



## Πανεπιστήμιο Πειραιώς – Τμήμα Πληροφορικής

Πρόγραμμα Μεταπτυχιακών Σπουδών

«Προηγμένα Συστήματα Πληροφορικής»

### Μεταπτυχιακή Διατριβή

Τίτλος Διατριβής	Βιβλιογραφική Επισκόπηση Συστημάτων Κοινωνικής Σύστασης  Literature Review of Social Recommendation Systems
Όνοματεπώνυμο Φοιτητή	Χρήστος Βακόντιος
Πατρώνυμο	Δημήτριος
Αριθμός Μητρώου	ΜΠΣΠ/ 15008
Επιβλέπων	Διονύσιος Σωτηρόπουλος, Επίκουρος Καθηγητής

Ημερομηνία Παράδοσης **Αύγουστος 2021**

---

Τριμελής Εξεταστική Επιτροπή

(υπογραφή)

(υπογραφή)

(υπογραφή)

Διονύσιος Σωτηρόπουλος  
Επίκουρος Καθηγητής

Γεώργιος Τσιχριτζής  
Καθηγητής

Ευθύμιος Αλέπης  
Αναπληρωτής  
Καθηγητής

## Περίληψη

Τα κοινωνικά δίκτυα αποτελούν πηγή ξεκάθαρων αλλά και σιωπηρών πληροφοριών για τους χρήστες- συνδρομητές τους. Ο όγκος των πληροφοριών με εμπορικό περιεχόμενο που μπορούν να βρεθούν πλέον στα κοινωνικά δίκτυα είναι τεράστιος. Σε αυτό το τομέα έχει επικεντρωθεί κάποια έρευνα που εστιάζει στο κατά πόσο μπορούν να υλοποιηθούν συστήματα αυτοματοποιημένων προτάσεων προς τους χρήστες λαμβάνοντας υπόψη την αλληλεπίδραση τους στα κοινωνικά δίκτυα. Τα συστήματα αυτά ονομάζονται social recommendation systems (SRs). Καθώς οι χρήστες ξοδεύουν όλο και περισσότερο χρόνο σε κοινωνικά δίκτυα και μοιράζονται πολύτιμες πληροφορίες για τον εαυτό τους και τις σχέσεις τους, η χρήση αυτών των πληροφοριών συμβάλει στην βελτιστοποίηση των μηχανισμών δημιουργίας αυτοματοποιημένων προτάσεων στα παραπάνω συστήματα.

Η βιβλιογραφική έρευνα στα πλαίσια αυτής της πτυχιακής εργασίας ανέδειξε ωστόσο ότι πρέπει ακόμα να λυθούν ορισμένα προβλήματα. Πρώτον, τα ανεπεξέργαστα δεδομένα απαιτούν περαιτέρω προσοχή, όπως πώς πρέπει να γίνει η εξαγωγή των πιο χρήσιμων λέξεων στα σχόλια ενός χρήστη, ποιες έννοιες ή συναισθήματα σχετίζονται με αυτές τις λέξεις, τι προτάσεις πρέπει να δημιουργηθούν για αυτές τις έννοιες ή συναισθήματα του χρήστη. Παραμένει επίσης άλυτο το πρόβλημα της ασάφειας, και της βελτιστοποίησης των φίλτρων είτε βάσει περιεχομένου ή βάσει των σχέσεων που διαμορφώνουν οι χρήστες στα κοινωνικά δίκτυα.

Η εργασία αυτή αναδεικνύει διαφορετικά μοντέλα που είτε προσπαθούν να μοντελοποιήσουν τον τρόπο που οι κοινωνικές σχέσεις ενός ατόμου μπορούν να επηρεάσουν τις αποφάσεις του στην αγορά ενός προϊόντος, πως χρήστες με επιρροή σε άλλα άτομα επηρεάζουν εν τέλει τις αποφάσεις των τελευταίων και πολύ σημαντικό, ποια είναι η συμβολή του παράγοντα της εμπιστοσύνης (ή μη) μεταξύ των χρηστών. Αναδεικνύουμε επίσης σχετικές μετρικές που έχουν καταγραφεί ως τώρα στη βιβλιογραφία καθώς και αρχιτεκτονικές SR συστημάτων.

## Abstract

Social networks have become a source of clear and also implicit information for their users. The amount of information with commercial content that can now be found on social networks is vast. In this area, some research has focused on whether automated recommendation systems can be implemented, taking into account the user interactions on social networks. These systems are called social recommendation systems (SRs). As users spend more and more time on social networks and share valuable information about themselves and their relationships, the use of this information helps to optimize the mechanisms for creating automated proposals in the above systems.

The literature review in the context of this dissertation highlighted the need to address some problems which remain open challenges. First of all, raw data requires further attention, such as how to extract the most useful words in a user's comments, what concepts or emotions are associated with those words, and what suggestions should be made about those user's concepts or emotions. The problem of ambiguity, and the optimization of filters, either on the basis of content or on the basis of the relationships that users form on social networks, also remains unresolved.

This work highlights different methods that either attempt to model how an individual's social relationships can influence their decisions to buy a product, how influencers ultimately influence the decisions of the latter and most importantly, what the contribution of the factor of trust (or not) between users. We also highlight relevant metrics that have been recorded so far in the literature as well as SR system architectures.

## Περιεχόμενα

Περίληψη .....	3
Abstract .....	4
1. Εισαγωγή.....	6
1.1 Προκλήσεις των SR συστημάτων .....	8
2. Συστήματα Προτάσεων.....	10
2.1 Κύρια χαρακτηριστικά.....	10
2.2 Σενάρια Προτάσεων.....	11
2.3 Η σημασία των SR συστημάτων στις online επιχειρήσεις.....	13
2.4 Η σημασία των SR συστημάτων για τους χρήστες .....	14
2.5 Κατηγοριοποίηση των SR συστημάτων .....	16
2.5.1 SR συστήματα βάσει περιεχομένου.....	16
2.5.2 SR συστήματα βάσει ανάλυσης κοινωνικών σχέσεων .....	17
3. Social Recommendation συστήματα .....	24
3.1 Μέσα κοινωνικής δικτύωσης.....	24
3.2 Κοινωνικά δίκτυα .....	28
3.3 SR Αρχιτεκτονικές.....	29
3.3.1 Σχηματισμός μοντέλου γειννίαςης βάση την εξάπλωση των επιρροών .....	30
3.3.2 Πρόβλεψη ταξινόμησης με βάση παράγοντες κοινωνικής κανονικοποίησης ...	31
3.3.3 Η Προσέγγιση Implicit Social Trust and Sentiment (ISTS).....	35
3.3.4 Η προσέγγιση H-ISTS .....	41
4. Μετρικές αξιολόγησης των Social Recommendation συστημάτων .....	44
4.1 Τεχνικές ανατροφοδότησης των SR συστημάτων .....	44
4.1.1 Τυπικοί μηχανισμοί ανατροφοδότησης .....	46
4.2 Αναγνώριση χρηστών με επιρροή .....	46
4.3 Μετρικές εμπιστοσύνης στα SR συστήματα.....	47
5. Συμπεράσματα.....	57
Βιβλιογραφία .....	60

## 1. Εισαγωγή

Οι πληροφορίες είναι ένα πολύ ισχυρό εργαλείο. Οι πληροφορίες που συλλέγονται από το κοινωνικό μας περίγυρο μας βοηθούν να σχηματίζουμε απόψεις όπως και να λαμβάνουμε αποφάσεις με βάση αυτές τις απόψεις. Όχι μόνο είμαστε επηρεασμένοι από το κοινωνικό μας περίγυρο, αλλά ένα ευρύτερο κοινωνικό σύνολο (πχ φίλοι φίλων) μπορούν να επηρεάσουν τις αποφάσεις μας.

Ο πρώτος που ασχολήθηκε με το ζήτημα της κοινωνικής επιρροής στις αποφάσεις των ανθρώπων ήταν ο Granovetter στα μέσα της δεκαετίας του 1960. Διαπιστώνουμε ότι οι κοινωνιολογικές έννοιες που ανέπτυξε ο Granovetter μπορούν να εφαρμοστούν στα τρέχοντα διαδικτυακά κοινωνικά δίκτυα επίσης. Το γεγονός ότι μόνο λίγα άτομα μπορούν να επηρεάσουν μεγάλες πλειοψηφίες ατόμων σε ένα κοινωνικό δίκτυο μπορεί να αποτελέσει ένα ισχυρό εργαλείο για τη μεγιστοποίηση της κοινωνικής προσέγγισης σε αυτό το δίκτυο. Αυτές οι πληροφορίες μπορούν να χρησιμοποιηθούν αποτελεσματικά, για παράδειγμα, στη διαφημιστική βιομηχανία με τη προβολή μερικών δημοφιλών διασημοτήτων στα μέσα κοινωνικής δικτύωσης για τη διαφήμιση αγαθών και υπηρεσιών προς τις ευρύτερες κοινωνικές μάζες. Επίσης κυβερνητικές υπηρεσίες μπορεί να ακολουθήσουν αυτή τη προσέγγιση προκειμένου να στείλουν ένα κοινωνικό μήνυμα στο μέγιστο δυνατό σύνολο όταν η κοινωνία αντιμετωπίζει κρούσματα και καταστάσεις έκτακτης ανάγκης.

Τα social recommendation (SR) συστήματα αξιοποιούν τις κοινωνικές σχέσεις για τη βελτίωση της διαδικασίας δημιουργίας μίας σειράς προτάσεων κατά προτεραιότητα (Aggarwal, 2016), με βάση την υπόθεση ότι οι προτιμήσεις ενός χρήστη είναι πιθανόν να είναι παρόμοιες ή επηρεασμένες από εκείνες των φίλων του (Guy, 2015). Έρευνες έχουν δείξει ότι οι χρήστες επηρεάζονται κυρίως από τους άμεσους «φίλους» στα κοινωνικά μέσα δικτύωσης ή στον περίγυρο τους.

Πράγματι, οι πιο σημαντικές δυνατότητες των τρεχόντων μέσων κοινωνικής δικτύωσης (ΜΚΔ) είναι αφενός το δωρεάν περιεχόμενο που παράγεται από διαφορετικούς χρήστες και το δεύτερο είναι το προηγμένο τεχνολογικό περιβάλλον μέσα στο πλαίσιο των οποίων λειτουργούν τα ΜΚΔ. Οι χρήστες μπορούν να προσεγγίσουν εύκολα τους ιστότοπους και να μοιραστούν τις απόψεις τους σε διαθέσιμες πλατφόρμες όπως το Facebook, ή το Twitter. Κατά συνέπεια, οι άνθρωποι αντιμετωπίζουν ένα τεράστιο ζήτημα που αφορά την δημιουργία πληροφοριών με εκθετικό τρόπο. Το ζήτημα αυτό είναι προβληματικό για δύο λόγους. Το πρώτο πρόβλημα είναι η δυσκολία των χρηστών στην εύρεση περιεχομένου που είναι σχετικό με τις ανάγκες τους ή εξεύρεση μίας λύσης μεταξύ ενός τεράστιου αριθμού εναλλακτικών λύσεων. Το δεύτερο πρόβλημα αφορά τις διαδικτυακές επιχειρήσεις και τη μεγάλη πρόκληση που αντιμετωπίζουν για την παροχή εξατομικευμένων προϊόντων και λύσεων στο τρέχον διαδικτυακό περιβάλλον.

Τα SR συστήματα παρέχουν εξατομικευμένες προτάσεις με πληροφορίες για προϊόντα ή υπηρεσίες που σχετίζονται με τις ανάγκες του χρήστη. Θεωρούνται ισχυρά εργαλεία που βοηθούν τους χρήστες να βρουν ενδιαφέροντα αντικείμενα που ταιριάζουν στις προτιμήσεις τους. Αν και τα συστήματα αυτά έχουν κάνει ουσιαστική πρόοδο στη θεωρία και την ανάπτυξη αλγορίθμων και έχουν επιτύχει πολλές εμπορικές επιτυχίες, οι ευρέως διαθέσιμες πληροφορίες στα μέσα κοινωνικής δικτύωσης δεν έχουν αξιοποιηθεί ακόμα σε μεγάλο

βαθμό. Καθώς οι επιλογές ενός χρήστη επηρεάζονται σε μεγάλο βαθμό από την εμπιστοσύνη των φίλων και τις απόψεις τους, στην έρευνα προτείνονται διάφορες εξειδικευμένες προτάσεις.

Τα SR συστήματα ορίζονται ως αποτελεσματικά εργαλεία που παρέχουν εξατομικευμένες προτάσεις πληροφοριών ή προϊόντων που σχετίζονται με τις ανάγκες των χρηστών. Στην πραγματικότητα, τα συστήματα αυτά έχουν την εξουσιοδότηση να παρέχουν προτάσεις τόσο σε ιδιώτες- πελάτες όσο και στις διαδικτυακές επιχειρήσεις όταν χρησιμοποιούν πλατφόρμες, όπως Amazon, Netflix και Ebay. Παραδοσιακά, τα συστήματα αυτά μπορούν να κατηγοριοποιηθούν σε α) συστήματα που προτείνουν κάποιου είδους φιλτράρισμα στο περιεχόμενο (Content Filtering), και β) συστήματα που κάνουν προτάσεις με βάση το περιεχόμενο (Content Based) (Adomavicius and Tuzhilin, 2005), (Ricci et al, 2011).

Οι λύσεις τύπου CF χρησιμοποιούν ένα ιστορικό αξιολόγησης χρηστών και κάνουν μία ανάλυση της ομοιότητας μεταξύ του προφίλ χρήστη και εκείνων των γειτόνων του. Για χρήστες που έχουν υψηλή ομοιότητα στο ιστορικό τους, υπάρχει η δυνατότητα να μοιραστούν τα ίδια ενδιαφέροντα στο μέλλον με βάση τα κοινά χαρακτηριστικά του προφίλ τους (Resnick et al, 1994), (Sarwar et al, 2001).

Οι λύσεις τύπου CB αναγνωρίζουν χαρακτηριστικά που εμφανίζονται στα περιεχόμενα στοιχείων που έχουν οι χρήστες χρησιμοποιήσει προηγουμένως, προτείνοντας στη συνέχεια περισσότερα στοιχεία που περιέχουν σχετικά χαρακτηριστικά στους χρήστες. Οι τεχνικές CB εφαρμόζονται ευρέως σε αντικείμενα που περιέχουν κείμενο, συνεπώς επιτρέπει στο σύστημα να αντλήσει λέξεις-κλειδιά που αφορούν τα πιο σημαντικά χαρακτηριστικά και περιλαμβάνονται στην αγορά ενός βιβλίου ή άρθρου και στη συνέχεια να προτείνουν περισσότερα βιβλία / άρθρα που περιέχουν παρόμοιες λέξεις-κλειδιά. Αυτές οι δυνατότητες και οι λέξεις-κλειδιά χρησιμοποιούνται για τη δημιουργία προφίλ με τα στοιχεία του χρήστη (Pazzani and Billsus, 1997), (Si and Jin, 2003).

Αρκετά πρόσφατα, οι ερευνητές σε αυτό το τομέα ενσωμάτωσαν και τις σχέσεις εμπιστοσύνης μεταξύ των χρηστών για τη βελτίωση της λειτουργίας των παραδοσιακών SR συστημάτων προτείνοντας περαιτέρω προϊόντα που βασίζονται σε αξιόπιστες συνδέσεις (Ma et al, 2011), (Kim and Phalak, 2012), (Birtolo and Ronca, 2013). Επίσης πρόσφατα, έχουν προταθεί διάφορες προσεγγίσεις για τον εμπλουτισμό των προτάσεων με βάση κριτικές προϊόντων στο διαδίκτυο. Αυτές οι προσεγγίσεις χρησιμοποιούν τεχνικές επεξεργασίας φυσικής γλώσσας για να προσδιορίσουν το συναίσθημα στις κριτικές που έχουν γράψει οι χρήστες σχετικά με τα προϊόντα (Pang et al, 2002), (Ganu et al, 2009), (Leung et al, 2011).

