδικασίες:

1. Εφαρμογή του συστήματος για πελάτες τουριστικού γραφείου που είχαν ως προορισμό τα Χανιά, για δύο συναπτά έτη (100 χρήστες), Από τους 100 πελάτες, οι 73 αποφάσισαν να κάνουν κράτηση σε ξενοδοχείο που αποτελούσε σύσταση του συστήματος. Οι 73 πελάτες που επέλεξαν κάποια από τις συστημένες επιλογές ξενοδοχείου, μετά το τέλος των διακοπών τους υπέβαλαν τη συνολική αξιολόγηση τους καθώς και τη βαθμολογία τους για τα κριτήρια σε σχέση με την εμπειρία τους.

Το 77% των χρηστών που βαθμολόγησαν το ξενοδοχείο τους παραπάνω από το μέσο όρο, αποδεικνύει ότι το σύστημα συστάσεων έκανε επιτυχή ταύτιση των προτιμησιακών μοντέλων ξενοδοχείων και χρηστών.

Επιπρόσθετα είναι χαρακτηριστικό ότι από το 23% των χρηστών που βαθμολόγησαν την εμπειρία τους με κατώτερο βαθμό από το μέσο όρο, το 88% βαθμολόγησαν μέχρι 10% χαμηλότερα από το μέσο όρο τα ξενοδοχεία τους. Αυτό σημαίνει ότι για την 5-βάθμια κλίμακα που χρησιμοποιούμε στο παρόν πλαίσιο, βαθμολόγησαν το ξενοδοχείο με 0.5 λιγότερο.

Μελετώντας τα ποσοστιαία σφάλματα για καθένα από τα κριτήρια παρατηρούμε ότι οι βαθμολογίες των πελατών κυμαίνονται σε πλαίσια ανώτερα των μέσων τιμών βαθμολογιών για όλα. Αποδεικνύεται λοιπόν ότι το μεγαλύτερο ποσοστό των πελατών που χρησιμοποίησαν τις συστάσεις, μείνανε παραπάνω ικανοποιημένοι από τους πελάτες με παρόμοιο προτιμησιακό μοντέλο.

Πιο συγκεκριμένα τα κριτήρια καθαριότητα και εξυπηρέτηση δείχνουν να ξεπέρασαν τις προσδοκίες των περισσότερων, όπως ακριβώς συνέβη και για τα υπόλοιπα τρία κριτήρια αλλά σε λίγο μικρότερο βαθμό.

Τα παραπάνω ποσοστά υποδηλώνουν ότι το μεθοδολογικό πλαίσιο που αναπτύχθηκε αποτελεί μια επιτυχημένη προσπάθεια και με περαιτέρω παρεμβάσεις ή δεδομένα θα μπορούσε να αυξήσει ακόμα περισσότερο τα ποσοστά επιτυχούς σύστασης.

2. Εφαρμογή του συστήματος με χρήστες 50 φυσικά πρόσωπα που είχαν στο παρελθόν επισκεφτεί καταλύματα στα Χανιά. Από αυτό το δείγμα επαληθεύτηκε ότι σε ένα ποσοστό 82%, οι χρήστες επιβεβαίωσαν την προτίμηση τους για τουλάχιστο ένα από τα ξενοδοχεία που τους προτάθηκαν κατά τη διαδικασία σύστασης.

Το σύστημα λόγω του τρόπου με τον οποίο καταλήγει σε σύσταση δεν απαιτεί προηγούμενα δεδομένα από το προφίλ του χρήστη ή τις επιλογές του. Συνεπώς ξεπερνά ένα συνήθη περιορισμό που τίθεται στα συστήματα συστάσεων που είναι η κρύα εκκίνηση (cold start). Η κρύα εκκίνηση αναφέρεται στην περίπτωση όπου δεν προϋπάρχουν δεδομένα σχετικά με κάποιο χρήστη στο σύστημα. Τότε συνήθως το σύστημα δυσκολεύεται ή δεν μπορεί να προβεί σε σύσταση, καθώς δεν έχει τρόπο να υπολογίσει την ομοιότητα (Hahsler, 2011).

Όσον αφορά την αποτελεσματικότητα των διαφορετικών μεθόδων συστάσεων που αναπτύχθηκαν, παρατηρείται ότι η μέθοδος συστάσεων με βάση το προτιμησιακό μοντέλο, το σταθμισμένο άθροισμα, το κατώφλι συναισθήματος και το μοντέλο φιλτραρίσματος (Recommendation Based on Preference Model, Weighted Sum, Sentiment Threshold and Filtering Model) καταλήγει σε συστάσεις που είναι πιο κοντά στις προτιμήσεις των χρηστών. Αυτό οφείλεται στο γεγονός ότι χρησιμοποιείται, πέραν του προτιμησιακού μοντέλου για την ταύτιση των προφίλ χρήστη-ξενοδοχείου, της ανάλυσης συναισθήματος και του σταθμισμένου αθροίσματος, το μοντέλο φιλτραρίσματος μέσω των ON/OFF κριτηρίων. Με αυτόν τον τρόπο μειώνονται οι πιθανές επιλογές ανάλογα με την επιλογή του χρήστη. Βέβαια, λόγω της χρήσης δύο κατ' ουσία διαφορετικών κατωφλίων, είναι μεγάλο το ενδεχόμενο να μην προκύψει κάποια σύσταση στο τέλος της διαδικασίας.

Στα λοιπά είδη σύστασης που παρουσιάστηκαν δεν εφαρμόστηκε το μοντέλο φιλτραρίσματος, συνεπώς προκύπτει τουλάχιστον ο ίδιος αριθμός συστάσεων.

Στο μοντέλο σύστασης με απλό άθροισμα (Recommendation Based on Preference Model with Single Sum) είναι δυνατόν να προκύπτουν αρκετά αποτελέσματα συστάσεων, πολλές φορές περισσότερα από τα αντίστοιχα των προηγούμενων μοντέλων, λόγω του τρόπου με τον οποίο υλοποιείται η σύσταση (Απλό άθροισμα και σταθμισμένο άθροισμα). Το συγκεκριμένο αποτέλεσμα για μικρό αριθμό προϊόντων (ξενοδοχείων) συμπίπτει με αυτό των προηγούμενων μεθόδων. Για ένα μεγαλύτερο δείγμα δεδομένων κρίνεται ανεπαρκές.

Το σύστημα συστάσεων παράγει σε πολύ μικρό χρόνο συστάσεις (<1 second). Κατά τους (Sarwar et al., 2001) είναι άκρως σημαντικό, στα συστήματα συστάσεων, να μετράται πόσο γρήγορα παράγουν τις συστάσεις τους, κυρίως για την βελτίωση της εμπειρίας του πελάτη.

Με βάση και την έρευνα των Pessemier et al., (2017), είναι πολύ σωστή αντιμετώπιση του προβλήματος των συστάσεων για ξενοδοχεία, με χρήση πολυκριτήριων μεθόδων, καθώς σε αντίθεση με άλλους τομείς έρευνας, υπάρχουν διαφορετικές παράμετροι που οι χρήστες θεωρούν πιο σημαντικές και οι οποίες θα πρέπει να ληφθούν υπόψη για μια ακριβή και αποδοτική τελική σύσταση. Η διαφορετικότητα του συστήματος που το κάνει μοναδικό είναι επίσης η ταυτόχρονη χρήση βαθμολογιών και κειμένου για να προσδιορίζει το επίπεδο ικανοποίησης του πελάτη καθώς και ότι τα δεδομένα εισόδου είναι ανώνυμα αλλά εξακριβωμένα ότι είναι αληθινά, συνεπώς εμπεριέχουν μεγάλο βαθμό αντικειμενικότητας.

Τέλος, ένα πολύ σημαντικό χαρακτηριστικό της παρούσας διατριβής είναι η ενσωμάτωση της ανάλυσης συναισθήματος σαν καθοριστικό κομμάτι της μεθοδολογίας συστάσεων. Η προστιθέμενη αξία της ενσωμάτωσης της συγκεκριμένης μεθόδου είναι μεγάλη, καθώς δίνει τη δυνατότητα εκμετάλλευσης πολλαπλών πηγών δεδομένων για προϊόντα ή υπηρεσίες που είναι δυνατό να βρεθούν στο διαδίκτυο σε μορφή κειμένου. Συνεπώς το αποτέλεσμα της ανάλυσης συναισθήματος μπορεί να είναι ακόμα πιο ακριβές και από πολύ μεγαλύτερο δείγμα δεδομένων σε σχέση με τις βαθμολογίες που αντίστοιχα αντλούνται, αλλά απαιτούν συγκεκριμένες προδιαγραφές όσον αφορά τη βαθμολόγηση των κριτηρίων.

## 7.1 Περιορισμοί

1. Απαίτηση ανατροφοδότησης από το χρήστη για εγκυρότητα λειτουργίας. Σε πολλές περιπτώσεις ο χρήστης είναι διστακτικός να κάνει ανατροφοδότηση με αποτέλεσμα να μην είναι δυνατή η τεκμηρίωση καλής λειτουργίας του συστήματος.
2. Η συνωνυμία ή πολυσημία λέξεων ή όρων είναι σύνηθες να καταλήγει σε ανακριβή αποτελέσματα. (Balabanović and Shoham, 1997b)
3. Ο περιορισμός αντικειμένου. Οι χρήστες μπορούν να δεχθούν συστάσεις μόνο για προϊόντα που υπάρχουν από το παρελθόν στη βάση προϊόντων, και έχουν ήδη λάβει πάνω από ένα ελάχιστο αριθμό κριτικών που είναι απαραίτητος προκειμένου να εφαρμοστεί το μοντέλο MUSA, για ανάλυση ικανοποίησης πελατών.
4. Περιορισμός στο φιλτράρισμα. Δεν υπάρχει δυνατότητα φιλτραρίσματος σε στοιχεία όπως η ποιότητα, το στυλ ή η οπτική γωνία σε ένα περιεχόμενο.
5. Ο περιορισμός στον αριθμό των κριτηρίων καθώς και στον αριθμό των πηγών για άντληση δεδομένων, καθώς είναι απαραίτητο προκειμένου να λειτουργήσει το σύστημα, η ολική αξιολόγηση που προσδίδεται από το χρήστη στο σύστημα να είναι ανεξάρτητη από τα υπόλοιπα κριτήρια και να μην υπολογίζεται σαν μέσος όρος των αξιολογήσεων των εκάστοτε κριτηρίων.
6. Η επεκτασιμότητα είναι ένα σύνηθες πρόβλημα όταν έχουμε να κάνουμε με μεγάλες βάσεις δεδομένων. Η μεγάλη ποσότητα δεδομένων σε μια βάση προκαλεί υπερφόρτωση του συστήματος. (Chen et al., 2008)

## 7.2 Προτάσεις για μελλοντική έρευνα

Η παρούσα διαδικτυακή διατριβή παρουσίασε ένα νέο μεθοδολογικό πλαίσιο ανοίγοντας πολλά ερωτήματα για περαιτέρω διερεύνηση στο τομέα των συστημάτων συστάσεων.

Η δυνατότητα σύστασης, με βάση βαθμολογίες και κριτικές χρηστών σε συνδυασμό με την ελάχιστη πληροφορία που καλείται να εισάγει ο χρήστης, αναμένεται να αποτελέσουν καταλυτικές παραμέτρους για την επέκταση χρήσης της μεθοδολογίας σε πλήθος διαφορετικών τομέων, πέραν του ξενοδοχειακού, που χρησιμοποιήσαμε σαν μελέτη περίπτωσης. Σημαντική εξέλιξη αποτελεί το γεγονός ότι τα τελευταία χρόνια οι κριτικές και οι βαθμολογίες χρηστών, αποτελούν αναπόσπαστο κομμάτι των περισσότερων διαδικτυακών πλατφορμών ανεξαρτήτως τομέα. Συνεπώς, η δημιουργία ενός πλήρους πλαισίου βέλτιστης επεξεργασίας και χρήσης τους για βελτίωση της εμπειρίας των χρηστών και της αποτελεσματικότητας των παρόχων υπηρεσιών, αναμένεται να τραβήξει το ενδιαφέρον όχι μόνο ερευνητών αλλά και επιχειρηματιών που δραστηριοποιούνται στο διαδίκτυο.

Ένα επόμενο βήμα, είναι η ενσωμάτωση του νέου συστήματος σε διαδικτυακή πλατφόρμα κρατήσεων/αγορών, προκειμένου να εφαρμοστεί η μεθοδολογία σε ένα ακόμα μεγαλύτερο δείγμα δεδομένων με πολλαπλούς χρήστες και με δυναμικά δεδομένα που τροποποιούνται σε πραγματικό χρόνο. Η ανατροφοδότηση από διαδικτυακούς χρήστες θα συντελέσει στην περαιτέρω βελτίωση του συστήματος και στην πιθανή χρήση περισσότερων κριτηρίων ή/και αποτελεσμάτων των μεθόδων για την επίτευξη καλύτερων συστάσεων.