Πράγματι τα συστήματα αυτά βελτιώνουν τις υπάρχουσες προσεγγίσεις αντλώντας δεδομένα από τα μέσα κοινωνικής δικτύωσης και από σύντομες δημοσιεύσεις φίλων σε blogs. Στην περίπτωση νέων χρηστών που δεν εμφανίζουν προηγούμενες προτιμήσεις, το σύστημα χαρτογραφεί τις προτεινόμενες προτάσεις σε αριθμητικά κλίμακες αξιολόγησης εφαρμόζοντας ορισμένα κύρια στοιχεία. Το πρώτο συστατικό είναι η μέτρηση της σιωπηρής εμπιστοσύνης μεταξύ φίλων με βάση τις δραστηριότητες, τις συνομιλίες τους και τη συμπεριφορά τους. Λόγω της ανάγκης προσαρμογής των απόψεων των φίλων, το μοντέλο κοινωνικής εμπιστοσύνης έχει σχεδιαστεί για να περιλαμβάνει τους αξιόπιστους φίλους και να τους δίνει το υψηλότερο βάρος συνεισφοράς σε συνάντηση. Το δεύτερο συστατικό

συνάγει μία βαθμολογία που αντικατοπτρίζει τα συναισθήματα, και τις απόψεις που εξάγονται από σύντομες δημοσιεύσεις των φίλων.

Τα συναισθήματα εξάγονται χρησιμοποιώντας τεχνικές συναισθηματικής ανάλυσης (Sentiment Analysis - SA). Για την επίτευξη της καλύτερης αναπαράστασης συναισθημάτων μπορεί να εφαρμοστούν διάφορες μεθοδολογίες κυρίως βασισμένες σε πιθανότητες. Ένα τρίτο συστατικό προσδιορίζει τον βαθμό αντίκτυπου των συναισθημάτων των φίλων και το επίπεδο εμπιστοσύνης τους χρησιμοποιώντας αλγόριθμους μηχανικής μάθησης, για παράδειγμα μοντέλα ταξινόμησης και μοντέλα παλινδρόμησης.

## 1.1 Προκλήσεις των SR συστημάτων

Στις παραδοσιακές τεχνικές προτάσεων οι πληροφορίες αξιολόγησης που χρησιμοποιούνται για την δημιουργία των προφίλ χρηστών-στοιχείων συνήθως επηρεάζονται αρνητικά από ένα παράγοντα βραδύτητας λόγω του γεγονότος ότι η βαθμολόγηση της συμπεριφοράς των ατόμων περιορίζεται σε ένα μικρό μέρος των διαθέσιμων στοιχείων, συνήθως μόνο το 1% των προϊόντων / υπηρεσιών και αυτό επηρεάζει την αποτελεσματικότητα των αλγορίθμων CF και CB [8]. Ως αποτέλεσμα, οι νέοι χρήστες βιώνουν το λεγόμενο cold-start πρόβλημα καθώς οι παραδοσιακοί αλγόριθμοι CF και CB παρέχουν λιγότερο ισχυρές προτάσεις για όσους δεν είχαν προηγούμενες επιλογές και αξιολογήσεις [5]. Συνεπώς νέα έρευνα μπορεί να διερευνήσει την ενίσχυση των παραπάνω πηγών δεδομένων και των προτάσεων ειδικά για νέους χρήστες με βάση τις κοινωνικές τους συνδέσεις στα ΜΚΔ.

Ο βαθμός κοινωνικής εμπιστοσύνης σε SR συστήματα που βασίζεται στην εμπιστοσύνη προϋποθέτει ότι οι χρήστες στα συστήματα είναι φίλοι μεταξύ τους και ότι καταχωρούν κάποια αξιολόγηση στα διάφορα αγαθά και υπηρεσίες. Αν και στα ΜΚΔ όπως το Facebook και το Twitter άτομα συνεργάζονται με τους φίλους, τους συγγενείς και τους συναδέλφους τους στον πραγματικό κόσμο, με αυτήν την πτυχή κοινωνικής εμπιστοσύνης δεν ισχύει σε πολλές περιπτώσεις στα τρέχοντα SR συστήματα που βασίζονται στην εμπιστοσύνη. Συνολικά, οι τρέχουσες μεμονωμένες διαδικτυακές σχέσεις στα ΜΚΔ δεν αποτελούν ακριβή αντανάκλαση των κοινωνικών σχέσεων [17].

Η εξαγωγή των απόψεων των χρηστών από γραπτές κριτικές σε ΜΚΔ για εμπορικά προϊόντα δεν έχει εφαρμοστεί πλήρως για την άντληση πιο εξατομικευμένων προτάσεων. Ωστόσο, τα σχόλια των φίλων περιλαμβάνουν πολλές απόψεις και ενδιαφέροντα για μία διαφορετική γκάμα προϊόντων και υπηρεσιών, όπως βιβλία, ταινίες και εστιατόρια και άλλα. Ωστόσο, λόγω της έμφυτης έλλειψης διαδικτυακών σχολίων και της περιορισμένης χρήσης χαρακτήρων στο περιβάλλον των ΜΚΔ, και της έλλειψης μεθόδων όπως η συχνότητα εμφάνισης μίας λέξης - κλειδί, και άλλων τεχνικών που χρησιμοποιούνται για την ταξινόμηση των χρηστών σε διαφορετικά προφίλ ενδέχεται να μην είναι οι πιο κατάλληλοι αλγόριθμοι.

Για παράδειγμα, οι ισότοποι αξιολόγησης σχετικά με εστιατόρια περιλαμβάνουν πολλά σχόλια για την καθοδήγηση του πελάτη και κριτικές όπως: τοποθεσία, γεύση, κ.λπ. Αυτό δίνει ένα πιο δομημένο μακρύ κείμενο που επιτρέπει στους πελάτες να περιγράψουν λεπτομερώς τις απόψεις τους. Επιπλέον, αρκετά συχνά σε αυτή την ερευνητική περιοχή εξετάζεται η άτυπη χρήση γλωσσικών χαρακτηριστικών στα ΜΚΔ, όπως emoticon εικόνες, καθώς



πιστεύουμε ότι παρέχουν μία καθαρή εικόνα για τα συναισθήματα και τις απόψεις στο κείμενο.

Επομένως ένα SR σύστημα καλείται να αξιοποιήσει τις πληροφορίες που παράγουν οι χρήστες και ο κοινωνικός τους περίγυρος στα ΜΚΔ. Συνεπώς παρουσιάζονται διάφορα ερευνητικά ερωτήματα:

- Πώς μπορούμε να αντλήσουμε συστάσεις από ΜΚΔ για συγκεκριμένα προϊόντα /υπηρεσίες; Υπάρχουν κατάλληλα εργαλεία για τη συλλογή τους; Πώς οι απόψεις των φίλων μπορούν να χρησιμοποιηθούν;
- Πώς μπορεί το μοντέλο εμπιστοσύνης να χιτιστεί επιλέγοντας φίλους που μπορούν να παρέχουν προτάσεις; Το φιλτράρισμα θα πρέπει να βασίζεται σε διάφορα χαρακτηριστικά και κριτήρια για την διαμόρφωση ενός δείκτη εμπιστοσύνης.
- Ποια είναι η πιο αποτελεσματική μέθοδος για να αντλήσουμε το σωστό συναίσθημα που περιγράφει τις απόψεις των φίλων ενός ατόμου από τις δημοσιεύσεις τους;
- Πώς μπορεί να σχεδιαστεί μια αποτελεσματική μέθοδος που να προβλέπει τις βαθμολογήσεις των νέων χρηστών;
- Πόσο αποτελεσματικό μπορεί να είναι το φιλτράρισμα όταν συγχωνεύουμε τις προτιμήσεις των απλών χρηστών με τις προτιμήσεις φίλων μέσα από το ΜΚΔ;
- Πόσο αποτελεσματική μπορεί να είναι η CF μέθοδος λαμβάνοντας υπόψη τη πρόκληση της διαφορετικότητας στα προφίλ χρηστών;

Το κεφάλαιο 2 παρουσιάζει μία εισαγωγή στα Συστήματα δημιουργίας αυτοματοποιημένων προτάσεων (Recommendation Systems – RS), τα κύρια χαρακτηριστικά και τη χρησιμότητα τους στη σύγχρονη ψηφιακή οικονομία. Το κεφάλαιο 3 εισάγει τη πτυχή των μέσων κοινωνικής δικτύωσης και πως εξελίσσουν τα παραπάνω συστήματα σε Κοινωνικά Συστήματα Προτάσεων (Social Recommendation Systems – SR). Παρουσιάζονται διαφορετικές αρχιτεκτονικές ενώ στο κεφάλαιο 4 παρουσιάζονται σημαντικές μετρικές που λαμβάνονται υπόψη στη διαμόρφωση προτάσεων όπως οι μηχανισμοί ανατροφοδότησης και η εμπιστοσύνη. Το κεφάλαιο 5 ολοκληρώνει την εργασία με τα Συμπεράσματα της εργασίας.

## 2. Συστήματα Προτάσεων

Τα Συστήματα Δημιουργίας Προτάσεων/Εισηγήσεων (RS) είναι ισχυρά εργαλεία και τεχνικές που επεξεργάζονται τους πόρους διαδικτυακών επιχειρήσεων για τη δημιουργία χρήσιμων προτάσεων στοιχείων που αντιστοιχούν στα ενδιαφέροντα και τις ανάγκες χρηστών. Τα είδη / προϊόντα μπορούν να είναι οποιοσδήποτε από τους πόρους των επιχειρήσεων που αγοράζουν οι χρήστες στο Διαδίκτυο, όπως βιβλία, ηλεκτρονικά είδη, είδη ρουχισμού, κ.λπ. Τα RS παράγουν προτάσεις και συμβάλουν στη λήψη προτάσεων από τον χρήστη για ποια είδη να αγοράσει ή ποιες ειδήσεις πρέπει να διαβάσει στο διαδίκτυο.

Σε γενικές γραμμές, ο μηχανισμός ενός συστήματος συστάσεων βασίζεται στην ανάλυση ιστορικών δεδομένων σχετικά με τα σχόλια κάθε χρήστη για στοιχεία που χρησιμοποιήθηκαν προηγουμένως σε μία παραγγελία. Προτείνει μια λίστα από αντικείμενα που εκτιμά ότι ταιριάζει στις προτιμήσεις του χρήστη. Οι χρήστες βομβαρδίζονται πλέον από έναν τεράστιο αριθμό επιλογών από προϊόντα που διατίθενται μέσα από τους διάφορους ιστότοπους. Ως εκ τούτου, τα RS χρησιμεύουν για να φιλτράρουν αυτές τις εναλλακτικές λύσεις με εξατομικευμένο τρόπο και προτείνουν εκείνα τα προϊόντα που πιθανότατα θα προσελκύσουν την προσοχή του χρήστη.

Τα RS έχουν εμφανιστεί ως ανεξάρτητη και σημαντική υπηρεσία [7] την τελευταία δεκαετία την οποία οι επιχειρήσεις ηλεκτρονικού εμπορίου έχουν ενσωματώσει ως υπηρεσία παροχής εξατομικευμένων συστάσεων και πλέον ως σημαντικό μέρος των συστημάτων τους για την ανάπτυξη της κερδοφορίας τους. Επιχειρήσεις όπως το Amazon.com, ebay.com, Netflix (netflix.com), Ringo (ringo.com) [19], LastFm (last.fm) και Pandora (pandora.com), και ιστότοπους προτάσεων ειδήσεων, όπως το Google News [20].

### 2.1 Κύρια χαρακτηριστικά

Οι πιο σημαντικές δυνατότητες των σύγχρονων μέσων κοινωνικής δικτύωσης (ΜΚΔ) ομαδοποιούνται σε δύο κύριες κατηγορίες. Η πρώτη είναι το δωρεάν περιεχόμενο που παράγεται από διαφορετικούς χρήστες και το δεύτερο είναι το προηγμένο τεχνολογικό περιβάλλον μέσα στο οποίο λειτουργούν τα ΜΚΔ. Οι χρήστες μπορούν να προσεγγίσουν εύκολα τους διάφορους ιστότοπους και να μοιραστούν τις απόψεις τους στα διαθέσιμα ΜΚΔ, όπως το Facebook, wiki και Twitter. Κατά συνέπεια, οι άνθρωποι έρχονται αντιμέτωποι με μία εκθετική αύξηση της πληροφορίας κάτι που δημιουργεί μία προβληματική κατάσταση για δύο λόγους.

Το πρώτο πρόβλημα αφορά τους ίδιους τους χρήστες και τη δυσκολία που αντιμετωπίζουν στην εύρεση περιεχομένου που είναι σχετικό με τα δικά τους ενδιαφέροντα μεταξύ ενός τεράστιου αριθμού εναλλακτικών λύσεων. Το δεύτερο σχετίζεται με τις διαδικτυακές επιχειρήσεις και τη μεγάλη πρόκληση που αντιμετωπίζουν στην παροχή εξατομικευμένων προϊόντων και προτάσεων προς ένα παγκόσμιο πελατολόγιο.

Τα SR συστήματα μπορούν να κατηγοριοποιηθούν σε δύο κύριες τεχνικές: εκείνα που κατευθύνουν ένα φίλτράρισμα του περιεχομένου (CF) βάσει συνεργασίας (ανταλλαγής απόψεων σε μία ομάδα), και εκείνα που διαμορφώνουν προτάσεις με βάση το περιεχόμενο (CB) (Felfernig et al, 2013). Τα CF συστήματα χρησιμοποιούν ένα ιστορικό με πληροφορίες από προηγούμενες αξιολογήσεις χρηστών και περιέχουν επίσης τα αποτελέσματα της σύγκρισης μεταξύ του προφίλ αξιολόγησης του χρήστη και εκείνων των εταίρων του (στην ίδια ομάδα). Αναδεικνύονται τυχόν ομοιότητες μεταξύ τους. Για χρήστες που έχουν αρκετές ομοιότητες στο ιστορικό τους, υπάρχει δυνατότητα κοινής χρήσης των ίδιων ενδιαφερόντων στο μέλλον βάσει των κοινών χαρακτηριστικών στα προφίλ τους (Yue Liu et al, 2020).

Τα CB συστήματα αναγνωρίζουν τα χαρακτηριστικά που εμφανίζονται στα περιεχόμενα των διάφορων στοιχείων που έχουν χρησιμοποιήσει οι χρήστες προηγουμένως, προτείνοντας τους στη συνέχεια περισσότερα στοιχεία (προϊόντα ή υπηρεσίες) που περιέχουν παρόμοια χαρακτηριστικά. Οι CB τεχνικές εφαρμόζονται ευρέως σε αντικείμενα που περιέχουν κείμενο όπου το σύστημα μπορεί να αντλήσει τις πιο σημαντικές δυνατότητες / λέξεις-κλειδιά που περιλαμβάνονται σε μια αγορά πχ ενός βιβλίου ή ενός άρθρου και στη συνέχεια να προσφέρουν περισσότερα βιβλία / άρθρα που περιέχουν παρόμοιες λέξεις-κλειδιά.

Αυτές οι δυνατότητες και οι λέξεις-κλειδιά χρησιμοποιούνται για τη δημιουργία προφίλ με τα στοιχεία του χρήστη (Atahan and Sarkar, 2010). Στο παρελθόν, οι ερευνητές σε αυτό το τομέα των SR συστημάτων ενσωμάτωσαν τις σχέσεις εμπιστοσύνης των χρηστών για τη βελτίωση των παραδοσιακών τεχνικών προτείνοντας περαιτέρω προϊόντα που βασίζονται στο βαθμό αξιοπιστίας μεταξύ των κοινωνικών σχέσεων (Ma et al 2011), (Kim and Phalak, 2012). Πρόσφατα, έχουν προταθεί διάφορες προσεγγίσεις για τον εμπλουτισμό των συστάσεων βάσει κριτικών που διαμορφώνουν οι χρήστες για προϊόντα και είναι διαθέσιμες στο διαδίκτυο. Αυτές οι προσεγγίσεις χρησιμοποιούν τεχνικές επεξεργασίας φυσικής γλώσσας για να προσδιορίσουν το συναίσθημα στις κριτικές που έχουν γράψει οι χρήστες για προϊόντα (Zou et al, 2015).