Μια τροποποίηση στο παρόν σύστημα που θα είχε ενδιαφέροντα αποτελέσματα προς ανάλυση, θα ήταν η εμφάνιση των ράβδων κύλισης στη διαδικασία καθορισμού των τιμών  $Z_{min}$ ,  $Z_{max}$ , της μεθόδου WAP καθώς και η χρήση Ασαφούς λογικής για τον χαρακτηρισμό των διαστημάτων  $[Z_{min}, Z_{max}]$  για συνεχόμενα κριτήρια. Στόχος είναι να βρεθεί τρόπος να αναπαρασταθεί με όσο το δυνατόν πιο φιλική και κατανοητή προσέγγιση προς τον απλό χρήστη η μέθοδος υπολογισμού των  $Z_{min}, Z_{max}$ . Με συγκεκριμένη προσέγγιση



αναμένεται να παρουσιαστεί μεγαλύτερη διακύμανση στα προτιμησιακά προφίλ χρηστών. Μια περαιτέρω βελτίωση της μεθόδου εύρεσης προτιμησιακού προφίλ χρήστη θα ήταν η χρήση νευρωνικού δικτύου που θα εκπαιδεύεται με βάση τις τιμές των  $Z_{min}$ ,  $Z_{max}$  των πρότερων χρήσεων, και σε συσχέτιση τόσο με την κατάταξη των κριτηρίων όσο και την επιλογή κατ'αποκοπή κριτηρίων από το χρήστη να καταλήγει σε πρόβλεψη των τιμών  $Z_{min}$ ,  $Z_{max}$ .

Με βάση την έρευνα των Manouselis and Costopoulou, (2007c) θα μπορούσαμε επίσης σαν επόμενο βήμα να προχωρήσουμε σε clustering (ομαδοποίηση) των χρηστών υπολογίζοντας την ομοιότητα μεταξύ τους. Εφαρμόζοντας τη μέθοδο Similarity per priority, δηλαδή υπολογίζοντας τις ομοιότητες μεταξύ των χρηστών σε συνάρτηση με τα βάρη του κάθε χρήστη, για κάθε κριτήριο, θα μπορούσε να δημιουργηθεί μια γειτονιά χρηστών με παρόμοιες προτιμήσεις, με το χρήστη στόχο. Αυτό θα έχει σαν αποτέλεσμα, πέραν των συστάσεων που προκύπτουν από τη σύγκριση του προτιμησιακού προφίλ ενός χρήστη με το προφίλ του προϊόντος, να προβαίνουμε και σε επιπλέον σύσταση, με βάσει τις πιο συχνές επιλογές της ομάδας χρηστών με τα 'πλησιέστερα' προτιμησιακά προφίλ. Στη συνέχεια θα γίνεται προσπάθεια να προβλεφθεί η συνολική αξία ενός αντικειμένου για τον χρήστη με βάση τις συνολικές αξίες των χρηστών που γειτνιάζουν.

Μια διαφορετική προσέγγιση στο πρόβλημα της σύστασης μπορεί να είναι η σύγκριση των βαρύκεντρων των αποτελεσμάτων της WAP και MUSA με δείκτες ομοιότητας (similarity) όπως η ευκλείδεια απόσταση, η απόσταση Manhattan ή η απόσταση Minkowsky. Εν συνεχεία κατάταξη των αποτελεσμάτων συστάσεων με βάση τη τιμή της ομοιότητας.

Τέλος, ένας σημαντικός στόχος είναι η ενσωμάτωση της ανάλυσης συναισθήματος στη μεθοδολογία με τη μορφή 'εικονικής βαθμολογίας', ώστε να έχει πιο καθοριστικό ρόλο στη διαμόρφωση του προφίλ ικανοποίησης προϊόντος.

## 8. Βιβλιογραφία

### 8. Βιβλιογραφία

- Adomavicius, G., Kwon, Y., 2015. Multi-Criteria Recommender Systems, in: Ricci, F., Rokach, L., Shapira, B. (Eds.), *Recommender Systems Handbook*. Springer US, Boston, MA, pp. 847–880. [https://doi.org/10.1007/978-1-4899-7637-6\\_25](https://doi.org/10.1007/978-1-4899-7637-6_25)
- Adomavicius, G., Kwon, Y., 2007. New Recommendation Techniques for Multicriteria Rating Systems. *IEEE Intell. Syst.* 22, 48–55. <https://doi.org/10.1109/MIS.2007.58>
- Adomavicius, G., Manouselis, N., Kwon, Y., 2011a. Multi-Criteria Recommender Systems, in: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (Eds.), *Recommender Systems Handbook*. Springer US, Boston, MA, pp. 769–803. [https://doi.org/10.1007/978-0-387-85820-3\\_24](https://doi.org/10.1007/978-0-387-85820-3_24)
- Adomavicius, G., Manouselis, N., Kwon, Y., 2011b. Multi-Criteria Recommender Systems, in: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (Eds.), *Recommender Systems Handbook*. Springer US, Boston, MA, pp. 769–803.
- Adomavicius, G., Tuzhilin, A., 2005a. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowl. Data Eng. IEEE Trans. On* 17, 734–749.
- Adomavicius, G., Tuzhilin, A., 2005b. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowl. Data Eng. IEEE Trans. On* 17, 734–749.
- Agresti, A., Natarajan, R., 2001. Modeling Clustered Ordered Categorical Data: A Survey. *Int. Stat. Rev.* 69, 345–371. <https://doi.org/10.1111/j.1751-5823.2001.tb00463.x>
- Anderson, J.L., Bettencourt, S.U., 1993. A Conjoint Approach to Model Product Preferences: The New England Market for Fresh and Frozen Salmon. *Mar. Resour. Econ.* 8, 31–49. <https://doi.org/10.1086/mre.8.1.42629045>
- Antunes, F., Costa, J.P., 2012. Integrating Decision Support and Social Networks. *Adv. Hum.-Comput. Interact.* 2012, 1–10. <https://doi.org/10.1155/2012/574276>
- Arabatzis, G., Grigoroudis, E., 2010. Visitors' satisfaction, perceptions and gap analysis: The case of Dadia–Lefkimi–Soufli National Park. *For. Policy Econ.* 12, 163–172. <https://doi.org/10.1016/j.forpol.2009.09.008>
- Archak, N., Ghose, A., Ipeiritos, P.G., 2007. Show me the money!: deriving the pricing power of product features by mining consumer reviews. *ACM Press*, p. 56. <https://doi.org/10.1145/1281192.1281202>
- Atkinson, D., 2010. Research interviews with people with mental handicaps. *Ment. Handicap Res.* 1, 75–90. <https://doi.org/10.1111/j.1468-3148.1988.tb00006.x>
- Balabanović, M., Shoham, Y., 1997a. Fab: content-based, collaborative recommendation. *Commun. ACM* 40, 66–72. <https://doi.org/10.1145/245108.245124>
- Balabanović, M., Shoham, Y., 1997b. Fab: content-based, collaborative recommendation. *Commun. ACM* 40, 66–72.
- Bana e Costa, C.A., Vansnick, J.-C., 1994. MACBETH—An interactive path towards the construction of cardinal value functions. *Int. Trans. Oper. Res.* 1, 489–500.
- Benkler, Y., 2006. *The wealth of networks: how social production transforms markets and freedom*. Yale University Press, New Haven [Conn.].
- Bennett, J., Lanning, S., 2007. The netflix prize, in: *Proceedings of KDD Cup and Workshop*. New York, NY, USA, p. 35.

- Berka, T., Plößnig, M., 2004. Designing recommender systems for tourism.
- Bobadilla, J., Serradilla, F., Hernando, A., 2009. Collaborative filtering adapted to recommender systems of e-learning. *Knowl.-Based Syst.* 22, 261–265. <https://doi.org/10.1016/j.knosys.2009.01.008>
- Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. *J. Comput. Sci.* 2, 1–8. <https://doi.org/10.1016/j.jocs.2010.12.007>
- Breese, J.S., Heckerman, D., Kadie, C., 1998a. Empirical analysis of predictive algorithms for collaborative filtering, in: *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers Inc., pp. 43–52.
- Breese, J.S., Heckerman, D., Kadie, C., 1998b. Empirical analysis of predictive algorithms for collaborative filtering, in: *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers Inc., pp. 43–52.
- Bruce, R.F., Wiebe, J.M., 1999. Recognizing subjectivity: a case study in manual tagging. *Nat. Lang. Eng.* 5, 187–205. <https://doi.org/10.1017/S1351324999002181>
- Burke, R., 2002. Hybrid Recommender Systems: Survey and Experiments. *User Model. User-Adapt. Interact.* 12, 331–370.
- Burton, J., Khammash, M., 2010. Why do people read reviews posted on consumer-opinion portals? *J. Mark. Manag.* 26, 230–255. <https://doi.org/10.1080/02672570903566268>
- Carenini, G., Smith, J., Poole, D., 2003. Towards more conversational and collaborative recommender systems. *ACM Press*, p. 12. <https://doi.org/10.1145/604045.604052>
- Chapphannarungsri, K., Maneeroj, S., 2009a. Combining Multiple Criteria and Multidimension for Movie Recommender System. *Hong Kong* 6.
- Chapphannarungsri, K., Maneeroj, S., 2009b. Combining multiple criteria and multidimension for movie recommender system, in: *Proceedings of the International MultiConference of Engineers and Computer Scientists*.
- Chathoth, P.K., 2007. The impact of information technology on hotel operations, service management and transaction costs: A conceptual framework for full-service hotel firms. *Spec. Issue Self-Cater. Accommod.* 26, 395–408. <https://doi.org/10.1016/j.ijhm.2006.03.004>
- Chen, G., Chen, L., 2014. Recommendation Based on Contextual Opinions, in: Dimitrova, V., Kuflik, T., Chin, D., Ricci, F., Dolog, P., Houben, G.-J. (Eds.), *User Modeling, Adaptation, and Personalization*. Springer International Publishing, Cham, pp. 61–73. [https://doi.org/10.1007/978-3-319-08786-3\\_6](https://doi.org/10.1007/978-3-319-08786-3_6)
- Chen, L., Hsu, F., Chen, M., Hsu, Y., 2008. Developing recommender systems with the consideration of product profitability for sellers. *Inf. Sci.* 178, 1032–1048. <https://doi.org/10.1016/j.ins.2007.09.027>
- Chen, L., Wang, F., 2013. Preference-based clustering reviews for augmenting e-commerce recommendation. *Knowl.-Based Syst.* 50, 44–59. <https://doi.org/10.1016/j.knosys.2013.05.006>
- Cho, Y.H., Kim, J.K., Kim, S.H., 2002. A personalized recommender system based on web usage mining and decision tree induction. *Expert Syst. Appl.* 23, 329–342.
- Choi, T.Y., Chu, R., 2001. Determinants of hotel guests' satisfaction and repeat patronage in the Hong Kong hotel industry. *Int. J. Hosp. Manag.* 20, 277–297. [https://doi.org/10.1016/S0278-4319\(01\)00006-8](https://doi.org/10.1016/S0278-4319(01)00006-8)
- Cohen, W.W., Fan, W., 2000. Web-collaborative filtering: recommending music by crawling the Web. *Comput Netw* 33, 685–698.
- Colomo-Palacios, R., García-Peñalvo, F.J., Stantchev, V., Misra, S., 2017. Towards a social and context-aware mobile recommendation system for tourism. *Pervasive Mob. Comput.* 38, 505–515. <https://doi.org/10.1016/j.pmcj.2016.03.001>
- Courbon, J.C., 1987. Design of Intelligent Dialogue in Decision Support Systems. *IFAC Proc.* Vol. 20, 15–19. [https://doi.org/10.1016/S1474-6670\(17\)59024-2](https://doi.org/10.1016/S1474-6670(17)59024-2)