## 2.2 Σενάρια Προτάσεων

Για τη καλύτερη δυνατή κατανόηση των διαφορετικών προσεγγίσεων στα SR συστήματα και ειδικά όσων συνδυάζουν γνώση μέσω των κοινωνικών δικτύων, είναι σημαντικό να αναφέρουμε τρία υπάρχοντα σενάρια στην παροχή προτάσεων:

- Στην πραγματική ζωή, οι άνθρωποι συνήθιζαν να ζητούν συμβουλές από τους φίλους και τους συναδέλφους τους. Αυτές οι προσωπικές συνδέσεις παρέχουν τις απαιτούμενες προτάσεις από στόμα σε στόμα.
- Μέσω του Διαδικτύου, τα RS συστήματα δίνουν προτάσεις στους χρήστες αντικείμενα που μπορεί να τους αρέσουν. Αυτοί που παρέχουν συμβουλές / προτάσεις χρησιμοποιούν κυρίως πληροφορίες μέσα από το ίδιο το σύστημα, όπως η ομοιότητα μεταξύ των προφίλ των εγγεγραμμένων χρηστών ή των αντικειμένων, π.χ. το Amazon.com, όπως φαίνεται στην Εικόνα 1. Αυτά τα συστήματα υιοθετούν πολλές διαφορετικές προσεγγίσεις. Για παράδειγμα, ένα σύστημα μπορεί να χρησιμοποιεί την ομοιότητα μεταξύ  $u_1$  και  $u_2$  ή την εμπιστοσύνη μεταξύ  $u_3$  και  $u_4$ . Ωστόσο, όλοι οι

χρήστες  $u_1$ ,  $u_2$ ,  $u_3$ ,  $u_4$  και  $u_5$  είναι εγγεγραμμένοι χρήστες στο σύστημα. Τα συστήματα μπορούν να εκμεταλλευτούν την ποικιλομορφία των πληροφοριών ξεκινώντας από τη δραστηριότητα της αγοράς, τις αξιολογήσεις των προϊόντων, τις κριτικές για τα προϊόντα και ολοκληρώνοντας με τη συμπεριφορά επικοινωνίας για τη δημιουργία προσωπικών προτάσεων.

- Πρόσφατα, ορισμένες διαδικτυακές επιχειρήσεις άρχισαν να προσφέρουν μια υπηρεσία στην οποία οι εγγεγραμμένοι χρήστες μπορούν να ενημερώσουν τους φίλους τους σε άλλα δημοφιλή μέσα κοινωνικής δικτύωσης για προϊόντα που τους άρεσαν ή αγόρασαν. Αυτές οι επιχειρήσεις χρησιμοποιούν μία από τις κορυφαίες δυνατότητες των σύγχρονων Web εφαρμογών, η οποία είναι το δωρεάν περιεχόμενο που δημιουργείται από τους χρήστες, και αξιοποιούν την ικανότητα να μοιράζονται τις απόψεις τους σε μικρό-ιστολόγια (τα λεγόμενα blogs). Για αυτό το λόγο παρατηρούμε ότι ποικίλες πληροφορίες σχετικά με αντικείμενα και υπηρεσίες στα κοινωνικά δίκτυα, μπορούν να προέρχονται από αυτό που έχουν βιώσει οι φίλοι όταν δοκίμασαν τα προϊόντα ή επισκέφτηκαν έναν εμπορικό ιστότοπο. Ως αποτέλεσμα, η ύπαρξη αυτής της υπηρεσίας αυξάνει το ποσό των πληροφοριών που ανταλλάσσονται μεταξύ φίλων και αυτό μπορεί να αξιοποιηθεί μεταξύ φίλων ως ένα πολύτιμο μέσο προτάσεων.



Εικόνα 1. Τα SR συστήματα αλλά και οι χρήστες συμβουλευονται τα προφίλ και το ιστορικό των χρηστών – συνδέσεων τους (Dimah et al 2015)

Ωστόσο, ένα διαφορετικό σενάριο για τη λήψη προτάσεων μπορεί να προσεγγιστεί ως εξής. Οι υποψήφιοι χρήστες αντί να συμβουλευονται την ομοιότητα μεταξύ ξένων και άγνωστων ατόμων που είναι εγγεγραμμένοι σε αυτόν που δίνει συμβουλές, μπορούν να στραφούν στους πραγματικούς φίλους τους για βοήθεια στην επιλογή αντικειμένων μεταξύ εναλλακτικών επιλογών. Πιο συγκεκριμένα, η εύρεση προϊόντων ή υπηρεσιών που μπορεί να προσελκύσουν την προσοχή των χρηστών μπορεί να πραγματοποιηθεί βάσει αξιόπιστων πηγών από τις επιλογές των φίλων των χρηστών.

Η διερεύνηση των επιλογών φίλων και η γνώση του βαθμού ικανοποίησής τους για εμπορικά προϊόντα, απαιτεί ανάλυση του τι γράφουν οι φίλοι για να περιγράψουν τα συναισθήματά τους για αυτά τα αντικείμενα. Τα blogs είναι εύκολες και απλές πλατφόρμες που επιτρέπουν στους φίλους να δημιουργούν και να δημοσιεύουν άμεσα σύντομα μηνύματα που καταγράφουν τη γνώμη τους για σχεδόν τα πάντα στον πραγματικό κόσμο. Για παράδειγμα, προκειμένου κάποιος να παρέχει σε ένα συνδεδεμένο χρήστη ενδιαφέρουσες επιλογές, παρόμοια με τη συνήθη διαδικασία που ακολουθούν οι άνθρωποι όταν δεν έχουν σχετικές πληροφορίες σχετικά με αντικείμενα που μπορεί να ενδιαφέρονται, είναι δυνατόν να ανατρέξει σε σχόλια που μπορεί να έχουν αναρτηθεί από τους διαδικτυακούς του φίλους.

Υποθέτοντας ότι εξετάζουμε προτεινόμενες ταινίες για παράδειγμα, ένας χρήστης Α και ένας φίλος του Β μπορούν να προτείνουν δύο ταινίες, «Ταινία 1» και «Ταινία 2» αντίστοιχα, σε ένα άλλο χρήστη Γ. Οι σχέσεις εμπιστοσύνης διαφέρουν μεταξύ του χρήστη Γ και των φίλων του Α και Β. Οι δραστηριότητες ενδοεπικοινωνιών του χρήστη Γ μπορούν να χρησιμοποιηθούν για να δείξουν ποιος είναι ο πιο αξιόπιστος φίλος από τους δύο, όπως η ενέργεια της αποστολής μηνυμάτων που λαμβάνονται από τους φίλους του, η αποθήκευση των αναρτήσεων των φίλων του σε μια λίστα αγαπημένων ή ο συνολικός αριθμός φίλων που εμφανίζονται στη λίστα φίλων. Μια άλλη πρόκληση είναι να αναλυθεί το επίπεδο του συναισθήματος που υπάρχει στις γραπτές δημοσιεύσεις των φίλων για αντικείμενα. Για παράδειγμα, εάν η σχέση εμπιστοσύνης μεταξύ του χρήστη Γ και του Α είναι χαμηλή και ο φίλος Α δείχνει υψηλή θετική γνώμη για μια συγκεκριμένη ταινία, ενώ ο φίλος Β έχει ισχυρότερη σχέση εμπιστοσύνης με τον χρήστη Γ, τότε η γνώμη του φίλου Β θα πρέπει να λαμβάνεται σοβαρά υπόψη και να προτείνεται.

Τα συστήματα προτάσεων μπορούν να χρησιμοποιούν τους φίλους των χρηστών για να βοηθήσουν στη διαδικασία προτάσεων. Μια προφανής συνέπεια της εκθετικής αύξησης των πληροφοριών που διατίθενται στους χρήστες είναι ότι τα συστήματα που απαρτίζονται από συνιστώσες αντιμετωπίζουν δύο προκλήσεις, η πρώτη πρόκληση είναι η δυσκολία που βιώνουν οι χρήστες να βρουν περιεχόμενο που να είναι συναφές με τα δικά τους ενδιαφέροντα, μεταξύ ενός τεράστιου αριθμού εναλλακτικών και η άλλη πρόκληση είναι η απαίτηση για μια σύγχρονη τεχνική που μπορεί να παρέχει εξατομικευμένες προτάσεις αξιοποιώντας πληροφορίες στο τρέχον περιβάλλον του διαδικτύου.

### 2.3 Η σημασία των SR συστημάτων στις online επιχειρήσεις

Οι διαδικτυακές επιχειρήσεις μπορούν να χρησιμοποιούν SR τεχνικές για τους ακόλουθους λόγους (Bobadilla et al, 2013):

- Αύξηση του αριθμού των πωλήσεων. Αυτός είναι κατά πάσα πιθανότητα ο βασικότερος λόγος που τα SR συστήματα εφαρμόζονται από τις διαδικτυακές επιχειρήσεις. Διαδικτυακά καταστήματα μπορούν να πουλήσουν ένα σχετικά μεγαλύτερο σύνολο προϊόντων σε σύγκριση με αυτά που το κάνουν χωρίς να χρησιμοποιούν SR τεχνικές (Bobadilla et al, 2011). Ακόμη και μη κερδοσκοπικοί οργανισμοί αποσκοπούν στην αύξηση του αριθμού των αντικειμένων που καταναλώνονται ή χρησιμοποιούνται από τους χρήστες. Για παράδειγμα, η χρήση

τεχνικών προτάσεων σε επιστημονικά περιοδικά στοχεύει στην αύξηση του αριθμού των άρθρων που διαβάζονται στον ιστότοπο.

- Αύξηση του εύρους των προϊόντων προς πώληση. Μια άλλη σημαντική συμβολή του SR συστήματος είναι να δώσει το κίνητρο στον χρήστη να εξερευνήσει νέα προϊόντα που ενδεχομένως να του αρέσουν. Για παράδειγμα, στον κλάδο της ενοικίασης ταινιών τα συστήματα μπορούν να νοικιάσουν τις πιο δημοφιλείς ταινίες εύκολα, αλλά με SR τεχνικές μπορούν να προωθήσουν και τις λιγότερο δημοφιλείς ταινίες προτείνοντας τις στους σωστούς χρήστες δηλαδή σε αυτούς που ταιριάζουν στις προτιμήσεις τους.
- Αύξηση της ικανοποίησης των χρηστών. Συστήματα που περιλαμβάνουν ένα συνδυασμό από ένα καλά σχεδιασμένο φιλικό περιβάλλον αλληλεπίδρασης με τον χρήστη και ένα σύστημα που παράγει προτάσεις που ενδιαφέρουν τον χρήστη είναι πιο πιθανό να βελτιώσουν την υποκειμενική κρίση του χρήστη για τη λειτουργία του συστήματος.
- Αύξηση της πιστότητας των χρηστών. Όσο πιο συχνά ο χρήστης επισκέπτεται ένα ιστότοπο τόσο πιο αποτελεσματικά το SR σύστημα θα κατανοήσει τις προτιμήσεις του χρήστη. Δεδομένου ότι το σύστημα αναλύει τις αλληλεπιδράσεις μεταξύ του χρήστη και του συστήματος τόσο διευρύνεται το προφίλ του χρήστη με περισσότερη γνώση και, κατά συνέπεια, οι μακροπρόθεσμοι χρήστες που επισκέπτονται τον ιστότοπο τακτικά αντιμετωπίζονται ως πολύτιμοι πελάτες που εξυπηρετούνται σύμφωνα με τις ανάγκες και τις προτιμήσεις τους, και αυτό με τη σειρά του αυξάνει την πιστότητα των χρηστών.
- Καλύτερη κατανόηση των επιθυμιών του χρήστη. Με τη συλλογή πληροφοριών για το χρήστη και τη διαμόρφωση εγγραφών καταγραφής σχολίων, αυτό δίνει τη δυνατότητα σε έναν πάροχο υπηρεσιών να χρησιμοποιήσει ξανά αυτήν τη γνώση σε άλλους διαφορετικούς στόχους, όπως η ενίσχυση των κριτηρίων παραγωγής προϊόντων που να ταιριάζουν με τα γούστα των χρηστών ή η βελτίωση της διαχείρισης των αποθεμάτων αντικειμένων αναλόγως. Για παράδειγμα, στο ταξιδιωτικό τομέα, η ανάλυση των δεδομένων που συλλέγει το SR σύστημα βοηθά τους ταξιδιωτικούς πράκτορες να αποφασίσουν αν πρέπει να προωθήσουν εμπορικά νέους προορισμούς βάσει των πληροφοριών που προέρχονται από τις επιλογές των πελατών.

Από την άλλη πλευρά, οι χρήστες μπορούν επίσης να χρησιμοποιούν ένα SR σύστημα για να επιτύχουν ένα διαφορετικό σύνολο στόχων όπως εξηγείται στην επόμενη παράγραφο.

## 2.4 Η σημασία των SR συστημάτων για τους χρήστες

Οι Herlocker et al. (2004) εξηγούν ότι οι χρήστες μπορούν επίσης να επωφεληθούν από διαφορετικά χαρακτηριστικά των συστημάτων προτάσεων που εφαρμόζονται όπως είναι τα παρακάτω:

- Αναζήτηση συγκεκριμένων προϊόντων ή αντικειμένων. Η βασική λειτουργία του συστήματος προτάσεων είναι να παρέχει μία σειρά συστάσεων που ταιριάζουν στις ανάγκες του χρήστη. Παρέχοντας στον χρήστη μια ταξινομημένη λίστα προτάσεων τον βοηθάει στην απόφαση του να επιλέξει μία από αυτές τις επιλογές λαμβάνοντας υπόψη την κατάταξη. Ωστόσο, ορισμένα συστήματα επιλέγουν να αποκρύψουν την προβλεπόμενη βαθμολογία των στοιχείων.
- Αναζήτηση των καλύτερων προϊόντων σύμφωνα με διάφορα κριτήρια. Ειδικά σε ιατρικές και χρηματοοικονομικές εφαρμογές, αυτοί οι τομείς είναι κρίσιμοι στις αποφάσεις των χρηστών και αφορούν μικρό αριθμό επιλογών. Ο χρήστης πρέπει να επωφεληθεί από όλες τις καλά εξεταζόμενες δυνατότητες και λαμβάνει υπόψη τυχόν κατάταξη ανάλογες περιγραφές σε τέτοιες περιπτώσεις.
- Αξιολόγηση της εμπειρίας προηγούμενων χρηστών. Η λίστα των προτάσεων μπορεί να δίνει έμφαση σε ορισμένα προϊόντα που περιέχουν τις κριτικές παλιότερων χρηστών και τις σχετικές προτιμήσεις τους.
- Προτάσεις για ακολουθία προϊόντων. Στη περίπτωση αυτή προτείνεται στο χρήστη μία σειρά προϊόντων που ταιριάζουν στα ενδιαφέροντα ενός χρήστη αντί για μία μεμονωμένη πρόταση. Για παράδειγμα αυτό γίνεται στην αγορά βιβλίων προτείνοντας μία σειρά σε ένα συγκεκριμένο θέμα, στο οποίο ο χρήστης έχει δηλώσει με κάποιο τρόπο ενδιαφέρον.
- Προτάσεις για δέσμη προϊόντων. Ο πελάτης λαμβάνει προτάσεις για ένα σύνολο προϊόντων που ταιριάζουν καλά μεταξύ τους. Για παράδειγμα, όταν παρέχετε στους χρήστες ένα πλάνο για υποψήφιο ταξίδι, που μπορεί να περιλαμβάνει πολλά αξιοθέατα, εστιατόρια, προορισμούς και καταλύματα υπηρεσίες ή ακόμη και ορισμένες ειδικές δραστηριότητες για παιδιά. Από τη πλευρά του πελάτη, αυτές οι εναλλακτικές λύσεις θα πρέπει να παρουσιάζονται ως μία ολοκληρωμένη πρόταση ταξιδιού.
- Καθοδήγηση στην περιήγηση σε ένα ιστότοπο. Ο απώτερος στόχος ορισμένων χρηστών είναι η περιήγηση στον κατάλογο χωρίς καμία πρόθεση αγοράς προϊόντων. Ως εκ τούτου, το σύστημα βοηθά στην κατανομή των προϊόντων σύμφωνα και με το εύρος της αναζήτησης του χρήστη.
- Επιλογή αξιόπιστου SR συστήματος. Ορισμένοι πελάτες θα προσπαθήσουν να δοκιμάσουν αυτά τα συστήματα για να αξιολογήσουν την ικανότητά τους στη δημιουργία αξιόπιστων προτάσεων. Ως εκ τούτου, ορισμένα συστήματα προσφέρουν αυτήν την υπηρεσία επιτρέποντας στους χρήστες να επαληθεύουν την απόδοσή τους.

- Βελτίωση του προφίλ των πελατών. Προκειμένου να επιτευχθεί ο στόχος της εξατομίκευσης, οι χρήστες τροφοδοτούν τα συστήματα με πληροφορίες σχετικά με τις προτιμήσεις και τις μη προτιμήσεις τους. Σε διαφορετική περίπτωση, τα συστήματα παρέχουν στον ενεργό χρήστη αυτό που παραδίδεται στους μέσους χρήστες.
- Αυτό-έκφραση. Ορισμένοι χρήστες σκοπεύουν να συνεισφέρουν με τις αξιολογήσεις τους μόνο για να εκφράσουν τις απόψεις και τις πεποιθήσεις τους. Ωστόσο, δεν δίνουν προσοχή στις προτάσεις άλλων.
- Συνεισφορά στη κοινή γνώμη. Το βασικό κίνητρο για ορισμένους χρήστες είναι να βοηθήσουν άλλους πελάτες με την εμπειρία τους και ωφελήσουν την κοινότητα με τις συνεισφορές τους. Για παράδειγμα, σε μια εταιρεία λιανικής, ένας χρήστης που έχει αγοράσει ένα αυτοκίνητο γνωρίζει ότι οι αξιολογήσεις θα υποστηρίξουν τα σχέδια άλλων αγοραστών και όχι των δικών τους για την επόμενη φορά που κάποιος θέλει να αγοράσει ένα αυτοκίνητο.
- Επηρεασμός της κοινής γνώμης. Ορισμένοι χρήστες έχουν τον απώτερο στόχο να επηρεάσουν τη γνώμη των άλλων για την αγορά συγκεκριμένων αντικειμένων που τους αρέσουν. Από την άλλη πλευρά, ορισμένα συστήματα υποφέρουν από κακόβουλους χρήστες που μπορεί να εκμεταλλευτούν το σύστημα μόνο για την προώθηση δικών τους συμφερόντων ή ακόμη και να υποβαθμίσουν συγκεκριμένα προϊόντα.

Στην πραγματικότητα, ένα καλά σχεδιασμένο SR σύστημα πρέπει να διατηρεί μια ισορροπία μεταξύ των αναγκών χρηστών και επιχειρήσεων ενώ προσπαθούν να ικανοποιήσουν και τα δύο.

## 2.5 Κατηγοριοποίηση των SR συστημάτων

### 2.5.1 SR συστήματα βάσει περιεχομένου

Οι τεχνικές αυτές (CB-SR) χρησιμοποιούν την περιγραφή των στοιχείων για δημιουργία αναπαραστάσεων των αντικειμένων και των προφίλ των χρηστών. Αναλύουν το περιεχόμενο ή τη περιγραφή στο προφίλ των χρηστών με βάση τα αντικείμενα που του άρεσαν πριν και μετά ταιριάζουν αυτό το περιεχόμενο με παρόμοιο περιεχόμενο άλλων αντικειμένων που δεν έχει δει ακόμα ο χρήστης. Με άλλα λόγια, αυτές οι μέθοδοι δημιουργούν συστάσεις αναλύοντας τα χαρακτηριστικά των αντικειμένων, και να προτείνουν παρόμοια αντικείμενα στο μέλλον.

Πολλά σύγχρονα CB-SR συστήματα δημιουργούν προτάσεις για στοιχεία που περιέχουν κείμενο, για παράδειγμα ιστότοποι, βιβλία, έγγραφα, ειδήσεις κ.λπ. Για παράδειγμα, σε μια εφαρμογή που συστήνει βιβλία, προκειμένου να προτείνονται βιβλία σε ένα συγκεκριμένο χρήστη, το σύστημα προσπαθεί να αναλύσει τις συσχετίσεις μεταξύ των βιβλίων που ο χρήστης έχει δώσει μεγάλη βαθμολογία στο παρελθόν (πχ για συγκεκριμένους συγγραφείς,



τύπους, θέματα κ.λπ.). Στη συνέχεια, μόνο βιβλία που έχουν υψηλό βαθμό ομοιότητας με αυτά που προτίμησε στο παρελθόν ο χρήστης θα προταθούν. Η προσέγγιση αυτή έχει παρόμοια χαρακτηριστικά με τις τεχνικές ανάκτησης πληροφοριών (Chowdhury, 2010). Η διαφορά μεταξύ τους είναι ότι σε αντίθεση με τις παραδοσιακές προσεγγίσεις ανάκτησης πληροφοριών, τα CB-SR συστήματα βασίζονται στη χρήση προφίλ χρηστών που περιέχουν πληροφορίες σχετικά με αντικείμενα που ένας χρήστης έχει προτιμήσει στο παρελθόν ή έχει δηλώσει ότι του αρέσουν. Το προφίλ του χρήστη μπορεί να δημιουργηθεί παρατηρώντας τη συμπεριφορά των χρηστών κατά τη πραγματοποίηση συναλλαγών και που σχετίζεται με ορισμένα αντικείμενα (Felfernig et al, 2013).

Οι τεχνικές CB έχουν πολλά πλεονεκτήματα: πρώτον, αυτές οι μέθοδοι βασίζονται μόνο στις προτιμήσεις των χρηστών που παρουσιάζονται στις αξιολογήσεις για τη δημιουργία προφίλ χρηστών και δεν απαιτείται εισαγωγή αξιολογήσεων άλλων χρηστών για τη δημιουργία αυτών των προφίλ. Δεύτερον, το σύστημα μπορεί να εκμεταλλευτεί τη λίστα λέξεων-κλειδιών για να δικαιολογήσει γιατί συνιστά ένα συγκεκριμένο αντικείμενο στον χρήστη. Επιπλέον, τα συστήματα αυτού του τύπου είναι σε θέση να προτείνουν νέα αντικείμενα που δεν έχουν λάβει ακόμη αξιολογήσεις, μια δυνατότητα που δεν έχουν άλλες τεχνικές. Παρά όλα αυτά τα πλεονεκτήματα, οι τεχνικές CB εξακολουθούν να παρουσιάζουν αρκετούς περιορισμούς [5, 26]:

- Περιορισμένη ανάλυση περιεχομένου: Αυτόματη εξαγωγή χαρακτηριστικών για την περιγραφή των στοιχείων σε κάποιο σύστημα θα είναι πολύ πιο δύσκολο να εφαρμοστεί όταν τα δεδομένα είναι γραφικά (εικόνες, ηχητικά αρχεία και βίντεο). Οι τεχνικές ανάκτησης πληροφοριών επιτυγχάνουν καλά αποτελέσματα εξαγωγή χαρακτηριστικών από έγγραφα κειμένου. Ωστόσο, κατά την περιγραφή εγγράφων από τις πιο σημαντικές λέξεις-κλειδιά τους, οι τεχνικές CB δεν μπορούν να διακρίνουν μεταξύ τους άρθρα υψηλής ποιότητας ή κακώς γραμμένα κείμενα όταν χρησιμοποιούνται οι ίδιες λέξεις-κλειδιά [19].
- Υπερβολική εξειδίκευση: Τα προτεινόμενα στοιχεία θα περιορίζονται σε αυτό που ο χρήστης έχει βαθμολογήσει στο παρελθόν. Με άλλους όρους, μόνο στοιχεία που ταιριάζουν με το ενσωματωμένο προφίλ του χρήστη θα προτείνονται στοιχεία με υψηλή βαθμολογία ομοιότητας. Για παράδειγμα, χρησιμοποιώντας ένα σύστημα σύστασης για εστιατόρια, ένας χρήστης χωρίς προηγούμενη εμπειρία της μεξικάνικης κουζίνας δεν θα λάμβανε ποτέ σύσταση για το καλύτερο Μεξικάνικο εστιατόριο στην πόλη.

### 2.5.2 SR συστήματα βάσει ανάλυσης κοινωνικών σχέσεων

Η τεχνική του Collaborative Filtering (CF) είναι μια από τις πρώτες και πιο επιτυχημένες τεχνικές, συμπεριλαμβανομένων εργαλείων όπως το σύστημα Tapestry, το σύστημα GroupLens και Video Recommender (Schafer et al, 2007). Οι CF τεχνικές λειτουργούν με την δημιουργία προφίλ των προτιμήσεων των πελατών (Shankar et al, 2017). Με άλλα λόγια, τα συστήματα αυτά προτείνουν προϊόντα σε ένα συγκεκριμένο χρήστη αντλώντας τις κοινές απόψεις από άλλους πελάτες. Αυτά τα συστήματα περιλαμβάνουν ένα συγκεκριμένο είδος

αλγορίθμων και τεχνικών για την αναγνώριση ενός συνόλου πελατών που ονομάζονται και γείτονες των χρηστών.

Οι γείτονες ορίζονται ως οι πελάτες που έχουν συναφή αξιολογήσεις προϊόντων και μοιράζονται ένα παρόμοιο στυλ κρίσης με έναν ενεργό χρήστη. Για παράδειγμα, σε ένα σύστημα συνεργατικών συστάσεων που σχετίζεται με ταινίες, το σύστημα προσπαθεί για να αναζητήσει γείτονες που έχουν παρόμοιες προτιμήσεις με έναν ενεργό χρήστη με βάση παρόμοιες ταινίες που έχουν αξιολογήσει στο παρελθόν. Στη συνέχεια, μόνο οι ταινίες που προτιμώνται περισσότερο από τους συνομήλικους του ενεργού χρήστη θα παρέχονται ως προτάσεις. Τα CF συστήματα βασίζονται στην υπόθεση ότι οι άνθρωποι που είχαν ταιριαστές προτιμήσεις στο παρελθόν θα εμφανίσουν τα ίδια μοτίβο και στο μέλλον.

Για παράδειγμα στον παρακάτω Πίνακα, απεικονίζονται τα προφίλ πέντε χρηστών. Παρουσιάζονται ως σειρές σε έναν πίνακα ταξινόμησης χρήστη-στοιχείου U1, U2, U3, U4, U5. Οι στήλες αντιπροσωπεύουν τα στοιχεία I1, I2, I3, I4. Τα προφίλ αξιολόγησης των χρηστών είναι  $pr1 = [4, f, 5, 5]$ ,  $pr2 = [4, 2, 1, \emptyset]$ ,  $pr3 = [5, \emptyset, \emptyset, 4]$ ,  $pr4 = [5, 3, \emptyset, \emptyset]$  και  $pr5 = [\emptyset, 3, 2, 1]$ . Για παράδειγμα, αν εξετάσουμε το U1 ως ενεργό χρήστη, κοιτάζοντας το προφίλ του, μπορούμε να δούμε ότι το U1 δεν αξιολόγησε το I2 και ο στόχος του συστήματος είναι να αποφασίσει αν το I2 είναι μια βολική πρόταση για το U1. Μόνο τρεις χρήστες U2, U4 και U5 έχουν χρησιμοποιήσει το I2 και έχουν ορισμένες απόψεις σχετικά με αυτό. Ωστόσο, τα U2 και U4 έχουν παρόμοιες προτιμήσεις σε ορισμένα στοιχεία με το U1. Επομένως, είναι καλύτερο να ελέγξουν τις προτιμήσεις τους για να εκτιμήσουν εάν το στοιχείο I2 είναι μια καλή ή κακή πρόταση για το U1. Αντίθετα, το U5 έχει διαφορετικές προτιμήσεις από το U1 και ως εκ τούτου, η αξιολόγηση του U5 θα έχει πολύ μικρή επίδραση στην δημιουργία μίας πρότασης, ή μάλλον, η άποψη του U5 θα αγνοηθεί. Επιπλέον, το U3 θα αποκλειστεί από τη διαδικασία σύστασης καθώς ο χρήστης αυτός δεν έχει χρησιμοποιήσει και αξιολογήσει το I2 αν και έχει παρόμοιες προτιμήσεις με το U1.

	I1	I2	I3	I4
U1	4	?	5	5
U2	4	2	1	$\emptyset$
U3	5	$\emptyset$	$\emptyset$	4
U4	$\emptyset$	3	2	1

Οι συνεργατικοί αλγόριθμοι φιλτραρίσματος μπορούν να ταξινομηθούν σε δύο κύριες υποκατηγορίες: σε αλγόριθμους βάσει μνήμης και βάσει μοντέλων μηχανικής μάθησης.

#### 2.5.2.1 Αλγόριθμοι βάσει μνήμης

Τεχνικές βασισμένες στη μνήμη ή τη γεινίαση προτιμήσεων (Breese et al, 2013) χρησιμοποιούν ολόκληρο τον πίνακα στοιχείων χρήστη για τον υπολογισμό προβλέψεων προτάσεων για στοιχεία που δεν προβάλλονται ακόμα στον χρήστη. Αυτά τα συστήματα

εντοπίζουν παρόμοιους χρήστες, ή τους λεγόμενους γείτονες ή πλησιέστερους γείτονες. Γενικά, με βάση τη μνήμη, οι αλγόριθμοι αυτοί εργάζονται σύμφωνα με τα ακόλουθα βήματα:

1. ζητούν από τους χρήστες να αξιολογήσουν ορισμένα στοιχεία για να μπορούν μετά να αναγνωρίσουν τις προτιμήσεις τους.
2. εφαρμόζουν στατιστικές τεχνικές για τον καθορισμό των γειτόνων του χρήστη που μοιράστηκαν παρόμοιες προηγούμενες προτιμήσεις.
3. μετά τον εντοπισμό χρηστών με παρόμοιες προτιμήσεις, τα συστήματα παρέχουν προβλέψεις για αξιολογήσεις στοιχείων που ακόμα δεν έχει δει ο χρήστης, και παρέχουν επίσης ανάλογες προτάσεις.

Στο δεύτερο βήμα, χρησιμοποιείται μια μετρική για τον υπολογισμό της ομοιότητας μεταξύ δύο προφίλ. Οι γείτονες με κοντινές προτιμήσεις μπορούν να κατανεμηθούν με τους εξής τρόπους:

1. καθορίζοντας έναν ορισμένο αριθμό γειτόνων  $N$
2. με προκαθορισμό ενός κατάλληλου ορίου, επομένως μόνο γείτονες των οποίων το επίπεδο ομοιότητας υπερβαίνει το κατώτατο όριο ενσωματώνονται στη συνολική πρόταση, και
3. εξαιρώντας γείτονες που έχουν εξαιρετικά ανόμοιες προτιμήσεις που αντικατοπτρίζονται από μία αρνητική βαθμολόγηση στο προφίλ τους.

Στο τρίτο βήμα, τα συστήματα μπορούν να επιλέξουν να παρέχουν τις συστάσεις ως σκορ μίας βαθμολογίας και οι τιμές να εκφράζονται χρησιμοποιώντας την ίδια κλίμακα που χρησιμοποιήθηκε για να εκφράσουν οι χρήστες τις απόψεις τους, για παράδειγμα, μια κλίμακα από 1 έως 5. Ένας άλλος διαφορετικός τρόπος σύστασης θα ήταν να δημιουργηθεί μια ταξινομημένη λίστα με προτεινόμενα είδη.

#### *2.5.2.2 Αλγόριθμοι βάσει μοντέλων μηχανικής μάθησης*

Οι CF τεχνικές χρησιμοποιούν επίσης τεχνικές μηχανικής μάθησης για να κάνουν έξυπνες προβλέψεις. Η υλοποίηση τέτοιων μοντέλων περιλαμβάνει τρία βασικά βήματα:

- Τα μοντέλα μαθαίνουν πρώτα το μοτίβο των συλλεγόμενων αξιολογήσεων χρηστών λαμβάνοντας υπόψη ένα σύνολο δεδομένων για την εκπαίδευσή τους.
- Τα μοντέλα δοκιμάζονται και όλες οι απαιτούμενοι παράμετροι ρυθμίζονται για την ικανοποίηση του των απαιτήσεων του προβλήματος; για παράδειγμα, ελαχιστοποιώντας το απόλυτο τετραγωνικό σφάλμα.
- Μετά τη δημιουργία αξιόπιστων και δοκιμασμένων μοντέλων, μπορούν να προβλεφθούν αξιολογήσεις και προτάσεις που είναι στη διάθεση των χρηστών μέσα από αυτά τα μοντέλα.