- Das, S.R., Chen, M.Y., 2001. Yahoo! for Amazon: Sentiment Parsing from Small Talk on the Web. SSRN Electron. J. <https://doi.org/10.2139/ssrn.276189>
- Dave, K., Lawrence, S., Pennock, D.M., 2003. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. ACM Press, p. 519. <https://doi.org/10.1145/775152.775226>
- Deshpande, M., Karypis, G., 2004. Item-based top-n recommendation algorithms. ACM Trans. Inf. Syst. TOIS 22, 143–177.
- Ding, X., Liu, B., Yu, P.S., 2008. A holistic lexicon-based approach to opinion mining. ACM Press, p. 231. <https://doi.org/10.1145/1341531.1341561>
- Dong, R., Schaal, M., O'Mahony, M.P., McCarthy, K., Smyth, B., 2013. Opinionated Product Recommendation, in: Delany, S.J., Ontañón, S. (Eds.), Case-Based Reasoning Research and Development. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 44–58. [https://doi.org/10.1007/978-3-642-39056-2\\_4](https://doi.org/10.1007/978-3-642-39056-2_4)
- Esparza, S.G., O'Mahony, M.P., Smyth, B., 2011. Effective Product Recommendation using the Real-Time Web, in: Bramer, M., Petridis, M., Hopgood, A. (Eds.), Research and Development in Intelligent Systems XXVII. Springer London, London, pp. 5–18. [https://doi.org/10.1007/978-0-85729-130-1\\_1](https://doi.org/10.1007/978-0-85729-130-1_1)
- European Commission, Directorate-General for the Information Society and Media (Eds.), 2010. Social networks overview: current trends and research challenges. EUR-OP, Luxembourg.
- Feinberg, G., 1977. Century of Matter. The Sciences 17, 27–29. <https://doi.org/10.1002/j.2326-1951.1977.tb01590.x>
- Figueira, José, Roy, B., 2002. Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. Eur. J. Oper. Res., EURO XVI: O.R. for Innovation and Quality of Life 139, 317–326.
- Figueira, Jose, Roy, B., 2002. Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. Eur. J. Oper. Res. 139, 317–326.
- Freeman, L.C., 2000. Visualizing social networks. J. Soc. Struct. 1, 4.
- Fuchs, M., Zanker, M., 2012. Multi-criteria Ratings for Recommender Systems: An Empirical Analysis in the Tourism Domain, in: E-Commerce and Web Technologies. Springer, pp. 100–111.
- Gamon, M., 2004. Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis, in: Proceedings of the 20th International Conference on Computational Linguistics. Association for Computational Linguistics, p. 841.
- Ganu, G., Kakodkar, Y., Marian, A., 2013. Improving the quality of predictions using textual information in online user reviews. Inf. Syst. 38, 1–15. <https://doi.org/10.1016/j.is.2012.03.001>
- Ghose, A., Ipeirotis, P.G., 2011. Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. IEEE Trans. Knowl. Data Eng. 23, 1498–1512. <https://doi.org/10.1109/TKDE.2010.188>
- Goldberg, D., Nichols, D., Oki, B.M., Terry, D., 1992. Using collaborative filtering to weave an information tapestry. Commun. ACM 35, 61–70.
- Green, P.E., Krieger, A.M., Wind, Y., 2001. Thirty Years of Conjoint Analysis: Reflections and Prospects. Interfaces 31, S56–S73. <https://doi.org/10.1287/inte.31.3s.56.9676>
- Green, P.E., Srinivasan, V., 1978. Conjoint Analysis in Consumer Research: Issues and Outlook. J. Consum. Res. 5, 103. <https://doi.org/10.1086/208721>
- Gretzel, U., Yoo, K.H., 2008. Use and Impact of Online Travel Reviews, in: O'Connor, P., Höpken, W., Gretzel, U. (Eds.), Information and Communication Technologies in Tourism 2008: Proceedings of the International Conference in Innsbruck, Austria, 2008. Springer Vienna, Vienna, pp. 35–46.

- Grigoroudis, E., Siskos, Y., 2010. Customer satisfaction evaluation: Methods for measuring and implementing service quality. Springer.
- Grigoroudis, E., Siskos, Y., 2002. Preference disaggregation for measuring and analysing customer satisfaction: The MUSA method. *Eur. J. Oper. Res.* 143, 148–170.
- Groh, G., Hauffa, J., 2011. Characterizing Social Relations Via NLP-Based Sentiment Analysis., in: ICWSM.
- Hahsler, M., 2011. recommenderlab: A Framework for Developing and Testing Recommendation Algorithms. Nov.
- Han, P., Xie, B., Yang, F., Shen, R., 2004. A scalable P2P recommender system based on distributed collaborative filtering. *Expert Syst. Appl.* 27, 203–210. <https://doi.org/10.1016/j.eswa.2004.01.003>
- Hatzivassiloglou, V., McKeown, K.R., 1997. Predicting the semantic orientation of adjectives. *Association for Computational Linguistics*, pp. 174–181. <https://doi.org/10.3115/976909.979640>
- Hinduja, A., National Institute of Technology, Raipur - 492001, Chhattisgarh, India, Pandey, M., National Institute of Technology, Raipur - 492001, Chhattisgarh, India, 2017. Multicriteria Recommender System for Life Insurance Plans based on Utility Theory. *Indian J. Sci. Technol.* 10, 1–8. <https://doi.org/10.17485/ijst/2017/v10i14/111376>
- Horst, R., Pardalos, P.M., Thoai, N.V., 1995. Introduction to global optimization, Nonconvex optimization and its applications. Kluwer Academic Publishers, Dordrecht ; Boston.
- Hu, M., Liu, B., 2004. Mining and summarizing customer reviews, in: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 168–177.
- Hung, L., 2005. A personalized recommendation system based on product taxonomy for one-to-one marketing online. *Expert Syst Appl* 29, 383–392.
- Jackson, M.O., n.d. Diffusion on Social Networks Matthew O. Jackson Caltech.
- Jakob, N., Weber, S.H., Müller, M.C., Gurevych, I., 2009. Beyond the stars: exploiting free-text user reviews to improve the accuracy of movie recommendations. ACM Press, p. 57. <https://doi.org/10.1145/1651461.1651473>
- Jay Kandampully, Dwi Suhartanto, 2000. Customer loyalty in the hotel industry: the role of customer satisfaction and image. *Int. J. Contemp. Hosp. Manag.* 12, 346–351. <https://doi.org/10.1108/09596110010342559>
- Jelassi, T., Kersten, G., Zionts, S., 1990. An Introduction to Group Decision and Negotiation Support, in: Bana e Costa, C.A. (Ed.), Readings in Multiple Criteria Decision Aid. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 537–568. [https://doi.org/10.1007/978-3-642-75935-2\\_23](https://doi.org/10.1007/978-3-642-75935-2_23)
- Jeong, B., Lee, J., Cho, H., 2009. User credit-based collaborative filtering. *Expert Syst. Appl.* 36, 7309–7312. <https://doi.org/10.1016/j.eswa.2008.09.034>
- Karacapilidis, N., Hatzieleftheriou, L., 2003a. Exploiting similarity measures in multi-criteria based recommendations, in: E-Commerce and Web Technologies. Springer, pp. 424–434.
- Karacapilidis, N., Hatzieleftheriou, L., 2003b. Exploiting similarity measures in multi-criteria based recommendations, in: E-Commerce and Web Technologies. Springer, pp. 424–434.
- Kasper, W., Vela, M., 2011. Sentiment Analysis for Hotel Reviews.
- Kim, K., Ahn, H., 2017. Recommender systems using cluster-indexing collaborative filtering and social data analytics. *Int. J. Prod. Res.* 55, 5037–5049. <https://doi.org/10.1080/00207543.2017.1287443>
- Kim, Y.S., Yum, B.-J., Song, J., Kim, S.M., 2005. Development of a recommender system based on navigational and behavioral patterns of customers in e-commerce sites. *Expert Syst. Appl.* 28, 381–393. <https://doi.org/10.1016/j.eswa.2004.10.017>

- King, B., Pizam, A., Milman, A., 1993. Social impacts of tourism. *Ann. Tour. Res.* 20, 650–665.  
[https://doi.org/10.1016/0160-7383\(93\)90089-L](https://doi.org/10.1016/0160-7383(93)90089-L)
- Konstan, J.A. (Ed.), 2004. Introduction to recommender systems: Algorithms and Evaluation. *ACM Trans. Inf. Syst.* 22, 1–4. <https://doi.org/10.1145/963770.963771>
- Konstan, J.A., Riedl, J., 2012. Recommender systems: from algorithms to user experience. *User Model. User-Adapt. Interact.* 22, 101–123. <https://doi.org/10.1007/s11257-011-9112-x>
- Kouloumpis, E., Wilson, T., Moore, J.D., 2011. Twitter sentiment analysis: The good the bad and the omg! *lcwsm* 11, 164.
- Kwon, Y., 2011. Computational Techniques For More Accurate and Diverse Recommendations. University of Minnesota.
- Lakiotaki, K., Delias, P., Sakkalis, V., Matsatsinis, N.F., 2009. User profiling based on multi-criteria analysis: the role of utility functions. *Oper. Res.* 9, 3–16.  
<https://doi.org/10.1007/s12351-008-0024-4>
- Lakiotaki, K., Matsatsinis, N., 2012. Analysing user behaviour in recommender systems. *Int. J. Electron. Bus.* 10, 1. <https://doi.org/10.1504/IJEB.2012.048740>
- Lakiotaki, K., Matsatsinis, N.F., Tsoukias, A., 2011. Multicriteria User Modeling in Recommender Systems. *IEEE Intell. Syst.* 26, 64–76.  
<https://doi.org/10.1109/MIS.2011.33>
- Lakiotaki, K., Tsafarakis, S., Matsatsinis, N., 2008. UTA-Rec: a recommender system based on multiple criteria analysis. *ACM Press*, p. 219.  
<https://doi.org/10.1145/1454008.1454043>
- Lam, S., Frankowski, D., Riedl, J., 2006. Do you trust your recommendations? An exploration of security and privacy issues in recommender systems. *Emerg. Trends Inf. Commun. Secur.* 14–29.
- Leal, F., Gonzalez-Velez, H., Malheiro, B., Burguillo, J.C., 2017. Profiling And Rating Prediction From Multi-Criteria Crowd-Sourced Hotel Ratings. *ECMS*, pp. 576–582.  
<https://doi.org/10.7148/2017-0576>
- Leung, C.W., n.d. Integrating Collaborative Filtering and Sentiment Analysis: A Rating Inference Approach 5.
- Levi, A., Mokryn, O., Diot, C., Taft, N., 2012. Finding a needle in a haystack of reviews: cold start context-based hotel recommender system, in: *Proceedings of the Sixth ACM Conference on Recommender Systems*. pp. 115–122.
- Li, L.-H., Hsu, R.-W., Lee, F.-M., 2012. Review of Recommender Systems and Their Applications. *資訊科技國際期刊* 6, 63–87.
- Linden, G., Smith, B., York, J., 2003. Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Comput.* 7, 76–80.  
<https://doi.org/10.1109/MIC.2003.1167344>
- Liu, B., 2012. *Sentiment Analysis and Opinion Mining*. Morgan Claypool Publ.
- Liu, L., Mehandjiev, N., Xu, D.-L., 2011. Multi-criteria service recommendation based on user criteria preferences. *ACM Press*, p. 77. <https://doi.org/10.1145/2043932.2043950>
- Malone, T.W., Grant, K.R., Turbak, F.A., Brobst, S.A., Cohen, M.D., 1987. Intelligent information-sharing systems. *Commun. ACM* 30, 390–402.  
<https://doi.org/10.1145/22899.22903>
- Manouselis, N., Costopoulou, C., 2007a. Analysis and Classification of Multi-Criteria Recommender Systems. *World Wide Web* 10, 415–441.  
<https://doi.org/10.1007/s11280-007-0019-8>
- Manouselis, N., Costopoulou, C., 2007b. Analysis and Classification of Multi-Criteria Recommender Systems. *World Wide Web* 10, 415–441.  
<https://doi.org/10.1007/s11280-007-0019-8>