Οι ερευνητές στον τομέα των SR συστημάτων έχουν προτείνει διαφορετικά μοντέλα μηχανικής μάθησης, για παράδειγμα:

- Ταξινόμηση: Οι συγγραφείς στο (Miyahara and Pazzani, 2002), χρησιμοποίησαν απλούς αλγόριθμους Bayesian CF σε μια συνεργατική διαδικασία φιλτραρίσματος των

προτάσεων. Χρησιμοποιώντας μια στρατηγική Naive Bayesian (NB), υπέθεσαν ότι τα χαρακτηριστικά είναι ανεξάρτητα, δεδομένης της κλάσης, για τον υπολογισμό της πιθανότητας μιας τάξης με όλες τις δυνατότητες.

- Ομαδοποίηση: CF αλγόριθμοι που βασίζονται σε τεχνικές ομαδοποίησης έχουν επίσης χρησιμοποιηθεί για τη βελτίωση της ποιότητας των προβλέψεων. Η βασική ιδέα πίσω από τις τεχνικές ομαδοποίησης είναι να δημιουργήσουν ομάδες με παρόμοια χαρακτηριστικά, όπως μια ομάδα χρηστών που έχουν παρόμοιες προτιμήσεις. Οι Sarwar et al. [8, 30] έχουν προτείνει τέτοιους αλγορίθμους βασισμένους στην ομαδοποίηση για τη βελτίωση της απόδοσης της ταχύτητας. Εφαρμόζουν αλγόριθμους δύο φάσεων:
  - i. ομαδοποίηση των ταξινομήσεων των χρηστών σε ομάδες, και
  - ii. χρήση αλγορίθμων βάσει μνήμης για διαμόρφωση πιθανών προτάσεων προς κάθε χρήστη εντός των ομάδων βάσει μόνο προτιμήσεων από άλλα μέλη της ίδιας ομάδας.

Τα πειράματα τους δείξαν ότι κάνοντας προβλέψεις για προτάσεις μέσα σε μικρότερες ομάδες, χρησιμοποιώντας k-means αλγορίθμους, βελτιώνεται η επεκτασιμότητα στις τεχνικές ομαδοποίησης τεχνικές σε σύγκριση με κλασικές τεχνικές CF, μειώνεται ο αριθμός των γειτονικών ομάδων που πρόκειται να αξιολογηθούν λόγω των στατικών προ-υπολογισμένων ομάδων και ως εκ τούτου, η διαδικασία πρόβλεψης που γίνεται σε πραγματικό χρόνο γίνεται πολύ ταχύτερη (Xue et al, 2005).

1. Παλινδρόμηση: CF αλγόριθμοι με βάση την παλινδρόμηση έχουν χρησιμοποιηθεί επίσης για τη πρόβλεψη προτάσεων με βάση τις αξιολογήσεις των χρηστών. Σχετικές έρευνες (Vucetic and Obradovic, 2005) πρότειναν ένα εύρος γραμμικών μοντέλων για την αναζήτηση ομοιοτήτων μεταξύ αντικειμένων. Αυτή η παλινδρομική προσέγγιση κατάφερε να συνδυάσει γραμμικά μοντέλα για τον υπολογισμό των προβλέψεων αξιολόγησης για έναν ενεργό χρήστη. Για να εκτιμηθούν οι παράμετροι της συνάρτησης γραμμικής παλινδρόμησης, οι συγγραφείς χρησιμοποίησαν μετρικές των ελάχιστων τετραγώνων. Έδειξαν ότι η προσέγγισή τους προσφέρει καλές επιδόσεις στην αντιμετώπιση της σποραδικότητας, ένα συχνά εμφανιζόμενο πρόβλημα στις CF μεθόδους.
2. Μήτρα παραγόντων: οι τεχνικές που έχουν προταθεί (Rennie et al, 2005) βασίζονται σε δυο πλεονεκτήματα αυτού του μοντέλου:
  - i. Έχει ένα ελκυστικό επίπεδο ακρίβειας και
  - ii. μία επαρκή επεκτασιμότητα σε μεγάλα σύνολα δεδομένων.

Η κεντρική ιδέα πίσω από αυτό το μοντέλο είναι η μοντελοποίηση τόσο των χρηστών όσο και των αντικειμένων ως εσωτερικού προϊόντος και παραγωγής ενός κοινού χώρου με παρόμοια χαρακτηριστικά. Τα χαρακτηριστικά αυτά προκύπτουν από την αλληλεπίδραση μεταξύ χρηστών και αντικειμένων και απεικονίζονται στη μήτρα αξιολογήσεων χρήστη-στοιχείου.

Αυτές οι δυνατότητες εξηγούν πώς ένας χρήστης μπορεί να αξιολογήσει ένα στοιχείο και το μοντέλο χρησιμοποιεί αυτές τις δυνατότητες για να δημιουργήσει τον πίνακα ταξινόμησης.

### 2.5.2.3 Προκλήσεις των CF αλγορίθμων

Οι CF τεχνικές έχουν πολλά χρήσιμα σημεία κατά την εφαρμογή τους. Ένα από τα πολλά σημαντικά πλεονεκτήματα τους είναι ότι μπορούν να παρέχουν προτάσεις χρησιμοποιώντας μόνο την άμεση πηγή δεδομένων που είναι τα προφίλ με την αξιολόγηση των χρηστών. Επιπλέον, όσο περισσότερο οι χρήστες ανταποκρίνονται στις συστάσεις, τόσο καλύτερα τα συστήματα μπορούν να προσαρμοστούν στις προτιμήσεις των χρηστών με την πάροδο του χρόνου. Με λίγα λόγια, τα εμπλουτισμένα προφίλ των χρηστών υποδηλώνουν προτάσεις καλύτερης ποιότητας.

Οι πιο απαιτητικές πτυχές στη χρήση CF προσεγγίσεων, για τις οποίες έχουν γίνει προσπάθειες για να ξεπεραστούν, περιλαμβάνουν:

- **επίπεδο ακριβής πρόβλεψης:** Τα περισσότερα μοντέλα προτάσεων επιδιώκουν υψηλά επίπεδα ακρίβειας στις προβλέψεις τους. Οι ερευνητές χρησιμοποιούν τις μετρικές σχετικά με την ακρίβεια με τη σειρά τους για να αποδείξουν πόσο επιτυχημένες (και αν) είναι οι προτεινόμενες προσεγγίσεις προτάσεων παρέχοντας προβλέψεις που πλησιάζουν στις προτιμήσεις των χρηστών (Bell and Koren, 2007).
- **Πρόβλημα ψυχρής εκκίνησης:** Και τα συστήματα σύστασης CB και CF υποφέρουν από το αποκαλούμενο πρόβλημα ψυχρής εκκίνησης. Για να μπορούν να παρέχουν ακριβείς προτάσεις, οι προτιμήσεις των χρηστών πρέπει να αναλύονται με τη συλλογή βαθμολογιών που δεν ισχύουν για τους νέες χρήστες που έχουν εισέλθει πρόσφατα στο σύστημα. Ως εκ τούτου, αυτό μπορεί να οδηγήσει σε αποτελέσματα με μη ικανοποιητικές προτάσεις. Παρομοίως, το πρόβλημα της ψυχρής εκκίνησης υπάρχει και στην περίπτωση νέων αντικειμένων. Τα συστήματα προτάσεων ενημερώνονται τακτικά με νέα αντικείμενα. Για να ενσωματωθούν νέα στοιχεία στη διαδικασία προτάσεων, αυτά τα στοιχεία πρέπει να βαθμολογούνται πρώτα με ένα ικανοποιητικό αριθμό χρηστών.
- **Το πρόβλημα της σποραδικότητας:** Μια κοινή παρατήρηση είναι ότι ο αριθμός των αντικειμένων που τα άτομα ταξινομούν σε οποιοδήποτε σύστημα προτάσεων είναι συνήθως πολύ μικρός σε σύγκριση με το σύνολο του αριθμού των στοιχείων που χρειάζονται πρόβλεψη της βαθμολογίας τους. Φαίνεται πως ο συνολικός αριθμός των αξιολογήσεων των χρηστών επηρεάζει την επιτυχία των προτάσεων των συνεργατικών μοντέλων. Για παράδειγμα, σε ταινίες που προτείνει το σύστημα, όταν μια ταινία έχει βαθμολογηθεί από λίγα άτομα ακόμα και αν έχει υψηλή βαθμολογία, αυτή η ταινία δεν θα προτείνεται εκτενώς. Επιπλέον, ένα κοινό πρόβλημα με χρήστες που έχουν ασυνήθιστες προτιμήσεις σε σύγκριση με το ευρύτερο κοινό είναι ότι τα συστήματα ενδέχεται να μην είναι σε θέση να εκχωρήσουν παρόμοιους χρήστες (Lu et al, 2015).

Άλλα ζητήματα αφορούν την επεκτασιμότητα των αλγορίθμων ως προς τεράστια και δυναμικά σύνολα δεδομένων. Αρκετές τεχνικές σε πραγματικά συστήματα ασχολούνται με μεγάλες αλληλεπιδράσεις μεταξύ χρηστών και στοιχείων (πχ προϊόντων) όπως βαθμολογίες, προτιμήσεις και κριτικές, κ.λπ. Αυτές οι λύσεις πρέπει να καθιερωθούν για να διασφαλιστεί ότι οι δοκιμές σε μη πραγματικό χρόνο επαρκούν κατά την εκτέλεση της πραγματικής εφαρμογής. Επιπλέον, η πλειονότητα των τρεχόντων ανεπτυγμένων συστημάτων δημιουργίας προτάσεων δημιουργεί συστάσεις όταν οι χρήστες το ζητούν ρητά. Ωστόσο, σήμερα, εμφανίζονται διαφορετικά σενάρια λόγω του γεγονότος ότι οι χρήστες είναι πάντα συνδεδεμένοι μέσω πολλαπλών υπολογιστών. Επομένως, υπάρχει ανάγκη για ένα σύστημα συστάσεων ικανό να εντοπίσει μη σαφή αιτήματα/ προτιμήσεις. Επίσης, η νέα γενιά συστημάτων συστάσεων θα πρέπει να προβλέπει πότε και πώς να δημιουργούνται προτάσεις που εστιάζουν στην πρόβλεψη παρά στην αντίδραση (Melguizo et al, 2009).

Η διαφορετικότητα είναι μια άλλη πρόκληση για έναν ενεργό χρήστη. Οι ερευνητές συμφωνούν ότι μια λίστα συστάσεων πρέπει να περιλαμβάνει ένα συγκεκριμένο επίπεδο ποικιλομορφίας στα προϊόντα ή είδη που προτείνει. Στα πρώτα στάδια της διαδικασίας δημιουργίας προτάσεων, οι χρήστες μπορεί να χρειαστεί να εξερευνήσουν διαφορετικά είδη και να δοκιμάσουν τις προτιμήσεις τους. Με άλλα λόγια, οι χρήστες μπορούν να χρησιμοποιήσουν το προτεινόμενο είδος ως εργαλείο ανακάλυψης γνώσης. Από ερευνητική άποψη, υπάρχουν κάποιες προσπάθειες για τη διαμόρφωση ενός πιο ξεκάθਾਰου ορισμού για τον όρο «ποικιλομορφία» και πώς να εξισορροπηθεί ο στόχος της ακρίβειας της σύστασης σε σχέση με την εξατομίκευση των προτιμήσεων των χρηστών έναντι του στόχου της διαφορετικότητας.

Επίσης το απόρρητο των χρηστών είναι ένα άλλο ζήτημα. Για την παροχή εξατομικευμένων προτάσεων, τα SR συστήματα συλλέγουν όσο το δυνατόν περισσότερα δεδομένα από τους χρήστες. Αυτές οι ενέργειες ενδέχεται να ενοχλήσουν ορισμένους χρήστες όταν το σύστημα φαίνεται να γνωρίζει και να κατανοεί τις πραγματικές προτιμήσεις του. Επομένως, είναι σημαντικό να παρέχουμε λύσεις με εμπιστευτική και λογική χρήση των δεδομένων του χρήστη. Είναι σημαντικό επίσης αυτές οι λύσεις να διασφαλίζουν την προστασία των χρηστών έναντι κακόβουλων επιθέσεων από τρίτους.

Οι μέθοδοι που χρησιμοποιούνται από συστήματα δημιουργίας προτάσεων μπορούν να χωριστούν σε δύο προσεγγίσεις:

- Αυτά που διαμορφώνουν ένα μακροπρόθεσμο προφίλ του χρήστη, με βάση δεδομένα της συμπεριφοράς χρηστών τους (πχ με τεχνικές όπως collaborative filtering) και
- Αυτά που διαμορφώνουν ένα βραχυπρόθεσμο προφίλ με τις προτιμήσεις χρηστών δηλαδή καταγράφουν προσωρινά τις προτιμήσεις των χρηστών.

Είναι σαφές ότι και οι δύο προσεγγίσεις είναι σημαντικές για διαφορετικούς λόγους. Στην πραγματικότητα, πρέπει να καταβληθούν νέες προσπάθειες για τον συνδυασμό και των δύο προσεγγίσεων για τη δημιουργία μοντέλων υβριδικού τύπου, τα οποία χρησιμοποιούν ακριβώς το περιεχόμενο των προτιμήσεων χρηστών. Επίσης, πρέπει να καταβληθούν προσπάθειες για την εξήγηση του όποιου συσχετισμού μεταξύ βραχυπρόθεσμων και μακροπρόθεσμων προτιμήσεων.

Η δυνατότητα δημιουργίας γενικών μοντέλων και συστημάτων επαναχρησιμοποίησης SR συστημάτων σε διαφορετικές εφαρμογές και τομείς. Αν και τα SR συστήματα σε ένα πεδίο εφαρμογών μπορούν να βασιστούν σε υβριδικές προσεγγίσεις, δεν μπορούν εύκολα να χρησιμοποιήσουν τα προφίλ προτιμήσεων των χρηστών για να παρέχουν προτάσεις σε διαφορετικό πεδίο εφαρμογών.

### 3. Social Recommendation συστήματα

#### 3.1 Μέσα κοινωνικής δικτύωσης

Με βάση την αλληλεπίδραση των χρηστών στον Ιστό, τα μέσα κοινωνικής δικτύωσης έχουν εξελιχθεί σε μεγάλο βαθμό τα τελευταία χρόνια, με αποτέλεσμα να έχει υπάρξει μια θεμελιώδη αλλαγή στον τρόπο που οι άνθρωποι ζουν και καταναλώνουν πληροφορίες. Η δύναμη δεν είναι πλέον στις ομάδες μάρκετινγκ ή δημοσίων σχέσεων. Το διαδίκτυο έχει αρκετά εκδημοκρατιστεί έτσι ώστε κάθε χρήστης να μπορεί να παίζει ρόλο στο διαδίκτυο σήμερα. Τα μέσα κοινωνικής δικτύωσης (ΜΚΔ) ανήκουν σε οικογένειες ιστότοπων κοινωνικών μέσων και καθίστανται απαραίτητα στην καθημερινή ζωή των ανθρώπων. Όλο και περισσότερο άτομα από διαφορετικές χώρες και ηλικιακές ομάδες τα χρησιμοποιούν σε καθημερινή βάση και ξοδεύουν χρόνο σε ιστότοπους κοινωνικών μέσων. Μάλιστα δημιουργούν δεδομένα ή δημοσιεύουν πληροφορίες που προσελκύουν την προσοχή των ερευνητών.

Για τους λόγους που αναφέραμε εκτενώς στο προηγούμενο κεφάλαιο, μία από τις πιο ελκυστικές πτυχές που προκύπτει για τους ερευνητές από αυτό το φαινόμενο είναι πώς να βελτιώσουν τις προτάσεις τους προς τους χρήστες βάσει αυτών των κοινωνικών πληροφοριών.

Τα μέσα κοινωνικής δικτύωσης αναφέρονται στην κοινωνική αλληλεπίδραση μεταξύ ατόμων μέσα στα οποία δημιουργούν, μοιράζονται ή ανταλλάσσουν πληροφορίες και ιδέες σε εικονικές κοινότητες και δίκτυα (Schafar et al, 2007). Τα κοινωνικά μέσα βασίζονται στις συνομιλίες (Victor et al, 2011), το περιεχόμενο που δημιουργείται από τον χρήστη, και όλο αυτό το περιεχόμενο μπορεί να εξαπλωθεί και να προσεγγίσει εκατομμύρια χρήστες, ειδικά αν το περιεχόμενο είναι ελκυστικό για τους υπόλοιπους χρήστες (Guy 2015).

Οι ιστότοποι κοινωνικών μέσων εμφανίστηκαν εδώ και πολλά χρόνια στο διαδίκτυο με αναφορές σε συστήματα από το 1997 (Victor et al, 2011). Πλέον υπάρχουν ήδη διαφορετικοί τύποι ΜΚΔ στο διαδίκτυο σήμερα. Κάποιοι από αυτούς είναι εφαρμογές όπως το Facebook, και Twitter. Ενώ άλλοι ειδικεύονται σε ορισμένους τομείς όπως: LinkedIn για επαγγελματικά προφίλ, YouTube και Instagram για κοινή χρήση βίντεο και φωτογραφιών, FarmVille για κοινωνικά παιχνίδια και TripAdvisor για ανταλλαγή απόψεων για μέρη και ταξίδια. Οι χρήστες κοινωνικών μέσων μπορούν να δημιουργήσουν το δικό τους προφίλ και κατόπιν χρησιμοποιούν εργαλεία για την προώθηση των λογαριασμών τους σε διαφορετικές πλατφόρμες που θα ενθαρρύνουν την αύξηση των ακόλουθων σε αυτούς τους λογαριασμούς.

Τα δε blog sites που δημιουργήθηκαν στα τέλη του 1990 γίνανε εξαιρετικά δημοφιλείς καθώς οι συγγραφείς τους είναι από απλοί χρήστες μέχρι επαγγελματίες συγγραφείς ή ακόμα και διασημότητες. Οι ιστότοποι Blogging είναι εύκολο να δημιουργηθούν και να διατηρηθούν. Σήμερα υπάρχουν πάνω από 570 εκατομμύρια blogs. Τα blogs γίνονται πηγή επηρεασμού της κοινής γνώμης και πολλές μηχανές δημιουργήθηκαν για την αναζήτηση ιστολογίων. Τα μέσα

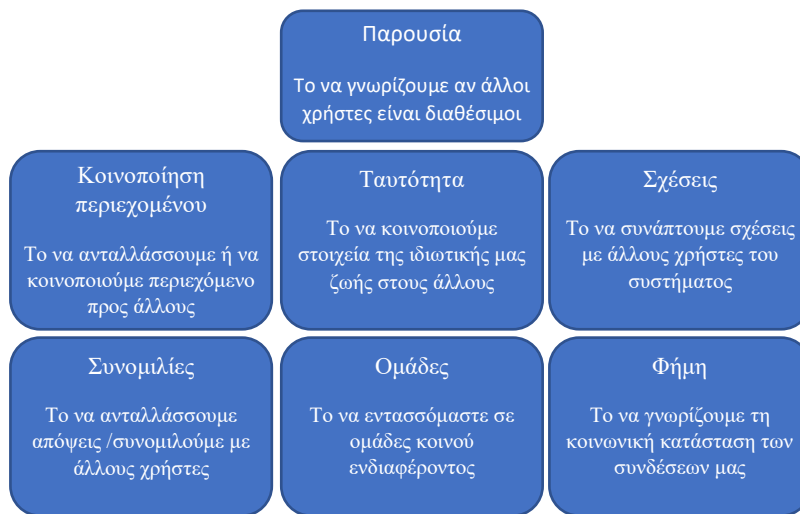


κοινωνικής δικτύωσης άνοιξαν επίσης την πόρτα για την κατάταξη των ιστότοπων με τις ψήφους των χρηστών για το περιεχόμενο, χρησιμοποιώντας blog sites όπως το Reddit.

Το micro blogging είναι μια άλλη πτυχή των κοινωνικών μέσων που προσφέρει ενημερώσεις σε πραγματικό χρόνο. Το πιο δημοφιλές εκ των οποίων είναι το Twitter. Οι χρήστες στέλνουν σύντομα μηνύματα (Tweets) που περιορίζονται σε 140 χαρακτήρες και είναι κυρίως σύντομες ενημερώσεις μίας κατάστασης για το τι κάνουν οι χρήστες, πού είναι, πώς αισθάνονται, συνδέσμους προς άλλους ιστότοπους, φωτογραφίες ή σύντομα βίντεο. Άλλοι σχετικοί ιστότοποι χρησιμοποιούν ένα μείγμα ενημερώσεων σε πραγματικό χρόνο και συγκεκριμένων πληροφοριών με τη τοποθεσία των χρηστών επιβραβεύοντας τους όταν ενημερώνουν για την άφιξη τους σε μία οποιαδήποτε τοποθεσία παγκοσμίως ή κατά την αποχώρησή τους. Τα σχόλια που προσθέτουν είναι προσβάσιμα προς όλους.

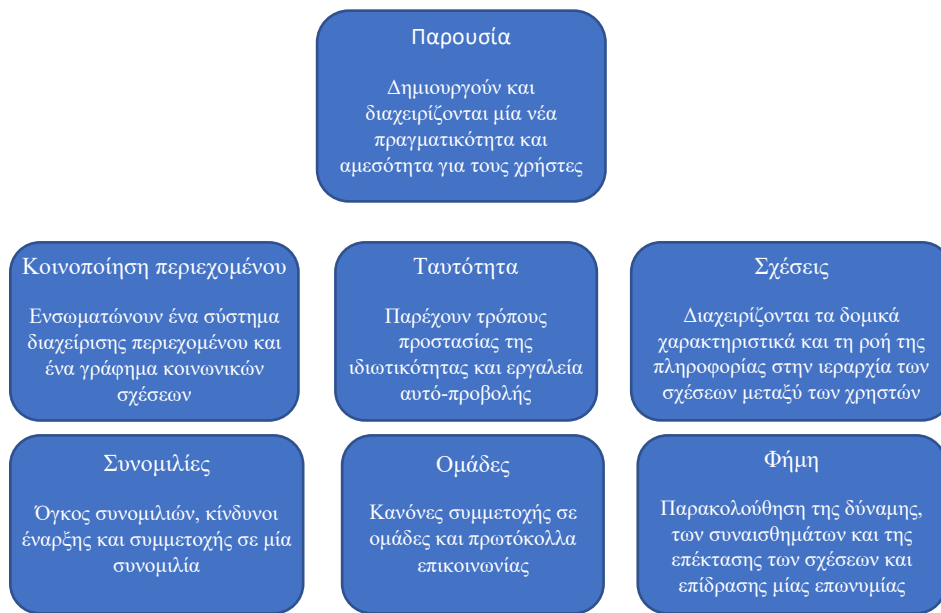
Με την ραγδαία άνοδο των κοινωνικών μέσων, η επικοινωνία έχει εκδημοκρατιστεί. Το μάρκετινγκ και οι δημόσιες σχέσεις είναι τώρα στα χέρια ατόμων και κοινοτήτων που δημιουργούν περιεχόμενο μέσω των διαφόρων καναλιών κοινωνικής δικτύωσης με ή χωρίς την άδεια επιχειρήσεων ή οργανισμών. Οι εταιρείες και οργανισμοί δεν μπορούν πλέον να αγνοήσουν την επίδραση των μέσων κοινωνικής δικτύωσης. Ένα πολύ γνωστό παράδειγμα είναι η υπόθεση της σπασμένης κιθάρας του Dave Carroll, ενός Καναδού μουσικού, από την United Airlines. Με βάση την απάντηση της αεροπορικής εταιρείας για την σπασμένη κιθάρα, ο Carroll ηχογράφησε ένα μουσικό βίντεο και μετά το μοιράστηκε στο YouTube εκείνη τη στιγμή και συγκέντρωσε περισσότερες από 14 εκατομμύρια προβολές. Ο οικονομικός προς την αεροπορική εταιρία εκτιμάται ότι ήταν γύρω στα 180 εκατομμύρια δολάρια!

Η αλληλεπίδραση των ατόμων στο διαδίκτυο και η αντίδραση τους προς μία εμπορική επωνυμία με θετικό τρόπο θεωρείται δωρεάν μάρκετινγκ που οδηγεί σε αυξανόμενη αναγνώριση της επωνυμίας, αύξηση πωλήσεων και ούτω καθεξής (Yang et al 2012). Αντίστοιχα, η αντίδραση των ατόμων με αρνητικό τρόπο μπορεί να προκαλέσει σημαντικές ζημιές (Yang Liu et al 2020). Ένα έξυπνα διατυπωμένο tweet, μία έξυπνη ανάρτηση στο διαδίκτυο, ένα βίντεο μπορεί να προκαλέσει το φαινόμενο της χιονοστιβάδας αν προωθηθεί σε εκατοντάδες φίλους και να καταστρέψει ένα προϊόν ή να βλάψει την τιμή της μετοχής μιας εταιρείας. Όλο αυτό το φαινόμενο αποκαλείται διαφορετικά η ηλεκτρονική έκδοση του από στόμα σε στόμα (electronic word of mouth – eWOM).



Εικόνα 2. Οι βασικές λειτουργίες των κοινωνικών μέσων (Kietzmann et al 2011)

Η κατανόηση της έννοιας τους του eWOM μπορεί να βοηθήσει στην περαιτέρω αξιοποίηση του από το ευρύτερο κοινό σύμφωνα με τις ανάγκες τους. Η Εικόνα 3 περιγράφει το πλαίσιο των λειτουργιών των κοινωνικών μέσων με επτά δομικά στοιχεία. Αυτά τα επτά στοιχεία μπορούν να βοηθήσουν στο να κατανοήσουμε τόσο τους υπάρχοντες ιστότοπους κοινωνικών μέσων, όσο και τις ανάγκες των οργανισμών. Αυτά τα επτά στοιχεία είναι η ταυτότητα, οι συνομιλίες, η κοινή χρήση, η παρουσία, οι σχέσεις, η φήμη και οι ομάδες. Κάθε στοιχείο επικεντρώνεται σε μια συγκεκριμένη πτυχή της εμπειρίας των χρηστών των κοινωνικών μέσων και των επιπτώσεών της στις επιχειρήσεις και τους οργανισμούς. Αυτά τα στοιχεία ποικίλλουν στην παρουσία τους και την επίδραση τους από το ένα κοινωνικό δίκτυο στο άλλο (Εικόνα 4).



Εικόνα 3. Οι επιπτώσεις των λειτουργιών των κοινωνικών μέσων (Kietzmann et al 2011)

- **Ταυτότητα:** Αντιπροσωπεύει τα όρια τα οποία θέτουν οι χρήστες σε πληροφορίες σχετικά με τη κοινοποίηση προσωπικών δεδομένων στα μέσα κοινωνικής δικτύωσης. Στοιχεία όπως το όνομα, την ηλικία, το φύλο, το επάγγελμα, η τοποθεσία ή ακόμα και πιο προσωπικές πληροφορίες όπως συναισθήματα, συμπάθειες και αντιπάθειες. Ωστόσο, οι ιστότοποι των κοινωνικών μέσων προσφέρουν διαφορετικές δυνατότητες κοινοποίησης πληροφοριών για τη ταυτότητα του χρήστη.
- **Συνομιλίες:** Είναι το κύριο χαρακτηριστικό των κοινωνικών μέσων. Οι άνθρωποι είναι εδώ για να επικοινωνήσουν με άλλους χρήστες και ομάδες. Οι χρήστες δημιουργούν tweets, δημοσιεύσεις, αναρτήσεις (blogs), αναζητούν ανθρώπους με κοινά ενδιαφέροντα, συναισθήματα, αναζητούν ένα ιδανικό βιβλίο, ή ακόμα και να χτίσουν αυτοεκτίμηση ή ανακαλύψουν νέες ιδέες ή δημοφιλή θέματα.
- **Κοινοποίηση:** Η δυνατότητα της κοινοποίησης αφορά το εύρος στο οποίο οι άνθρωποι ανταλλάσσουν, διανέμουν και λαμβάνουν περιεχόμενο. Ωστόσο, η κοινοποίηση από μόνη της είναι ένας τρόπος σύνδεσης στα κοινωνικά μέσα. Το ερώτημα είναι αν αυτό μπορεί να οδηγήσει στη δημιουργία σχέσεων μεταξύ των ανθρώπων μέσα από αυτή την αλληλεπίδραση. Αυτό εξαρτάται μεταξύ πολλών στοιχείων από τη λειτουργικότητα των ιστότοπων κοινωνικών μέσων και από τα στοιχεία που κοινοποιούνται.
- **Παρουσία:** Παρουσία σημαίνει αν τελικοί χρήστες μπορούν να γνωρίζουν ο ένας για τον άλλον εάν είναι προσβάσιμοι, αν ο χρήστης είναι συνδεδεμένος, ή δεν είναι συνδεδεμένος, απασχολημένος ή κρυμμένος. Καθώς ο εικονικός και ο πραγματικός κόσμος συνδέονται πολύ στις μέρες μας, αυτό υποδηλώνει επίσης πως οι άνθρωποι είναι στον πραγματικό κόσμο. Πού ήταν ή πού είναι ή που πηγαίνουν. Για παράδειγμα, το Facebook προσφέρει τη δυνατότητα να προσθέσει όλα τα μέρη που ο χρήστης είχε επισκεφθεί μέσα στον χρόνο.

- Σχέσεις: Η σχέση αντιπροσωπεύει σε ποιο επίπεδο οι χρήστες συνδέονται μεταξύ τους, και είναι μια σχέση που παρέχεται από το σύστημα, όπως οι φίλοι στο Facebook, ή το να ακολουθούν ο ένας τον άλλον στο Twitter. Μια σημαντική πτυχή είναι κατά πόσο η σχέση αυτή είναι πραγματική, με άλλα λόγια δύο φίλοι στο Facebook πραγματικά επικοινωνούν εκεί; Ή είναι απλώς μία από τις λίστες φίλων.
- Φήμη: Η φήμη αντιπροσωπεύει σε ποιο βαθμό οι χρήστες μπορούν να καθορίσουν την κοινωνική κατάσταση του άλλου (συμπεριλαμβανομένων των ίδιων) χρησιμοποιώντας τα κοινωνικά μέσα. Αν και η φήμη μπορεί να ερμηνευθεί διαφορετικά από τον καθένα, αλλά στις περισσότερες περιπτώσεις θεωρείται ως μέτρο εμπιστοσύνης όπως αναφερθήκαμε σε προηγούμενο κεφάλαιο.
- Ομάδες: Το στοιχείο των ομάδων υποδεικνύει σε ποιο βαθμό οι χρήστες μπορούν να δημιουργήσουν κοινότητες και υποκοινότητες. Όταν το δίκτυο είναι πιο κοινωνικό, οι χρήστες μπορούν να δημιουργήσουν όλο και μεγαλύτερες ομάδες φίλων, οπαδών.