- Manouselis, N., Costopoulou, C., 2007c. EXPERIMENTAL ANALYSIS OF DESIGN CHOICES IN MULTIATTRIBUTE UTILITY COLLABORATIVE FILTERING. *Int. J. Pattern Recognit. Artif. Intell.* 21, 311–331. <https://doi.org/10.1142/S021800140700548X>
- Manouselis, N., Matsatsinis, N.F., 2001. Introducing a Multi-agent, Multi-criteria Methodology for Modeling Electronic Consumer's Behavior: The Case of Internet Radio, in: Klusch, M., Zambonelli, F. (Eds.), *Cooperative Information Agents V*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 190–195. [https://doi.org/10.1007/3-540-44799-7\\_19](https://doi.org/10.1007/3-540-44799-7_19)
- Martineau, J., Finin, T., 2009. Delta TFIDF: An Improved Feature Space for Sentiment Analysis. *Icwsn* 9, 106.
- Matsatsinis, N., Moraitis, P., Psomatakis, V., Spanoudakis, N., 2003. AN AGENT-BASED SYSTEM FOR PRODUCTS PENETRATION STRATEGY SELECTION. *Appl. Artif. Intell.* 17, 901–925. <https://doi.org/10.1080/714858317>
- Matsatsinis, N.F., Siskos, Y., 1999. MARKEK: An intelligent decision support system for product development decisions. *Eur. J. Oper. Res.* 113, 336–354. [https://doi.org/10.1016/S0377-2217\(98\)00220-3](https://doi.org/10.1016/S0377-2217(98)00220-3)
- Maystre, L.Y., Pictet, J., Simos, J., 1994. Méthodes multicritères ELECTRE: description, conseils pratiques et cas d'application à la gestion environnementale. Presses polytechniques et universitaires romandes, Lausanne.
- McGlohon, M., Glance, N.S., Reiter, Z., 2010. Star Quality: Aggregating Reviews to Rank Products and Merchants., in: *ICWSM*.
- McLaughlin, M.R., Herlocker, J.L., 2004. A collaborative filtering algorithm and evaluation metric that accurately model the user experience. *ACM Press*, p. 329. <https://doi.org/10.1145/1008992.1009050>
- Meador, C.L., Guyote, M.J., Keen, P.G.W., 1984. Setting Priorities for DSS Development. *MIS Q.* 8, 117. <https://doi.org/10.2307/249348>
- Mirlas, L., 2010. Multisite commerce: proven principles for overcoming the business, organizational, and technical challenges. IBM Press/Pearson, Upper Saddle River, NJ.
- Moreno, A., Sebastiá, L., Vansteenwegen, P., n.d. Recommender Systems in Tourism 2.
- Moshfeghi, Y., Piwowarski, B., Jose, J.M., 2011. Handling data sparsity in collaborative filtering using emotion and semantic based features. *ACM Press*, p. 625. <https://doi.org/10.1145/2009916.2010001>
- Mudambi, S.M., Schuff, D., 2010. What makes a helpful review? A study of customer reviews on Amazon. com.
- Musat, C.-C., Liang, Y., Faltings, B., n.d. Recommendation Using Textual Opinions 7.
- Musto, C., de Gemmis, M., Semeraro, G., Lops, P., 2017. A Multi-criteria Recommender System Exploiting Aspect-based Sentiment Analysis of Users' Reviews. *ACM Press*, pp. 321–325. <https://doi.org/10.1145/3109859.3109905>
- Nasukawa, T., Yi, J., 2003. Sentiment analysis: capturing favorability using natural language processing. *ACM Press*, p. 70. <https://doi.org/10.1145/945645.945658>
- Ng, V., Dasgupta, S., Arifin, S.M., 2006. Examining the role of linguistic knowledge sources in the automatic identification and classification of reviews, in: *Proceedings of the COLING/ACL on Main Conference Poster Sessions*. Association for Computational Linguistics, pp. 611–618.
- Nilashi, M., Dalvi-Esfahani, M., Roudbaraki, M., Ramayah, T., Ibrahim, O., Oct2016. A Multi-Criteria Collaborative Filtering Recommender System Using Clustering and Regression Techniques. *J. Soft Comput. Decis. Support Syst.* 3, 24–30.
- O'Connor, B., Balasubramanyan, R., Routledge, B.R., Smith, N.A., 2010. From tweets to polls: Linking text sentiment to public opinion time series. *ICWSM* 11, 1–2.
- O'Mahony, M.P., Smyth, B., 2009. Learning to recommend helpful hotel reviews. *ACM Press*, p. 305. <https://doi.org/10.1145/1639714.1639774>

- Pang, B., Lee, L., 2004. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. *Association for Computational Linguistics*, pp. 271-es. <https://doi.org/10.3115/1218955.1218990>
- Parasuraman, A., Zeithaml, V.A., Berry, L.L., 1985. A Conceptual Model of Service Quality and Its Implications for Future Research. *J. Mark.* 49, 41. <https://doi.org/10.2307/1251430>
- Parrott, W.G., 2001. *Emotions in social psychology : essential readings*. Philadelphia (Pa.) : Psychology press.
- Pazzani, M.J., Billsus, D., 2007. Content-Based Recommendation Systems, in: Brusilovsky, P., Kobsa, A., Nejdl, W. (Eds.), *The Adaptive Web*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 325–341. [https://doi.org/10.1007/978-3-540-72079-9\\_10](https://doi.org/10.1007/978-3-540-72079-9_10)
- Pennock, D.M., Horvitz, E., Giles, C.L., others, 2000. Social choice theory and recommender systems: Analysis of the axiomatic foundations of collaborative filtering, in: *AAAI/IAAI*. pp. 729–734.
- Pero, Š., Horváth, T., 2013. Opinion-Driven Matrix Factorization for Rating Prediction, in: Carberry, S., Weibelzahl, S., Micarelli, A., Semeraro, G. (Eds.), *User Modeling, Adaptation, and Personalization*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–13. [https://doi.org/10.1007/978-3-642-38844-6\\_1](https://doi.org/10.1007/978-3-642-38844-6_1)
- Pessemier, T.D., Dhondt, J., Martens, L., 2017. Hybrid group recommendations for a travel service. *Multimed. Tools Appl.* 76, 2787–2811. <https://doi.org/10.1007/s11042-016-3265-x>
- Plantíé, M., Montmain, J., Dray, G., 2005a. Movies recommenders systems: automation of the information and evaluation phases in a multi-criteria decision-making process, in: *Database and Expert Systems Applications*. pp. 633–644.
- Plantíé, M., Montmain, J., Dray, G., 2005b. Movies recommenders systems: automation of the information and evaluation phases in a multi-criteria decision-making process, in: *Database and Expert Systems Applications*. pp. 633–644.
- Poirier, D., Fessant, F., Tellier, I., 2010. Reducing the Cold-Start Problem in Content Recommendation through Opinion Classification. *IEEE*, pp. 204–207. <https://doi.org/10.1109/WI-IAT.2010.87>
- Raad, E., Chbeir, R., Dipanda, A., 2011. Discovering relationship types between users using profiles and shared photos in a social network. *Multimed. Tools Appl.* <https://doi.org/10.1007/s11042-011-0853-7>
- Raghavan, S., Gunasekar, S., Ghosh, J., 2012. Review quality aware collaborative filtering. *ACM Press*, p. 123. <https://doi.org/10.1145/2365952.2365978>
- Ranjbar Kermany, N., Alizadeh, S.H., 2017. A hybrid multi-criteria recommender system using ontology and neuro-fuzzy techniques. *Electron. Commer. Res. Appl.* 21, 50–64. <https://doi.org/10.1016/j.elerap.2016.12.005>
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J., 1994. GroupLens: an open architecture for collaborative filtering of netnews. *ACM Press*, pp. 175–186. <https://doi.org/10.1145/192844.192905>
- Resnick, P., Varian, H.R., 1997. Recommender systems. *Commun. ACM* 40, 56–58. <https://doi.org/10.1145/245108.245121>
- Roy, B., 1990. The Outranking Approach and the Foundations of Electre Methods, in: Bana e Costa, C.A. (Ed.), *Readings in Multiple Criteria Decision Aid*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 155–183. [https://doi.org/10.1007/978-3-642-75935-2\\_8](https://doi.org/10.1007/978-3-642-75935-2_8)
- Santosa, P.I., Cahyono, A., Auliasari, K., 2012. A Multi-Criteria Recommender System For Tourism Destination 9.



- Sarwar, B., Karypis, G., Konstan, J., Reidl, J., 2001. Item-based collaborative filtering recommendation algorithms. ACM Press, pp. 285–295.  
<https://doi.org/10.1145/371920.372071>
- Schafer, J.B., Konstan, J., Riedl, J., 1999. Recommender systems in e-commerce. ACM Press, pp. 158–166. <https://doi.org/10.1145/336992.337035>
- Schärli, A., 1996. *Pratiquer Electre et Prométhée*. Lausanne. Press. Polytech. Univ. Romandes.
- Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M., 2005. CROC: a new evaluation criterion for recommender systems. *Electron. Commer. Res.* 5, 51–74.
- Shani, G., Gunawardana, A., 2011. Evaluating recommendation systems, in: *Recommender Systems Handbook*. Springer, pp. 257–297.
- Shardanand, U., Maes, P., 1995. Social information filtering: algorithms for automating “word of mouth.” ACM Press, pp. 210–217. <https://doi.org/10.1145/223904.223931>
- Shearer, C., 2000a. The CRISP-DM model: the new blueprint for data mining. *J Data Warehous.* 5, 13–22.
- Simonson, I., Rosen, E., 2014. *Absolute value: what really influences customers in the age of (nearly) perfect information*, First edition. ed. HarperBusiness, New York, NY.
- Simos, J., 1990a. *L'évaluation environnementale: Un processus cognitif négocié*. Thèse de doctorat, DGF EPFL, Lausanne.
- Simos, J., 1990b. *Evaluer l'impact sur l'environnement: Une approche originale par l'analyse multicritère et la négociation*, Presses Polytechniques et Universitaires Romandes, Lausanne. ed.
- Simos, J., 1990a. *L'évaluation environnementale: Un processus cognitif negocié*. These de doctorat. DGF-EPFL, Lausanne.
- Siskos, Y., 1985. Analyses de régression et programmation linéaire. *Révue Stat. Appliquée* XXXII 41–55.
- Siskos, Y., Matsatsinis, N., Baourakis, G., 2001. Multicriteria analysis in agricultural marketing: The case of French olive oil market. *Eur. J. Oper. Res.* 130, 315–331.  
[https://doi.org/10.1016/S0377-2217\(00\)00043-6](https://doi.org/10.1016/S0377-2217(00)00043-6)
- Skogland, I., others, 2004. Understanding switchers and stayers in the lodging industry.
- Sprague, R.H., 1980. A Framework for the Development of Decision Support Systems. *MIS Q.* 4, 1. <https://doi.org/10.2307/248957>
- Spyridakos, A., Tsotsolas, N., 2015. Estimating criteria weights exploiting priorities of the criteria and techniques of robustness analysis. Presented at the 4th Int. Symp. 26th Natl. Conf. Oper. Res, Chania, Greece, pp. 225–229.
- Spyridakos, A., Tsotsolas, N., Yannacopoulos, D., Siskos, E., 2015. Estimating criteria weights exploiting priorities of the criteria and techniques of robustness analysis. *Proceeding 4th Int. Symp. 26th Natl. Conf. Oper. Res.* June 4-6, Chania-Greece, 225–229.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M., 2011. Lexicon-based methods for sentiment analysis. *Comput. Linguist.* 37, 267–307.
- Tang, T.Y., McCalla, G., 2009. The pedagogical value of papers: a collaborative-filtering based paper recommender. *J. Digit. Inf.* 10.
- Tobias H Engler, Patrick Winter, Michael Schulz, 2015. Understanding online product ratings: A customer satisfaction model. *Elsevier* 27, 113–120.
- Tokuhisa, R., Inui, K., Matsumoto, Y., 2008. Emotion classification using massive examples extracted from the web, in: *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*. Association for Computational Linguistics, pp. 881–888.
- Traub, M., Kowald, D., Lacic, E., Schoen, P., Supp, G., Lex, E., 2015. Smart booking without looking: providing hotel recommendations in the TripRebel portal. ACM Press, pp. 1–4. <https://doi.org/10.1145/2809563.2809616>

- Tribe, J., Snaith, T., 1998. From SERVQUAL to HOLSAT: holiday satisfaction in Varadero, Cuba. *Tour. Manag.* 19, 25–34.
- Truong, T.-H., Foster, D., 2006. Using HOLSAT to evaluate tourist satisfaction at destinations: The case of Australian holidaymakers in Vietnam. *Tour. Manag.* 27, 842–855. <https://doi.org/10.1016/j.tourman.2005.05.008>
- Tsafarakis, S., Lakiotaki, K., Matsatsinis, N., 2010. Applications of MCDA in Marketing and e-Commerce, in: Zopounidis, C., Pardalos, P.M. (Eds.), *Handbook of Multicriteria Analysis*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 425–448. [https://doi.org/10.1007/978-3-540-92828-7\\_15](https://doi.org/10.1007/978-3-540-92828-7_15)
- Tsotsolas, N., Spyridakos, A., Siskos, E., Salmon, I., 2016a. Criteria weights assessment through prioritizations (WAP) using linear programming techniques and visualizations. *Oper. Res.* <https://doi.org/10.1007/s12351-016-0280-7>
- Tsotsolas, N., Spyridakos, A., Siskos, E., Salmon, I., 2016b. Criteria weights assessment through prioritizations (WAP) using linear programming techniques and visualizations. *Oper. Res.* 1–16. <https://doi.org/10.1007/s12351-016-0280-7>
- Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.M., 2010. Predicting elections with twitter: What 140 characters reveal about political sentiment. *Icwsn* 10, 178–185.
- Verma, R., Stock, D., McCarthy, L., 2012. Customer Preferences for Online, Social Media, and Mobile Innovations in the Hospitality Industry. *Cornell Hosp. Q.* 53, 183–186. <https://doi.org/10.1177/1938965512445161>
- Wang, F., Pan, W., Chen, L., 2013. Recommendation for New Users with Partial Preferences by Integrating Product Reviews with Static Specifications, in: Carberry, S., Weibelzahl, S., Micarelli, A., Semeraro, G. (Eds.), *User Modeling, Adaptation, and Personalization*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 281–288. [https://doi.org/10.1007/978-3-642-38844-6\\_24](https://doi.org/10.1007/978-3-642-38844-6_24)
- Wang, P., 2012. A Personalized Collaborative Recommendation Approach Based on Clustering of Customers. *Phys. Procedia* 24, 812–816. <https://doi.org/10.1016/j.phpro.2012.02.121>
- Wiebe, J., Wilson, T., Cardie, C., 2005. Annotating Expressions of Opinions and Emotions in Language. *Lang. Resour. Eval.* 39, 165–210. <https://doi.org/10.1007/s10579-005-7880-9>
- Yang, J.-M., Li, K.F., 2009. Recommendation based on rational inferences in collaborative filtering. *Knowl.-Based Syst.* 22, 105–114. <https://doi.org/10.1016/j.knosys.2008.07.004>
- Ye, Q., Law, R., Gu, B., Chen, W., 2011. The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Comput. Hum. Behav.* 27, 634–639. <https://doi.org/10.1016/j.chb.2010.04.014>
- Yu, H., Hatzivassiloglou, V., 2003. Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences. *Association for Computational Linguistics*, pp. 129–136. <https://doi.org/10.3115/1119355.1119372>
- Zhang, W., Ding, G., Chen, L., Li, C., Zhang, C., 2013. Generating virtual ratings from chinese reviews to augment online recommendations. *ACM Trans. Intell. Syst. Technol.* 4, 1–17. <https://doi.org/10.1145/2414425.2414434>
- Zhang, Z., Varadarajan, B., 2006. Utility scoring of product reviews. *ACM Press*, p. 51. <https://doi.org/10.1145/1183614.1183626>
- Ziegler, C.-N., McNeel, S.M., Konstan, J.A., Lausen, G., 2005. Improving recommendation lists through topic diversification, in: *Proceedings of the 14th International Conference on World Wide Web*. ACM, pp. 22–32.
- Ματσατσίνη, Ν., 2010. Συστήματα υποστήριξης αποφάσεων. Νέες Τεχνολογίες, Αθήνα.

## 9. Παράρτημα

### 9. Παράρτημα

Αποτελέσματα ερωτηματολογίου για τον προσδιορισμό των τιμών  $Z_{min}$ ,  $Z_{max}$  της μεθόδου WAP

Πίνακας 9-1 Αποτελέσματα ερωτηματολογίου για τις τιμές  $Z_{min}$ ,  $Z_{max}$

	ZminA vrg	ZmaxAvr g	ZminAv rg	ZmaxAv rg	ZminAv rg	ZmaxAv rg	ZminAv rg	ZmaxAvr g	ZminAv rg	ZmaxAv rg
	1.0243	1.3512	1.3873	1.7978	1.8289	2.3051	2.3161	2.8055	2.8138	3.4009
	Very Weak Preference		Weak Preference		Preference		Strong Preference		Very strong Preference	
	Zmin	Zman	Zmin	Zmax	Zmin	Zman	Zmin	Zmax	Zmin	Zmax
1	1.0202	1.1505	1.439	2.0769	1.7027	2.4483	2.5088	2.8462	2.9216	3.4444
2	1.0202	1.2472	1.5974	1.8571	2.125	2.125	2.2258	2.8462	3	3.3478
3	1.0202	1.4691	1.6316	1.5974	1.7027	2.5714	2.1746	3	3.3478	3.4444
4	1.0202	1.2472	1.4691	1.5	1.7397	2.2787	2.3333	3.2553	2.9216	3.7619
5	1.0202	1.3529	1.3256	1.6316	1.9412	2.125	2.3898	2.6364	2.7037	2.9216
6	1.0202	1.4096	1.3256	1.8169	2.125	2.5088	2.3333	2.8462	2.5714	3.3478
7	1.0202	1.5316	1.439	1.8169	1.9412	2.1746	2.8462	2.5714	2.6364	3.3478
8	1.0202	1.4096	1.439	1.7027	1.6316	2.5088	2.5088	2.3333	2.7037	2.8462
9	1.0202	1.4096	1.3529	1.8571	1.7027	2.1746	2.3898	2.6364	3.2553	2.8462
10	1.0408	1.1978	1.2472	1.5974	1.8169	1.9851	2.1746	2.8462	2.6364	3.1667
11	1.0408	1.2989	1.3529	1.6316	1.7027	2.125	2.2258	2.9216	2.7037	3.3478
12	1.0202	1.5	1.4096	2.0769	1.8169	2.5714	2.1746	3.3478	2.7037	3.7619
13	1.0408	1.3529	1.5316	1.7397	1.8169	2.4483	2.3898	2.3898	2.5714	3.4444
14	1.0202	1.4096	1.4096	1.8571	1.8169	2.2787	2.3333	2.8462	2.8462	2.9216
15	1.0202	1.3256	1.3529	1.8169	1.7397	1.9851	2.2787	3.3478	2.9216	2.8462
16	1.0202	1.1505	1.3256	1.7027	1.9412	2.1746	2.3333	2.7736	2.7736	3.2553
17	1.0202	1.3256	1.3529	1.6316	1.9412	2.2787	2.1746	3.3478	2.5714	3.4444
18	1.0202	1.3256	1.5974	1.8169	1.8571	1.9851	2.2258	2.3333	2.8462	3
19	1.0202	1.2472	1.5974	1.8986	1.8571	2.4483	2.5088	2.8462	2.9216	3.4444
20	1.0202	1.3256	1.2472	1.7027	1.9851	1.9851	2.2787	2.6364	3.7619	3.7619
21	1.0202	1.439	1.3529	1.6316	1.7027	2.125	2.3333	2.3333	2.7736	3.3478
22	1.0408	1.1277	1.2472	1.8986	1.8571	2.8462	2.1746	2.7736	3.1667	3.3478
23	1.0408	1.1978	1.439	1.8986	1.8571	1.9851	2.2787	2.3333	2.7736	3.878
24	1.0202	1.381	1.3529	1.6316	1.9412	2.1746	2.4483	3.3478	2.5714	3.878
25	1.0202	1.4691	1.2472	2.0769	1.7027	2.1746	2.1746	3.3478	2.7736	3.3478
26	1.0408	1.1739	1.5974	2.0769	1.7027	2.5088	2.4483	2.3333	2.9216	3.2553
27	1.0202	1.2989	1.1739	2.1746	1.8571	2.8462	2.2258	3.3478	2.7736	3.4444
28	1.0202	1.381	1.3256	1.7397	2.1746	2.5714	2.3333	2.7736	2.5714	3
29	1.0202	1.2222	1.5974	2.0303	1.5316	1.9851	2.2258	2.6364	2.7736	3.4444
30	1.0202	1.3256	1.3256	1.7027	1.9412	2.1746	2.4483	2.8462	2.7736	3.4444
31	1.0408	1.4096	1.3529	1.7397	2.125	2.125	2.2787	2.3333	2.6364	3.4444
32	1.0202	1.1739	1.7397	1.9412	2.125	2.1746	2.1746	2.8462	2.8462	3.3478
33	1.0202	1.1978	1.3529	1.7027	1.8169	2.2787	2.3333	2.6364	2.8462	3.4444
34	1.0202	1.2989	1.6316	2.125	2.125	2.4483	2.8462	2.7736	2.6364	3.2553
35	1.0408	1.2989	1.3529	1.6316	1.7397	2.2787	2.2258	2.3333	2.9216	3.4444

36	1.0202	1.2989	1.3256	2.125	1.7027	2.4483	2.5088	2.9216	2.9216	3.4444
37	1.0202	1.4691	1.3529	2.125	1.7397	2.3898	2.3333	3.3478	2.7736	3.4444
38	1.0202	1.4691	1.3256	1.8571	1.7027	2.3898	2.3333	2.7736	2.9216	3.3478
39	1.0408	1.2989	1.3256	1.7397	1.8169	2.1746	2.2787	3	2.7037	3.878
40	1.0202	1.5	1.2472	1.6316	1.7397	2.1746	2.2787	2.3333	2.7037	3.878
41	1.0202	1.5316	1.3529	1.6316	1.6316	2.1746	2.1746	2.7736	3	3.878
42	1.0202	1.4691	1.3529	2.0769	1.7027	2.5714	2.2787	2.3333	2.9216	3.5455
43	1.0202	1.3256	1.2472	1.6316	1.8169	2.4483	2.2258	2.7736	2.5714	3.5455
44	1.0202	1.4096	1.5	1.6316	1.7397	2.5714	2.2258	3.3478	2.7037	3.5455
45	1.0408	1.5	1.1739	1.7027	1.8169	2.4483	2.2787	2.8462	2.6364	3.3478
46	1.0202	1.4691	1.3529	1.7397	1.8169	1.9851	2.2787	2.9216	2.9216	3.2553
47	1.0202	1.2989	1.3529	1.6316	1.7397	2.3898	2.1746	3.3478	2.7037	3.2553
48	1.0202	1.5316	1.3529	1.7027	1.8169	2.3333	2.2787	2.7736	2.5714	3.7619
49	1.0202	1.4096	1.2472	1.7027	1.7027	2.8462	2.2258	2.3333	2.6364	3.5455
50	1.0202	1.5	1.3256	1.6316	1.8571	2.0303	2.1746	2.8462	2.9216	3.3478

### Αποτελέσματα του πειράματος με χρήση του συστήματος από ταξιδιωτικό γραφείο

Βαθμολογίες των 73 χρηστών για τα 5 κριτήρια καθώς και συνολική βαθμολογία Πίνακας 9-2. Τα ονόματα των ξενοδοχείων έχουν αντικατασταθεί με αρίθμηση, καθώς έχουν χρησιμοποιηθεί μόνο για την υλοποίηση του πειράματος και όχι για εμπορικούς λόγους.

**Πίνακας 9-2: Βαθμολογίες των 73 χρηστών του πειράματος για 5 κριτήρια και συνολική βαθμολογία**

User s	User Overall Rating	Cleanlines s	Servic e	Comfor t	Conditio n	Neighbourhoo d	Hotels	
U1	4	5	3	3	4	4	Hotel 1	3Stars
U2	5	5	4	5	5	5	Hotel 2	3Stars
U3	5	5	5	5	5	4	Hotel 3	3Stars
U4	5	5	5	5	4	5	Hotel 4	3Stars
U5	4	4	4	4	4	5	Hotel 5	3Stars
U6	5	5	5	5	5	5	Hotel 6	5Stars
U7	5	5	5	5	5	5	Hotel 7	4Stars
U8	5	5	5	5	4	5	Hotel 8	4Stars
U9	5	5	4	5	5	5	Hotel 9	5Stars
U10	4	5	5	4	4	5	Hotel	4Star

# Παράρτημα

							10	s
U11	5	5	5	5	5	4	Hotel 11	4Stars
U12	4	4	5	5	4	4	Hotel 12	4Stars
U13	5	5	5	5	5	5	Hotel 13	4Stars
U14	4	5	5	4	4	5	Hotel 14	5Stars
U15	5	5	5	5	5	5	Hotel 15	5Stars
U16	4	5	5	4	4	4	Hotel 16	5Stars
U17	5	4	5	4	4	5	Hotel 17	4Stars
U18	4	4	5	5	5	4	Hotel 18	3Stars
U19	5	5	5	5	5	5	Hotel 19	4Stars
U20	4	5	5	4	4	4	Hotel 20	3Stars
U21	4	4	5	4	3	4	Hotel 21	3Stars
U22	5	5	5	5	4	5	Hotel 22	3Stars
U23	5	5	5	5	5	5	Hotel 23	4Stars
U24	5	5	5	4	5	5	Hotel 24	2Stars
U25	4	5	5	3	4	3	Hotel 25	3Stars
U26	5	4	5	5	5	5	Hotel 26	3Stars
U27	5	5	5	5	5	4	Hotel 27	5Stars
U28	5	5	5	5	5	5	Hotel 28	5Stars
U29	4	4	4	2	4	4	Hotel 29	4Stars
U30	5	5	5	5	4	4	Hotel 30	5Stars
U31	4	4	4	4	5	3	Hotel 31	5Stars
U32	5	5	5	5	5	5	Hotel 32	4Stars
U33	5	5	5	5	5	5	Hotel 33	4Stars
U34	5	5	5	4	5	5	Hotel 34	4Stars
U35	4	5	4	5	5	5	Hotel 35	3Stars
U36	4	4	5	4	4	5	Hotel 36	3Stars
U37	4	3	4	4	4	4	Hotel 37	3Stars