### 3.2 Κοινωνικά δίκτυα

Παρακάτω δίνονται ορισμένα στατιστικά στοιχεία για τη χρήση των κοινωνικών δικτύων το 2021<sup>1</sup>:

- 3,96 δισεκατομμύρια άνθρωποι χρησιμοποιούν σήμερα κοινωνικά δίκτυα παγκοσμίως, σχεδόν διπλάσια από τα 2,07 δισεκατομμύρια το 2015
- Ο μέσος άνθρωπος έχει 8,6 λογαριασμούς κοινωνικών μέσων το 2020, από 4,8 το 2014
- Ο ρυθμός αύξησης των κοινωνικών μέσων από το 2015 είναι κατά μέσο όρο 12,5% από έτος σε έτος. Ωστόσο, η ανάπτυξη σημειώνει πτώση με τα στοιχεία 2019-2020 να αποκαλύπτουν ρυθμό ανάπτυξης 9,2%
- Ανά περιοχή, η ανάπτυξη των κοινωνικών μέσων το 2019-2020 ηγείται από την Ασία: + 16,98%, Αφρική + 13,92%, Νότια Αμερική + 8,00%, Βόρεια Αμερική + 6,96%, Ευρώπη + 4,32% και Αυστραλασία + 4,9%
- Το 50,64% των 7,77 δισεκατομμυρίων ανθρώπων στον κόσμο χρησιμοποιούν κοινωνικά μέσα, σε ένα κοινό ηλικίας 13+, το 63% είναι ενεργοί χρήστες
- Το 83,36% των χρηστών του Διαδικτύου είναι στα κοινωνικά μέσα. Ωστόσο, ένα 90,71% των χρηστών του διαδικτύου μέσω κινητών τηλεφώνων είναι ενεργά στα δίκτυα
- Από τα 3,96 δισεκατομμύρια χρηστών στα κοινωνικά δίκτυα κοινωνικών μέσων, το 99% έχει πρόσβαση σε ιστότοπους ή εφαρμογές μέσω κινητής συσκευής, με μόνο 1,32% να έχει πρόσβαση σε πλατφόρμες αποκλειστικά μέσω υπολογιστή
- Σε παγκόσμιο επίπεδο, ο μέσος χρόνος που αφιερώνει ένα άτομο στα μέσα κοινωνικής δικτύωσης την ημέρα είναι 2 ώρες και 24 λεπτά. εάν κάποιος εγγραφεί στα 16 και ζήσει ως τα 70, θα περάσει 5,7 χρόνια της ζωής του σε αυτό
- Το Facebook είναι το κορυφαίο κοινωνικό δίκτυο με 2,7 δισεκατομμύρια ενεργούς χρήστες κάθε μήνα, ακολουθούμενο από το YouTube (2 δισεκατομμύρια), το

<sup>1</sup> Πηγή: <https://backlinko.com/social-media-users>

WhatsApp (2 δισεκατομμύρια), το FB Messenger (1,3 δισεκατομμύρια) και το WeChat (1,2 δισεκατομμύρια)

- Το 70% του συνολικού πληθυσμού των ΗΠΑ έχει λογαριασμούς κοινωνικών δικτύων, συνολικού αριθμού 231,47 εκατομμυρίων ανθρώπων και έχει αυξηθεί κατά 3,1% μεταξύ 2019 και 2020
- Στις ΗΠΑ, οι μηνιαίοι ενεργοί χρήστες στο διάστημα 2019-2020 έχουν αυξηθεί κατά 1-2% από το Facebook, το Instagram και το LinkedIn. Αν και, το Pinterest, το Twitter, το WhatsApp και το Snapchat είναι όλα κάτω από 1-3%.
- Στις ΗΠΑ, το 76% όλων των γυναικών χρησιμοποιούν κοινωνικά μέσα, σε σύγκριση με το 72% όλων των ανδρών

### 3.3 SR Αρχιτεκτονικές

Τα SR συστήματα έχουν αποσπάσει μεγάλη προσοχή στο πεδίο της έρευνας σε μια προσπάθεια να αξιοποιήσουν τις κοινωνικές σχέσεις για να βελτιώσουν τη διαδικασία διαμόρφωσης αυτοματοποιημένων προτάσεων. Το παραπάνω βασίζεται στην υπόθεση ότι οι προτιμήσεις των χρηστών επηρεάζονται περισσότερο από αυτές που προτιμούν οι συνδεδεμένοι φίλοι τους παρά από εκείνες των άγνωστων χρηστών (Weng et al 2010). Αυτό είναι που αναλύσαμε επίσης ως κοινωνιολογική έννοια σχετικά με την κοινωνική επιρροή (Miller McPherson, and Smith-Lovin, 2001) στη παραπάνω ενότητα. Οι Tang et al (2013) αποδίδουν στον όρο social recommendation την εξής ερμηνεία «οποιαδήποτε πρόταση με επιπλέον δεδομένα που προτάσσονται από τις κοινωνικές σχέσεις, δηλαδή η ενίσχυση μίας υπάρχουσας μηχανής προτάσεων με επιπλέον κοινωνικά σήματα".

Οι διάφορες προτεινόμενες προσεγγίσεις μπορούν να κατηγοριοποιηθούν ανάλογα με τον τύπο των κοινωνικών σχέσεων (εμπιστοσύνη, φιλία κ.λπ.), τον τύπο του υποκείμενου αλγορίθμου SR (βάσει μοντέλου, βάσει μνήμης, κ.λπ.), και το επίπεδο ολοκλήρωσης των κοινωνικών πληροφοριών στη διαδικασία διαμόρφωσης μίας πρότασης. Μια κοινή προσέγγιση είναι η ενίσχυση των SR συστημάτων με βάση μοντέλα κοινωνικών συνδέσεων που συχνά εκφράζουν τον βαθμό εμπιστοσύνης. Αυτό μπορεί να γίνει αξιολογώντας από κοινού διάφορους παράγοντες:

- Ορίζεται ένα κοινό διάνυσμα προτιμήσεων μέσα στο οποίο οι χρήστες κάνουν τις επιλογές τους (Tang et al, 2013)
- Μία εισήγηση μπορεί να προέρχεται από τον γραμμικό συνδυασμό δύο ή περισσότερων συστημάτων (π.χ. (Ma et al, 2009), (Tang et al, 2012)) μέσω της εφαρμογής ensemble μεθόδων
- Μέσω της διαδικασίας κανονικοποίησης δίνεται προτεραιότητα στις αξιολογήσεις που βασίζονται σε κοινωνικά δίκτυα (π.χ. (Jamali and Ester, 2010), (Ma et al, 2011)).

Ένας εναλλακτικός τρόπος προσέγγισης περιλαμβάνει τρόπους ενίσχυσης της συνεργασίας που βασίζονται στη μνήμη. Η διαδικασία φιλτραρίσματος σχηματίζει τον «κοινωνικό»

περίγυρο του χρήστη χρησιμοποιώντας ομοιότητες που απορρέουν από τις αξιολογήσεις χρηστών ή / και τις κοινωνικές σχέσεις τους και εστιάζουν στην εμπιστοσύνη (Massa and Avesani, 2005), (Varlamis et al, 2013). Το μεγαλύτερο μέρος της έρευνας σε SR αξιοποιεί την ομοιότητα των χρηστών για τη δημιουργία προτάσεων αντί να ενσωματώσουν τις κοινωνικές πληροφορίες στον ίδιο τον αλγόριθμο δημιουργίας συστάσεων.

Πιο πρόσφατα, οι Deng et al. (2017) εισήγαγαν ένα πίνακα παραγοντοποίησης βασισμένη σε βαθιά μάθηση για την αντιμετώπιση των ζητημάτων που σχετίζονται με την εξαίρεση διανυσμάτων με λανθάνουσα χαρακτηριστικά. Χρησιμοποίησαν την εμπιστοσύνη του χρήστη προς άλλους χρήστες που ανήκουν στον στενό περίγυρο του για να ορίσουν τις ομάδες γειτνίασης. Χρησιμοποίησαν την επίδραση της κοινότητας στη δημιουργία εμπιστοσύνης μεταξύ των χρηστών και το ενσωμάτωσαν ως παράγοντα δημιουργίας εμπιστοσύνης στο παραπάνω πίνακα παραγοντοποίησης.

### 3.3.1 Σχηματισμός μοντέλου γειτνίασης βάση την εξάπλωση των επιρροών

Η διάδοση της επιρροής μπορεί να οριστεί ως ένα πρόβλημα ελαχιστοποίησης: πρέπει να προσδιοριστεί ο ελάχιστος αριθμός των πιο σημαντικών κόμβων, που μπορούν να επηρεάσουν το μέγιστο αριθμό των υπόλοιπων κόμβων (αν όχι ολόκληρο το δίκτυο). Έρευνα έχει δείξει ότι αυτές οι πληροφορίες αξιοποιούνται για τη δημιουργία SRs βάσει γραφημάτων που είναι πιο ακριβείς από τις παραδοσιακές τεχνικές που βασίζονται μόνο στην ταξινόμηση. Η τεχνική αυτή ακολουθεί τη παρακάτω προσέγγιση:

- Ορίζει τα βάρη στο κατευθυνόμενο γράφημα: το βάρος του κάθε άκρου υπολογίζεται με βάση χαρακτηριστικά των δύο κόμβων που συνδέονται στο κάθε άκρο. Τα βάρη των άκρων χρησιμοποιούνται στη συνέχεια για να ποσοτικοποιήσουν το βαθμό της επίδρασης που διαδίδεται κατά μήκος των κόμβων στο γραφικό κοινωνικό δίκτυο.
- Καθορίζει τις συνθήκες για τη διάδοση της επιρροής, για παράδειγμα κατώτατα ή ανώτατα άκρα στις τιμές παραμέτρων προκειμένου να προσδιοριστεί η κατάσταση του κατά πόσον ένας κόμβος επηρεάζεται από έναν άλλο κόμβο ή όχι.
- Ταξινομεί κόμβους με βάση συγκεκριμένες μεθοδολογίες όπως αυτές που αναφέρθηκαν στην παραπάνω ενότητα.
- Δημιουργεί το μοντέλο γειτνίασης με βάση το βαθμό διάδοσης επιρροών για τη δημιουργία προτάσεων προς τους χρήστες. Με αυτό το τρόπο δημιουργεί ένα εξατομικευμένο υποσύνολο χρηστών που βασίζεται σε κοινωνικά δίκτυα και με τη σειρά του παράγει συστάσεις με μεγαλύτερη ακρίβεια.

Με βάση τα παραπάνω, ένα κοινωνικό δίκτυο ορίζεται ως ένα κατευθυνόμενο γράφημα με βάρη  $G = (V; E; W)$ , όπου το  $V$  είναι το σύνολο όλων των κόμβων του γραφήματος (δηλαδή οι χρήστες), το  $E$  είναι το σύνολο των άκρων (δηλαδή οι κοινωνικές τους συνδέσεις), και  $W$  είναι το σύνολο των βαρών στις ακμές. Στο κατευθυνόμενο γράφημα, ένα άκρο από τον κόμβο  $u$  στον κόμβο  $v$  ( $u \rightarrow v$ ) δηλώνει ότι ο χρήστης  $u$  «ακολουθεί» τον χρήστη  $v$  στο κοινωνικό δίκτυο.

Τα βάρη των άκρων καθορίζονται στο δίκτυο από δομικά χαρακτηριστικά των δύο κόμβων που σχηματίζουν το άκρο. Για παράδειγμα, για τους γειτονικούς κόμβους  $u$  και  $v$ , το βάρος άκρου  $w(uv)$  αντιπροσωπεύει την επίδραση του  $v$  στο  $u$  για αυτήν τη σύνδεση και ορίζεται ως:

$$Eπιρροή(uv) = \text{βαθμός σπουδαιότητας}(v) / \text{βαθμός συνδέσεων}(u)$$

όπου ο βαθμός σπουδαιότητας ( $v$ ) είναι ο αριθμός των συνδέσεων που καταλήγουν στον  $v$ . Ο κόμβος  $v$  θεωρείται κόμβος που «επηρεάζει» και ο  $u$  θεωρείται κόμβος που «επηρεάζεται» εφόσον  $w(uv) \geq$  κατώτατο όριο. Εάν αυτή η συνθήκη ισχύει, τότε ο κόμβος  $v$  θεωρείται κόμβος γεννήτρια για τη δημιουργία συστάσεων και ο κόμβος  $u$  ανήκει στη πλειάδα γειτνιαζόντων κόμβων του  $v$ .

### 3.3.2 Πρόβλεψη ταξινόμησης με βάση παράγοντες κοινωνικής κανονικοποίησης

Η γενική ιδέα του Matrix Factorization (MF) είναι να μοντελοποιηθεί η αλληλεπίδραση χρήστη-στοιχείου με παράγοντες που αντιπροσωπεύουν λανθάνοντα χαρακτηριστικά του χρήστη και στοιχεία στο σύστημα. Μετά από αυτό το μοντέλο έχει εκπαιδευτεί από τα διαθέσιμα δεδομένα, για να χρησιμοποιηθεί για την πρόβλεψη των αξιολογήσεων των χρηστών σε νέα αντικείμενα. Το μοντέλο αυτό χρησιμοποιήθηκε ευρέως για τη σύνδεση του κοινωνικού δικτύου με συστήματα συστάσεων.

#### SoRec

Το μοντέλο αυτό στοχεύει στην αντιμετώπιση της αραιότητας δεδομένων και το πρόβλημα της έλλειψης δεδομένων για την πραγματοποίηση προβλέψεων με ακρίβεια με την ενσωμάτωση κοινωνικών δικτύων (SN) και συστημάτων συστάσεων (Ma et al, 2008). Το μοντέλο ορίζεται από τη παρακάτω φόρμουλα

$$S_{u,v}^* = S_{u,v} \sqrt{\frac{d_v^-}{d_u^- d_v^-}}$$

Στην οποία το  $d_u^-$  είναι ο αριθμός των χρηστών που ο χρήστης  $u$  ακολουθεί ή εμπιστεύεται.

$d_v^-$  είναι ο αριθμός των χρηστών που ακολουθούν ή εμπιστεύονται τον χρήστη  $v$ . Ο πίνακας αξιολόγησης στοιχείων τελικού χρήστη λαμβάνεται από το μοντέλο με τον τύπο:

$$\hat{R} = r_m + QP^T$$

Στην εξίσωση, οι μεταβλητές  $P$  και  $Q$  περιέχουν την κατάταξη και το offset. Η κοινωνική πτυχή προστίθενται στο μοντέλο με την ακόλουθη εξίσωση:

$$S^* = s_m + QZ^T$$

Όπου το  $Z$  περιέχει τις κοινωνικές πληροφορίες στο μοντέλο, και ο πίνακας  $Q$  διαμοιράζεται μεταξύ των δύο εξισώσεων. Αυτός είναι ο λόγος για τον οποίο αυτός ο πίνακας πρέπει να

περιέχει και τα δύο: πληροφορίες για τα στοιχεία (προϊόντα / υπηρεσίες) και πληροφορίες για το χρήστη προκειμένου να επιτευχθεί ακριβής πρόβλεψη και για τα δυο. Δεδομένου ότι ο πίνακας  $Z$  δεν απαιτείται για τη διαδικασία πρόβλεψης, μπορεί να υπολογιστεί μετά την εκμάθηση των δύο πινάκων  $P$  και  $Q$ .

Στη γενική ιδέα του SoRec είναι πιθανό δύο χρήστες  $Q_u, Q_v$  να έχουν ομοιότητες αν είναι συνδεδεμένοι μεταξύ τους. Στο αρχικό μοντέλο εξετάζεται μόνο ο παράγοντας της θετικής εμπιστοσύνης. Ωστόσο, το μοντέλο επιτρέπει επίσης αρνητικές τιμές αντιπροσωπεύοντας τη δυσπιστία μεταξύ δύο χρηστών. Μια τέτοια συνάρτηση βελτιστοποιείται χρησιμοποιώντας τη μέθοδο καθόδου κλίσης.

### Μοντέλο κοινωνικής εμπιστοσύνης

Το μοντέλο αυτό είναι ένας γραμμικός συνδυασμός της μεθόδου παραγοντοποίησης μήτρας και του μοντέλου που αντικατοπτρίζει ένα κοινωνικό δίκτυο (Ma et al 2009). Η πρόβλεψη επιτυγχάνεται με την ενσωμάτωση των δύο μητρών  $P$  και  $Q$  σε έναν τύπο:

$$\hat{R}_{u,i} = r_m + \alpha Q_u P_i^T + (1 - \alpha) \sum_{v \in F_u^-} S_{u,v} Q_v P_i^T$$

$F_u^-$  είναι το σύνολο των άμεσων φίλων του χρήστη  $u$ . Η εξίσωση περιέχει 3 στοιχεία, τα δύο πρώτα είναι τα ίδια στην παραδοσιακή μέθοδο συνεργατικού φιλτραρίσματος: χαρακτηριστικά καθολικής μετατόπισης  $r_m$  και πρόβλεψης με βάση τον χρήστη  $u$  και λανθάνοντα στοιχεία. Ο τελευταίος όρος είναι το σταθμισμένο άθροισμα των προβλεπόμενων αξιολογήσεων για το στοιχείο  $i$  από όλους τους φίλους του χρήστη  $u$ . Έτσι το μοντέλο αυτό ενσωματώνει την κοινωνική επιρροή στην διαδικασία. Η παράμετρος  $\alpha$  αντιπροσωπεύει το επίπεδο της κοινωνικής επιρροής στη διαδικασία πρόβλεψης. Η κοινωνική επιρροή δεν επηρεάζει αν το  $\alpha$  είναι ίσο με 1 και έχει την μεγαλύτερη επιρροή όταν είναι ίσο με 0.