## Παράρτημα

U38	4	5	5	4	4	3	Hotel 38	3Stars
U39	5	5	5	5	5	4	Hotel 39	4Stars
U40	5	5	4	5	4	5	Hotel 40	4Stars
U41	4	5	5	4	4	2	Hotel 41	5Stars
U42	5	4	5	5	4	5	Hotel 42	5Stars
U43	4	5	5	5	4	3	Hotel 43	4Stars
U44	4	4	4	4	4	4	Hotel 44	4Stars
U45	5	5	5	5	5	5	Hotel 45	4Stars
U46	5	5	5	4	5	4	Hotel 46	4Stars
U47	5	5	5	5	4	5	Hotel 47	5Stars
U48	4	5	4	3	4	5	Hotel 48	4Stars
U49	5	5	5	4	4	4	Hotel 49	3Stars
U50	5	5	5	5	5	4	Hotel 50	3Stars
U51	5	5	5	5	5	3	Hotel 51	3Stars
U52	5	5	5	5	5	5	Hotel 52	2Stars
U53	5	5	5	5	5	5	Hotel 53	4Stars
U54	5	5	5	5	5	4	Hotel 54	5Stars
U55	5	5	5	5	5	5	Hotel 55	4Stars
U56	5	5	5	4	4	5	Hotel 56	3Stars
U57	5	5	5	4	4	5	Hotel 57	3Stars
U58	4	5	5	4	4	4	Hotel 58	4Stars
U59	4	4	4	3	3	3	Hotel 59	4Stars
U60	4	4	4	4	3	4	Hotel 60	4Stars
U61	5	5	5	5	5	5	Hotel 61	5Stars
U62	5	5	5	5	5	5	Hotel 62	5Stars
U63	5	5	4	5	5	5	Hotel 63	5Stars
U64	4	5	5	4	5	5	Hotel 64	5Stars
U65	4	4	5	4	4	5	Hotel	4Stars

							65	s
U66	5	5	5	5	5	5	Hotel 66	4Stars
U67	5	5	5	4	4	4	Hotel 67	5Stars
U68	5	5	5	4	4	5	Hotel 68	5Stars
U69	5	5	5	5	5	3	Hotel 69	5Stars
U70	5	4	5	4	4	5	Hotel 70	4Stars
U71	5	5	5	5	5	5	Hotel 71	4Stars
U72	4	5	4	5	5	5	Hotel 72	5Stars
U73	4	4	4	3	4	3	Hotel 73	4Stars

Πίνακας με τους μέσους όρους των συνολικών βαθμολογιών και βαθμολογιών των κριτηρίων για τα ξενοδοχεία που επελέγησαν από τους χρήστες του πειράματος.

Οι βαθμολογίες προέρχονται από την αρχική άντληση δεδομένων.

**Πίνακας 9-3 Μέσοι όροι Βαθμολογιών κριτηρίων και συνολικών βαθμολογιών για ξενοδοχεία**

Hotels		Hotel Rating	Cleanliness	Service	Comfort	Condition	Neighbourhood
Hotel 1	3Stars	3.58	3.5	3.416	3.416	3.25	4
Hotel 2	3Stars	4.5	4.411	4.1764	4	4.17	4.764
Hotel 3	3Stars	3.57	3.714	3.857	3.428	3.571	3.428
Hotel 4	3Stars	4.58	5	4.833	4.17	4.666	3.75
Hotel 5	3Stars	4.01	4.223	4.247	3.776	3.788	4.141
Hotel 6	5Stars	4.23	4.477	4.136	4.022	4.25	4.227
Hotel 7	4Stars	4.92	4.918	4.945	4.864	4.945	4.946
Hotel 8	4Stars	4.45	4.581	4.657	4.351	4.324	4.513
Hotel 9	5Stars	4.64	4.794	4.64	4.551	4.653	4.23
Hotel 10	4Stars	4.45	4.581	4.657	4.351	4.324	4.513
Hotel 11	4Stars	4.69	4.644	4.711	4.533	4.6	4.311
Hotel 12	4Stars	3.95	4.21	4.263	3.368	4.105	3.684
Hotel 13	4Stars	4.45	4.581	4.657	4.351	4.324	4.513
Hotel 14	5Stars	4.33	4.666	3.857	4.286	4.571	3.381
Hotel 15	5Stars	4.57	4.857	4.57	4.428	4.428	4.428
Hotel 16	5Stars	4.23	4.477	4.136	4.022	4.25	4.227
Hotel 17	4Stars	4.32	4.818	4.636	4.091	4.136	4.818
Hotel 18	3Stars	4.36	4.48	4.64	4.44	4.56	3.72
Hotel 19	4Stars	4.92	4.918	4.945	4.864	4.945	4.946
Hotel 20	3Stars	4.25	4.606	4.553	4.095	4.213	4

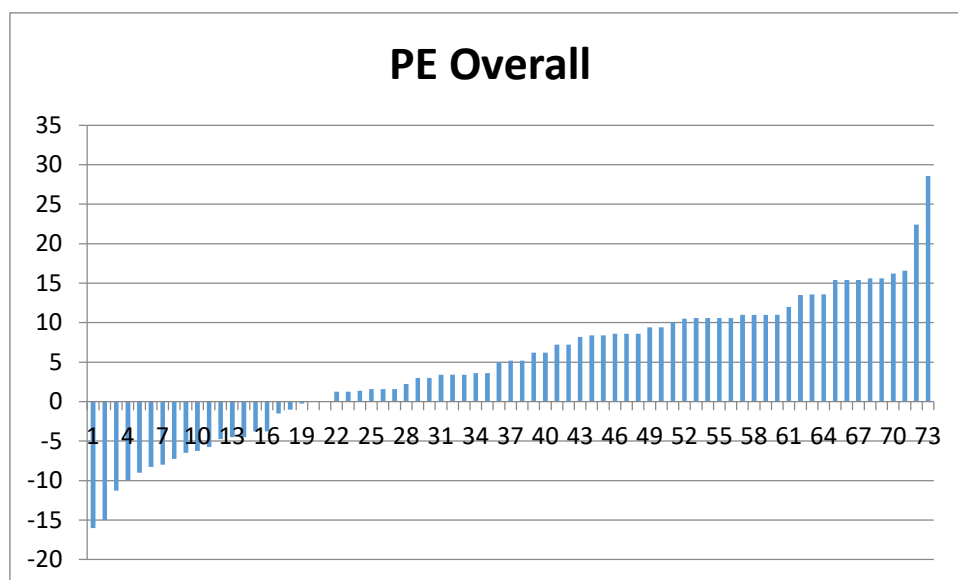
## Παράρτημα

Hotel 21	3Stars	4	4.14	4.5	3.34	3.78	4.24
Hotel 22	3Stars	4.74	4.519	4.563	4.039	4.236	4.804
Hotel 23	4Stars	4.82	4.765	4.882	4.823	4.823	4.235
Hotel 24	2Stars	4.85	4.917	4.284	4.426	4.815	4.509
Hotel 25	3Stars	4.26	4.404	4.399	4.051	4.165	3.413
Hotel 26	3Stars	4.74	4.519	4.563	4.039	4.236	4.804
Hotel 27	5Stars	4.47	4.77	4.6	4.371	5	4.071
Hotel 28	5Stars	4.57	4.857	4.57	4.428	4.428	4.428
Hotel 29	4Stars	3.88	4.1	4.074	3.298	3.627	4
Hotel 30	5Stars	4.64	4.794	4.64	4.551	4.653	4.23
Hotel 31	5Stars	4.32	4.488	4.488	3.953	4.116	4.209
Hotel 32	4Stars	4.75	4.75	4.5	4.625	4.75	4.25
Hotel 33	4Stars	4.83	4.847	4.796	4.796	4.83	4.508
Hotel 34	4Stars	4.32	4.818	4.636	4.091	4.136	4.818
Hotel 35	3Stars	4.06	4.371	4.4	4.057	4.142	4.086
Hotel 36	3Stars	4.04	4.4	4.36	3.52	3.88	4.56
Hotel 37	3Stars	4.15	4.25	4.35	3.79	4.166	3.958
Hotel 38	3Stars	4.29	4.56	4.707	4.026	4.268	3.951
Hotel 39	4Stars	4.82	4.765	4.882	4.823	4.823	4.235
Hotel 40	4Stars	3.88	4.1	4.074	3.298	3.627	4
Hotel 41	5Stars	4.18	4.705	4.294	4.06	4.05	3.647
Hotel 42	5Stars	4.23	4.477	4.136	4.022	4.25	4.227
Hotel 43	4Stars	4.18	4.705	4.294	4.06	4.05	3.647
Hotel 44	4Stars	3.91	4.318	4.227	4.1818	4.23	3.636
Hotel 45	4Stars	4.92	4.918	4.945	4.864	4.945	4.946
Hotel 46	4Stars	4.69	4.644	4.711	4.533	4.6	4.311
Hotel 47	5Stars	4.93	5	5	4.928	5	4.7857
Hotel 48	4Stars	4.4	4.6	4.266	4.266	4.466	4.46
Hotel 49	3Stars	4.47	4.756	4.621	4.378	4.27	4.324
Hotel 50	3Stars	4.53	4.812	4.656	4.312	4.406	4.218
Hotel 51	3Stars	4.58	5	4.833	4.17	4.666	3.75
Hotel 52	2Stars	4.59	4.656	4.5625	4.468	4.5	4.218
Hotel 53	4Stars	4.45	4.581	4.657	4.351	4.324	4.513
Hotel 54	5Stars	4.22	4.351	4.675	3.918	4.081	3.972
Hotel 55	4Stars	4.83	4.847	4.796	4.796	4.83	4.508
Hotel 56	3Stars	4.47	4.756	4.621	4.378	4.27	4.324
Hotel 57	3Stars	4.53	4.631	4.157	4.473	4.526	4.157
Hotel 58	4Stars	4	4	4.25	4	3.875	4.375
Hotel 59	4Stars	3.46	3.8	4.057	3.371	3.257	3.285
Hotel 60	4Stars	4.15	4.259	4.407	3.666	4.074	4.037
Hotel 61	5Stars	4.19	4.437	4.375	4.062	4.05	4.25
Hotel 62	5Stars	4.57	4.857	4.57	4.428	4.428	4.428
Hotel 63	5Stars	4.22	4.351	4.675	3.918	4.081	3.972
Hotel 64	5Stars	4.64	4.794	4.64	4.551	4.653	4.23
Hotel 65	4Stars	4.6	4.746	4.657	4.174	4.412	4.571
Hotel 66	4Stars	4.83	4.847	4.796	4.796	4.83	4.508

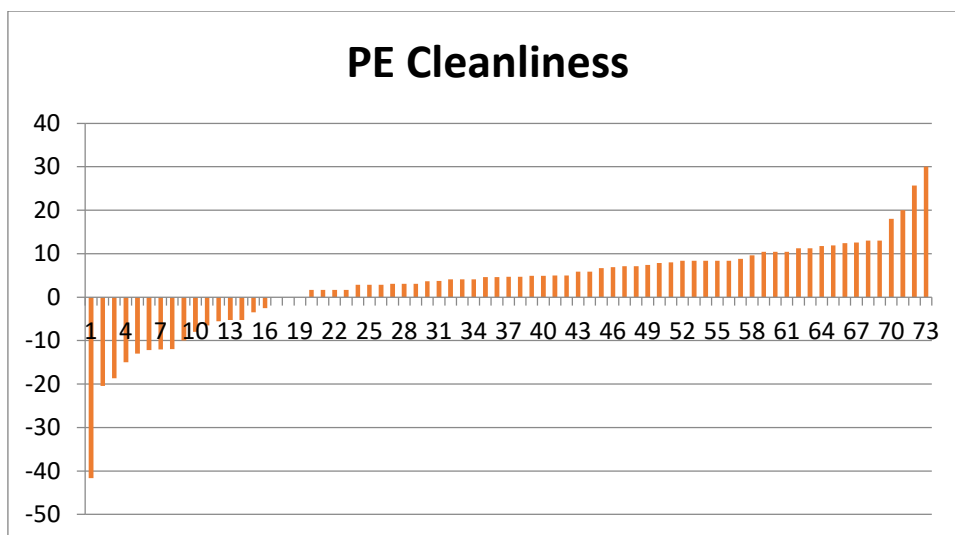


Hotel 67	5Stars	4.47	4.77	4.6	4.371	5	4.071
Hotel 68	5Stars	4.23	4.477	4.136	4.022	4.25	4.227
Hotel 69	5Stars	4.17	4.38	4.024	4.333	4.309	3.547
Hotel 70	4Stars	4.4	4.6	4.266	4.266	4.466	4.46
Hotel 71	4Stars	4.45	4.581	4.657	4.351	4.324	4.513
Hotel 72	5Stars	4.19	4.437	4.375	4.062	4.05	4.25
Hotel 73	4Stars	3.95	4.21	4.263	3.368	4.105	3.684

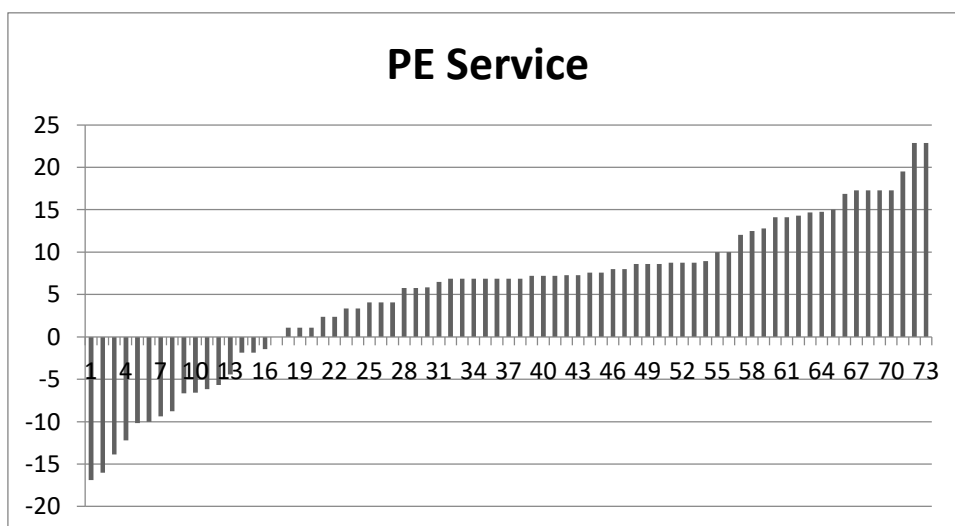
Γραφικές για τις μετρικές σφάλματος με βάση τα δεδομένα των βαθμολογιών από χρήστες του πειράματος και βαθμολογιών αντλημένων από τη πλατφόρμα κρατήσεων για τα αντίστοιχα ξενοδοχεία.



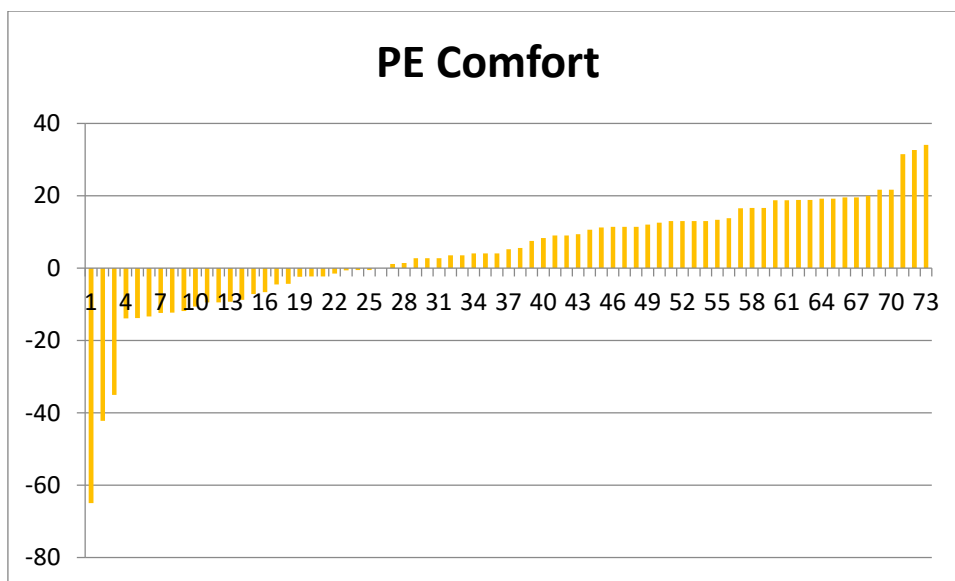
Γραφική 9-1 Ποσοστιαίο Σφάλμα για overall Rating



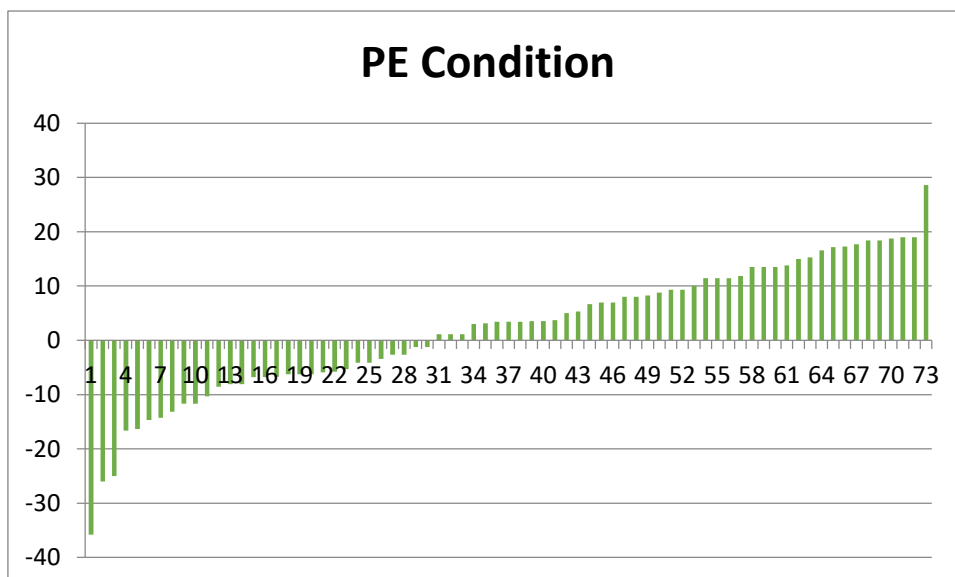
Γραφική 9-2 Ποσοστιαίο Σφάλμα για Καθαριότητα



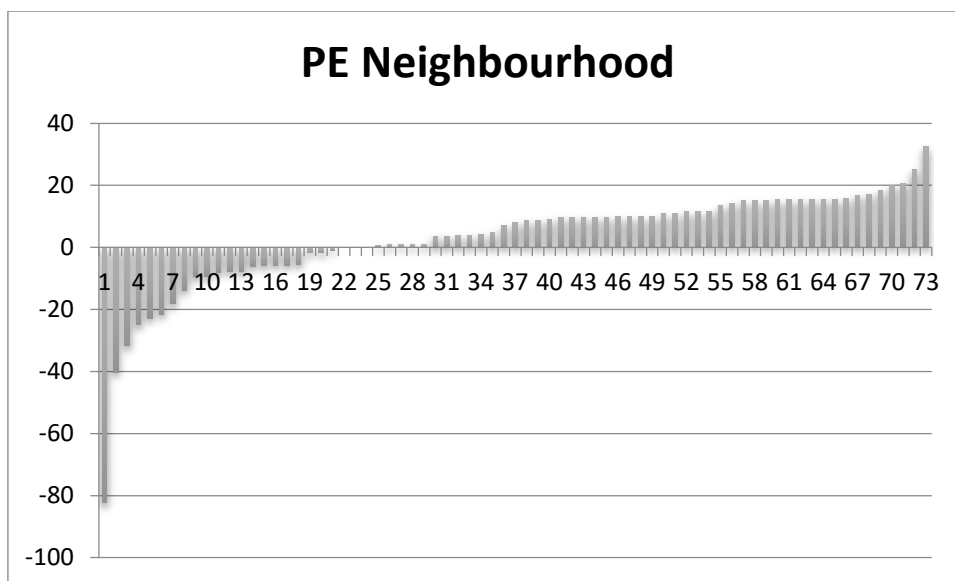
Γραφική 9-3 Ποσοστιαίο Σφάλμα για την Εξυπηρέτηση



Γραφική 9-4 Ποσοστιαίο Σφάλμα για Άνεση

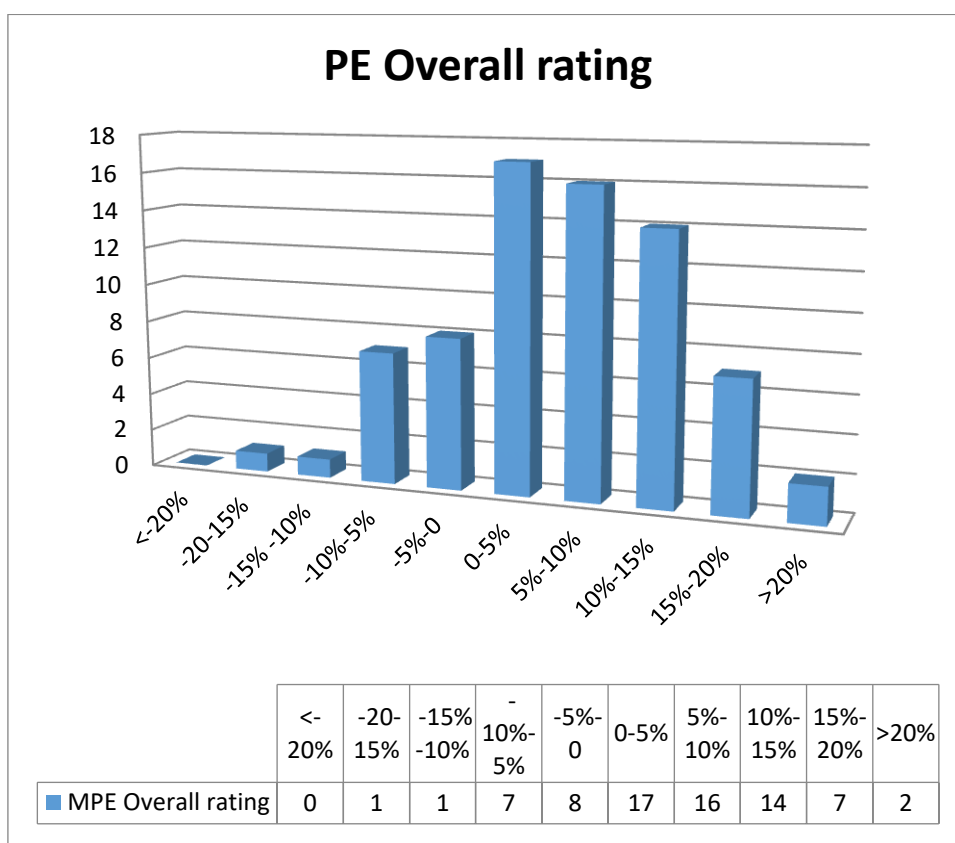


Γραφική 9-5 Ποσοστιαίο Σφάλμα για Κατάσταση

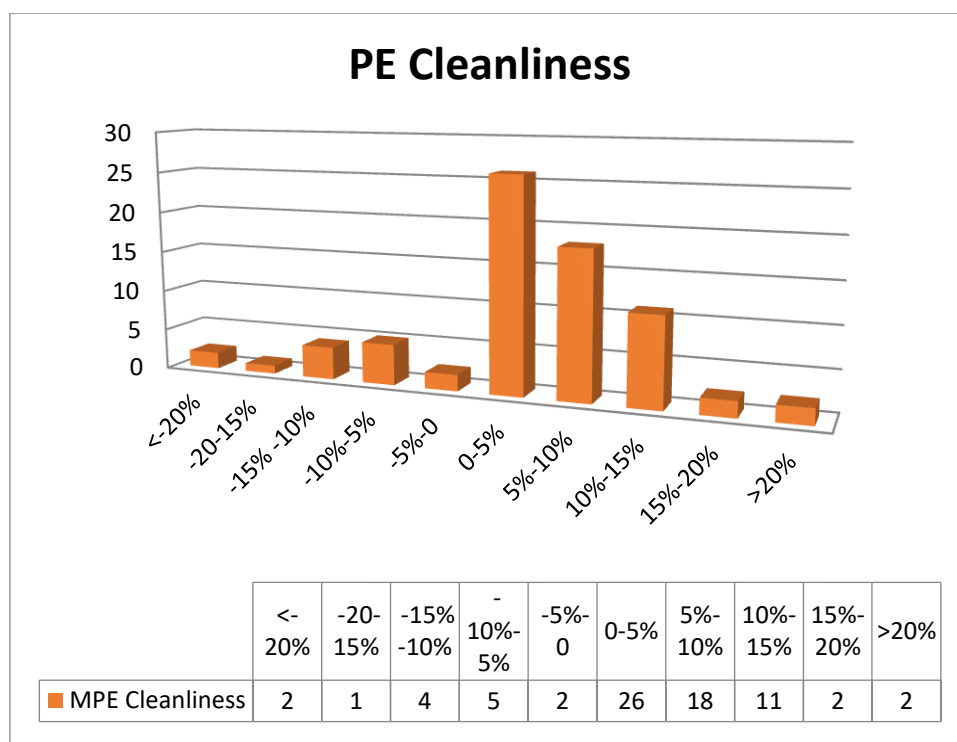


Γραφική 9-6 Ποσοστιαίο Σφάλμα για Γειτονιά

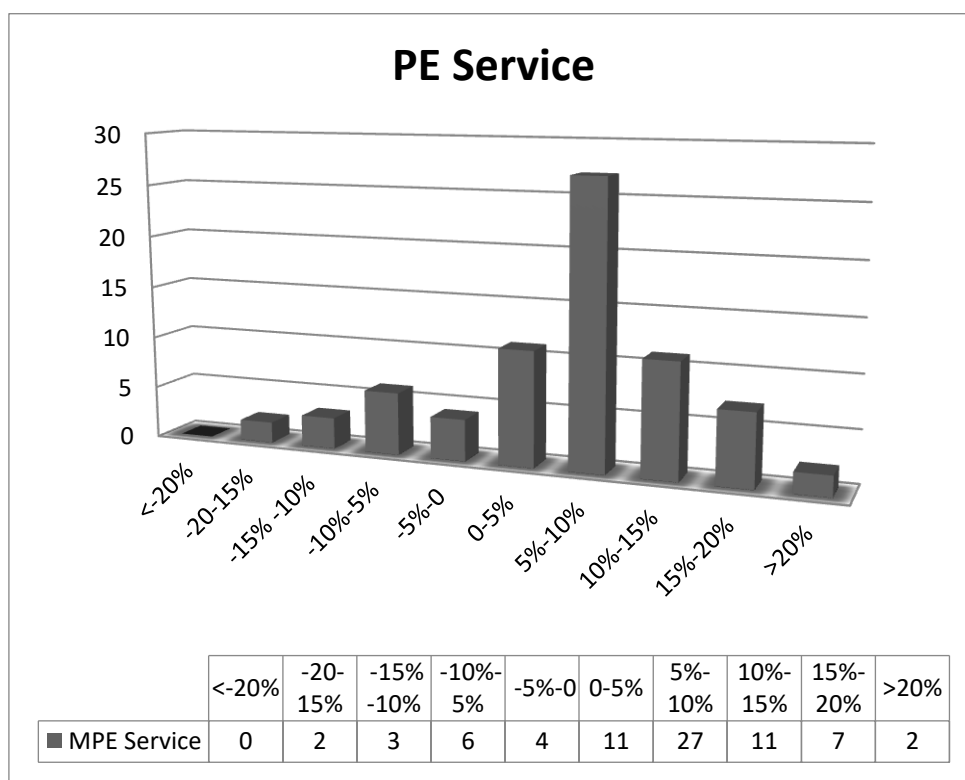
### Μέτρηση τιμών PE για τιμές -20% ως 20%



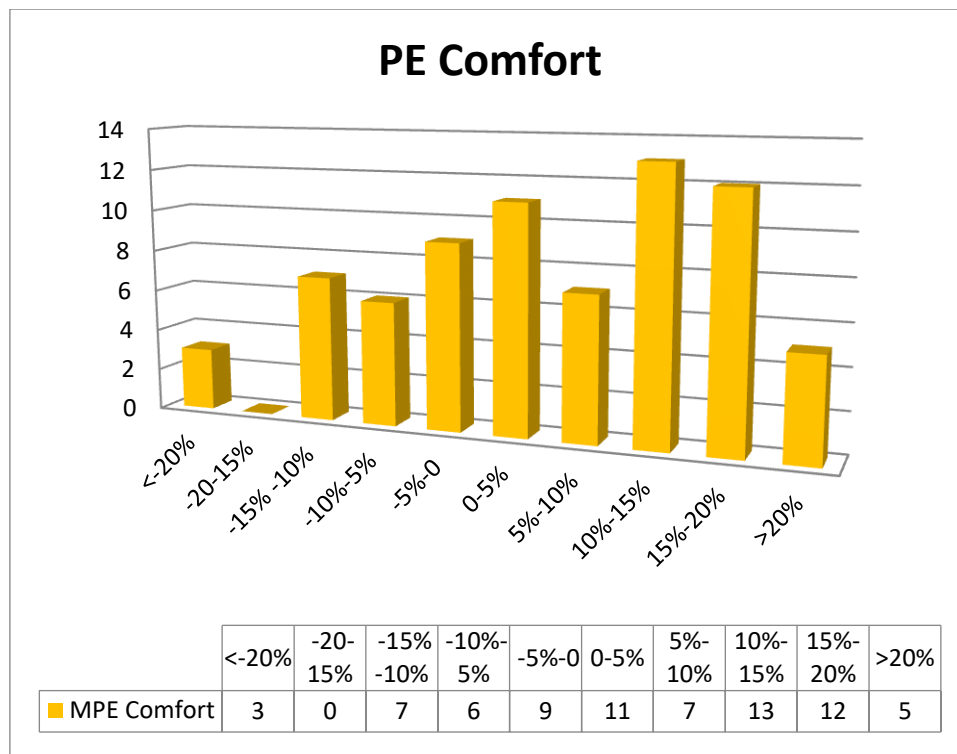
Γραφική 9-7 Μέτρηση τιμών PE για τιμές -20% ως 20% για overall Rating



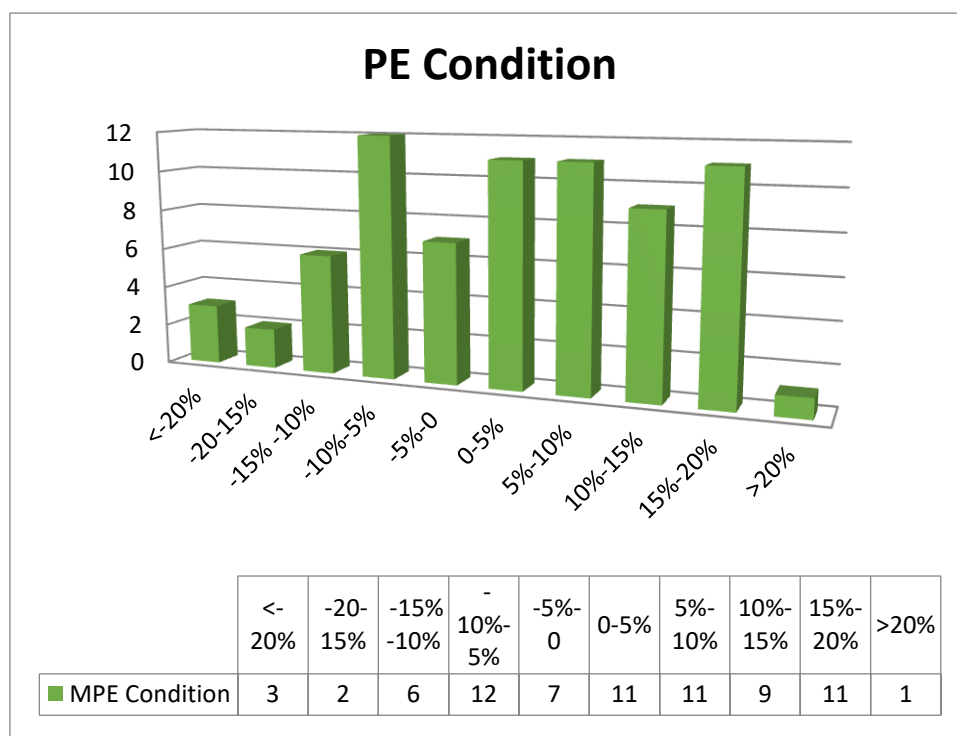
Γραφική 9-8 Μέτρηση τιμών PE για τιμές -20% ως 20% για Καθαριότητα



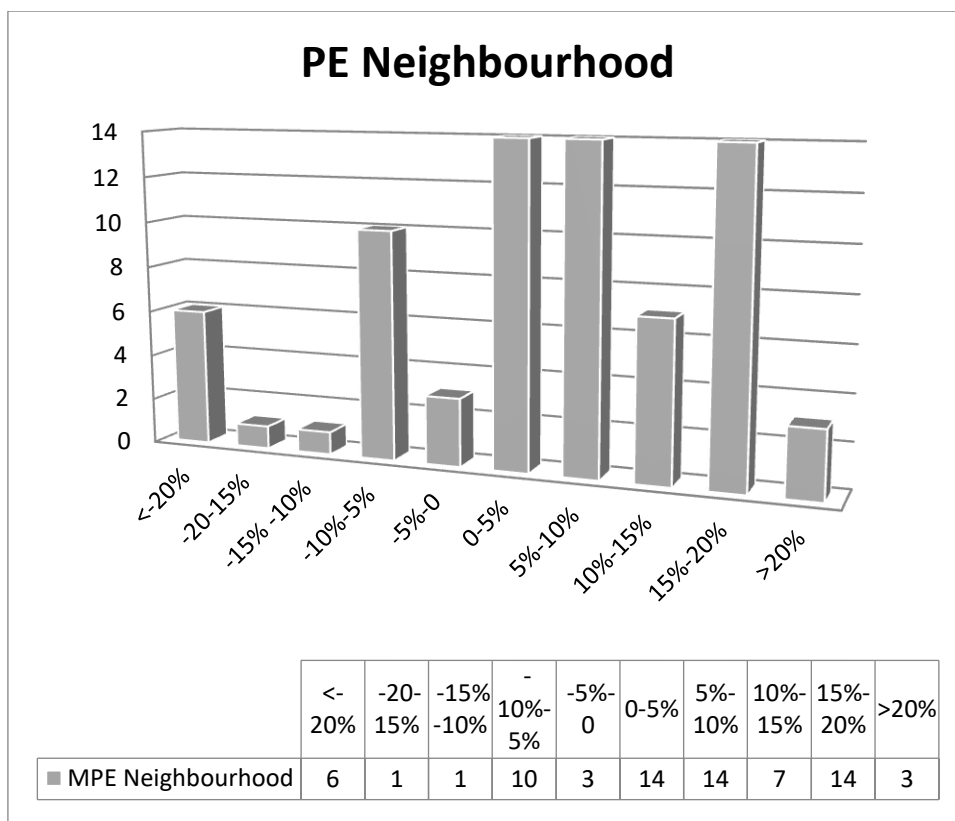
Γραφική 9-9 Μέτρηση τιμών PE για τιμές -20% ως 20% για Εξυπηρέτηση



Γραφική 9-10 Μέτρηση τιμών PE για τιμές -20% ως 20% για Άνεση



Γραφική 9-11 Μέτρηση τιμών PE για τιμές -20% ως 20% για Κατάσταση



Γραφική 9-12 Μέτρηση τιμών PE για τιμές -20% ως 20% για Γειτονιά

## **Σύντομο Βιογραφικό Σημείωμα**

Ο Ιωάννης Ζήσος είναι απόφοιτος της Σχολής Ηλεκτρονικών Μηχανικών και Μηχανικών Η/Υ του Πολυτεχνείου Κρήτης, με μεταπτυχιακό τίτλο στο τμήμα Ηλεκτρολόγων Μηχανικών του Μετσόβιου Πολυτεχνείου και MBA στο École des hautes études commerciales de Paris (HEC Paris).

Δραστηριοποιείται στο τομέα της πανεπιστημιακής εκπαίδευσης και διδάσκει στο Πανεπιστήμιο Δυτικής Αττικής, στη σχολή Διοίκησης και Οικονομίας. Στο παρελθόν έχει εργαστεί σαν σύμβουλος διαχείρισης χαρτοφυλακίου για αμερικανική εταιρεία διαχείρισης κεφαλαίων, σαν σύμβουλος σε product marketing και analytics στην Amadeus IT group στην Γαλλία και σαν υπεύθυνος διαχείρισης αναπτυξιακών συγχρηματοδοτούμενων έργων για διεθνή και ελληνικά έργα. Επίσης είχε την τιμή να υπηρετήσει σαν σύμβουλος ψηφιακής στρατηγικής του Ειδικού Γραμματέα Διαχείρισης Ευρωπαϊκών Πόρων του Υπουργείου Παιδείας της Ελλάδας, επί διετίας. Έχει διατελέσει επί δετία σαν ερευνητικός και επιστημονικός συνεργάτης διδάσκοντας σε προπτυχιακά και μεταπτυχιακά προγράμματα πολλαπλών πανεπιστημίων.

Τα ερευνητικά του ενδιαφέροντα συμπεριλαμβάνουν τις μεθοδολογίες συστημάτων συστάσεων, ανάλυση συναισθήματος και κειμένου, στρατηγικό μάρκετινγκ, μάρκετινγκ στο τομέα του τουρισμού, ανάλυση συμπεριφοράς, στρατηγικό management, κοινωνικά δίκτυα, νευρωνικά δίκτυα.