### Μοντέλο κοινωνικής μήτρας παραγοντοποίησης

Το μοντέλο αυτό, που προτάθηκε στο [31], έχει παρουσιάσει καλύτερα αποτελέσματα από τα δύο μοντέλα SoRec και STE. Αυτό το μοντέλο αντιμετωπίζει τη μεταβατικότητα της σχέσης εμπιστοσύνης στο κοινωνικό δίκτυο. Ως εξάρτηση του διανύσματος δυνατοτήτων του χρήστη στους γείτονές του, τα διανύσματα μπορούν να μεταδοθούν στο κοινωνικό δίκτυο, το διάνυσμα του χρήστη μπορεί να διαμορφωθεί κατά πάσα πιθανότητα ανάλογα με όλους τους χρήστες του δικτύου. Κάθε σειρά της μήτρας κανονικοποιείται στο 1, οπότε για κάθε χρήστη έχουμε την ακόλουθη εξίσωση:

$$\sum_u S_{u,v}^* = 1$$

### Σύστημα προτάσεων βάσει κύκλου κοινωνικών σχέσεων

Τα περισσότερα από τα υπάρχοντα συστήματα λαμβάνουν υπόψη τους το σύνολο του κοινωνικού δικτύου προκειμένου να παράγει μια πρόταση ενώ ορισμένοι ερευνητές εξετάζουν μια άλλη πτυχή συστημάτων συστάσεων. Στο [32] ο Γιανγκ πρότεινε έναν σύστημα



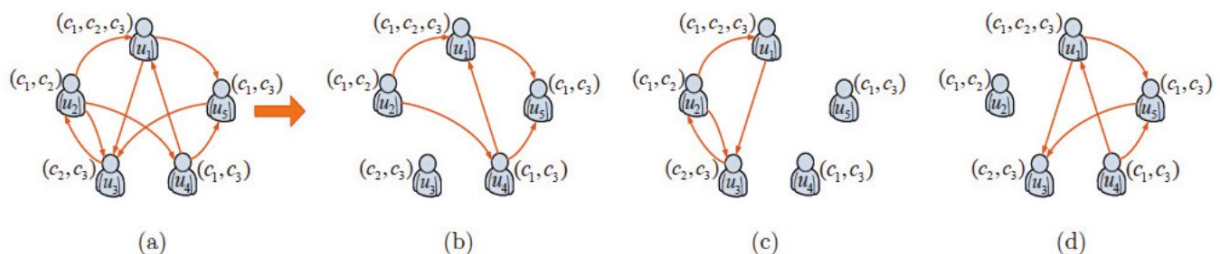
συστάσεων ανάλογα με τον κύκλο σχέσεων. Η υπόθεση είναι ότι οι άνθρωποι είναι πολύπλευροι, διαδικτυακοί ή εκτός σύνδεσης. Στην πραγματική ζωή, οι χρήστες δρουν σε διαφορετικές ομάδες φίλων σε διαφορετικούς τομείς. Για παράδειγμα, ο τομέας τεχνολογίας είναι διαφορετικός από τον τομέα ρούχων.

Υποθέτοντας ότι μπορούμε να γνωρίζουμε τους κύκλους εμπιστοσύνης των χρηστών στους διαφορετικούς τομείς, τότε μπορούμε να χρησιμοποιήσουμε τις πληροφορίες μόνο σε αυτόν τον κύκλο για να προτείνουμε στο χρήστη προτάσεις μόνο για αυτόν τον τομέα. Ωστόσο, αυτός ο ορισμός του κύκλου δεν είναι ο ίδιος με το Google + ή Facebook, όπου ο κύκλος μπορεί να περιέχει φίλους σε διαφορετικές κατηγορίες. Ο Yang et al (2012) προτείνει ένα σύνολο αλγορίθμων για συμπερίληψη συγκεκριμένων κατηγοριών κύκλων φίλων και μετά για να συμπεράνει τις τιμές εμπιστοσύνης κάθε συνδέσμου με βάση τις αξιολογήσεις των χρηστών δημιουργεί κύκλους σχέσεων με βάση τις αξιολογήσεις τους. Ακολουθεί την υπόθεση ότι ο χρήστης εμπιστεύεται έναν φίλο σε μια συγκεκριμένη κατηγορία αλλά όχι σε όλες, και η προσθήκη των πληροφοριών κάθε φίλου έχει ως αποτέλεσμα να υπάρχει θόρυβος στην δημιουργία προβλέψεων.

Για τον προσδιορισμό ενός κύκλου ανά χρήστη: σε κάθε κατηγορία, ένας χρήστης  $v$  βρίσκεται στον κύκλο του χρήστη  $u$ , εάν και μόνο εάν ισχύουν οι δύο συνθήκες:

- $S_{u,v} > 0$  στο κυρίως κοινωνικό δίκτυο, δηλαδή οι  $u$ , και  $v$  είναι φίλοι
- $n(M_u^{(c)}) > 0$  και  $n(M_v^{(c)}) > 0$  κατά τη στιγμή της αξιολόγησης

Αυτό σημαίνει ότι και οι δύο χρήστες  $u$  και  $v$  πρέπει να έχουν αξιολογήσει στοιχεία που σχετίζονται με την ίδια κατηγορία για να βρίσκονται στον ίδιο κύκλο.



Εικόνα 4. Παραδείγματα απεικόνιση κύκλων κοινωνικών σχέσεων. Κάθε χρήστης επισημαίνεται με κατηγορία στην οποία έχει υποβάλλει μια αξιολόγηση: α) Είναι το αρχικό κοινωνικό δίκτυο, και β), γ), δ) είναι οι κύκλοι που προκύπτουν σύμφωνα με τις κατηγορίες (Yang et al 2012)

Ο κύριος στόχος αυτής της έρευνας είναι να παράγει αποτελεσματικές προτάσεις, οι οποίες επιτυγχάνονται προβλέποντας ακριβείς αξιολογήσεις χρηστών και περιλαμβάνει τις παρακάτω τεχνικές:

- κοινωνική κανονικοποίηση ως ένα από τους παράγοντες στον πίνακα παραγοντοποίησης στον οποίο αναφερθήκαμε στην προηγούμενη ενότητα.
- μετρική ομοιότητας: ομοιότητες μεταξύ των χρηστών καθορίζουν τα βάρη στις ακμές του γραφήματος και συνδυάζονται με τον παράγοντα της κοινωνικής

κανονικοποίησης για να καθορίσουν τον βαθμό διάδοσης της επιρροής μέσα στο δίκτυο.

Υπάρχουν πολλοί τρόποι που προτείνονται στη βιβλιογραφία για τον εντοπισμό σημαντικών κόμβων σε ένα γράφημα. Οι περισσότερες προσεγγίσεις χρησιμοποιούν μια μετρική που βασίζεται σε γραφήματα όπως η κεντρικότητα (βαθμός, ιδιοδιανύσματα, κ.λπ.) ή PageRank για δημιουργία μίας αρχικής κατάταξης κόμβων:

- PageRank (PR): Ο Page et al. (1999) εισήγαγε το PageRank για την κατάταξη ιστοσελίδων στο διαδίκτυο που ωστόσο χρησιμοποιείται και στα κοινωνικά δίκτυα για την κατάταξη κόμβων.
- Βαθμός εξωστρέφειας (Outdegree): χρησιμοποιείται για τη μέτρηση της επιρροής στο κοινωνικό γράφημα.
- Upper-bound PageRank (UB-PR): Ο Liu et al. (2017) πρότεινε αυτή τη τεχνική ως ένα από τα αρχικά κριτήρια κατάταξης κόμβων στο γράφημα.

Ο παρακάτω αλγόριθμος, που ονομάζεται Threshold-Bounded Influence Propagation in Digraph (TB-IP), παίρνει ως είσοδο ένα γράφημα  $G$  και εξάγει τα σύνολα  $M$  και  $D$ . Ο αλγόριθμος δέχεται τις ακόλουθες παραμέτρους:

- a. ένα κατώτατο άκρο  $thr$
- b. το μέγιστο επιτρεπόμενο βάθος διάδοσης επιρροής
- c. ένα παράγοντα επιβράδυνσης της διάδοσης της επιρροής και
- d. μια στρατηγική κατάταξης για την αρχικοποίηση των κόμβων

Ο αλγόριθμος ξεκινά ταξινομώντας όλους τους κόμβους σε φθίνουσα σειρά με βάση τους εκχωρημένους σε κατάταξη κόμβους  $r(v)$ . Στη συνέχεια, ξεκινώντας με τον υψηλότερο κόμβο, εξετάζει αν οι άμεσες συνδέσεις του κόμβου θα πρέπει να προστεθούν στη γειτονιά του και ως εκ τούτου να θεωρηθούν επηρεασμένες από αυτόν ή όχι. Αυτό καθορίζεται εξετάζοντας εάν το βάρος της εκάστοτε ακμής  $w(uv)$  είναι πάνω από το καθορισμένο άκρο ή όχι. Εάν τουλάχιστον ένας συνδεδεμένος κόμβος πληροί την συνθήκη, ο κόμβος  $v$  προστίθεται στο σεντ  $M$  "κόμβοι που επηρεάζουν" και οι άλλοι κόμβοι προστίθενται στο σύνολο  $D$  "κόμβοι που επηρεάζονται".

Εάν ο αλγόριθμος έχει ρυθμιστεί για να εξετάσει κόμβους που συνδέονται έμμεσα με τον κόμβο  $v$  (το βάθος καθορίζεται από την παράμετρο  $maxhor$ ), καθένας από τους κόμβους που προστέθηκαν στο σύνολο  $NG(v)$  στο προηγούμενο βήμα χρησιμοποιούνται για την εύρεση των απευθείας συνδεδεμένων κόμβων τους. Ωστόσο, σε αυτήν την περίπτωση, τα αντίστοιχα βάρη των ακμών ενημερώνονται από τον παράγοντα hopping  $f$ , πριν αξιολογηθεί έναντι του κάτω άκρου. Όταν ένας κόμβος ικανοποιεί τη συνθήκη και δεν έχει προστεθεί προηγουμένως στο σύνολο  $D$ , τότε προστίθεται σε αυτό το σύνολο, όσο και στη γειτονιά του  $v$ ,  $NG(v)$ .

Ο αλγόριθμος σταματά αυτόν τον βρόχο όταν έχει επιτευχθεί είτε το μέγιστο βάθος είτε δεν υπάρχουν άλλοι κόμβοι που να επηρεάζονται στο τρέχον επίπεδο. Αυτή η διαδικασία επαναλαμβάνεται για καθένα από αυτούς τους κόμβους, όπως έχουν επιλεγεί από τη λίστα κατάταξης, και εφόσον δεν έχουν ήδη προστεθεί στο σύνολο  $D$ .

```

Require: A weighted and directed social network  $G = (V, E, W)$ 
Ensure: Influenced Vertices  $D \subset V$  and Influential Vertices  $M \subset V$ 
1: Initialize:  $thr, maxhop, visit = 0, hop = 0, decay = 0.1, D = \emptyset, M = \emptyset$ 
2:  $\forall v \in V, r(v) = \text{compute\_rank}(v)$ 
3: Sort  $\forall v \in r(v)$  in non-increasing order
4: for each  $v \in S$  following order do
5:   if  $w(u_v) > thr$  and  $\exists u \in V \setminus D$  then
6:      $N_G(v) = N_G(v) \cup u$ 
7:      $D = D \cup u$ 
8:      $M = M \cup v$ 
9:      $visit = 1$ 
10:  while  $hop \leq maxhop$  and  $visit = 1$  do
11:     $hop = hop + 1$ 
12:     $visit = 0$ 
13:    for each  $u \in N_G(v)$  do
14:       $w(p_u) = w(p_u) * (1 + hop * decay)$ 
15:      if  $w(p_u) > thr$  and  $\exists p \in V \setminus D$  then
16:         $N_G(v) = N_G(v) \cup p$ 
17:         $D = D \cup p$ 
18:         $visit = 1$ 
19:  output  $D, M$ 

```

Εικόνα 5. Threshold-Bounded Influence Propagation in Digraph (TB-IP) (Gulati et al, 2018)

### 3.3.3 Η Προσέγγιση Implicit Social Trust and Sentiment (ISTS)

Το ISTS είναι ένα σύστημα προτάσεων έμμεσης κοινωνικής εμπιστοσύνης και συναισθημάτων που μπορεί να χρησιμοποιήσει τα μέσα κοινωνικής δικτύωσης ενός χρήστη για να αντλήσει τις προτιμήσεις του χρήστη, ακόμη και στην περίπτωση που δεν υπάρχει διαθέσιμο ιστορικό πληροφοριών αξιολόγησης.

Τα γενικά χαρακτηριστικά αυτής της προτεινόμενης προσέγγισης είναι τα ακόλουθα:

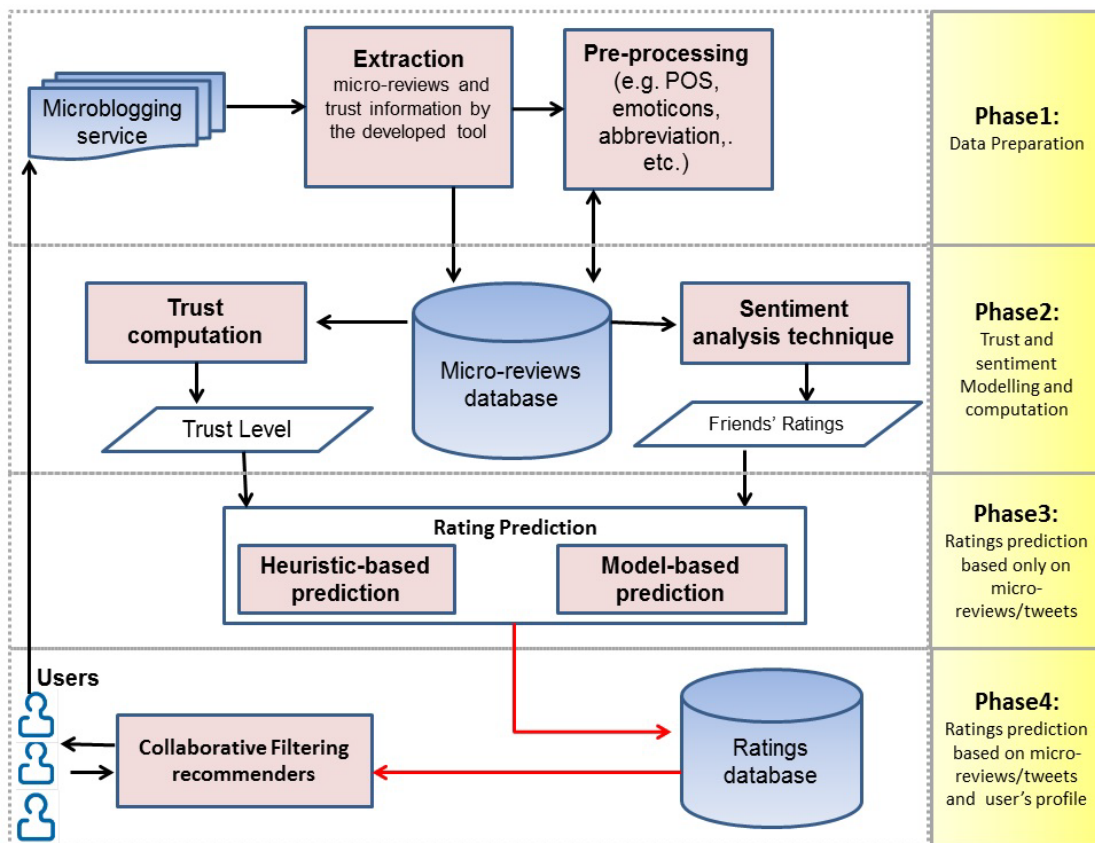
- Προχωρά ένα βήμα πιο πέρα από τον εμπορικό ιστότοπο και βασίζεται στη σύνδεση χρηστών με φίλους για να τους συμβουλευτούν σχετικά με πληροφορίες και εμπειρία προκειμένου να αποκτήσουν περισσότερες γνώσεις που μπορούν να βελτιώσουν τη διαδικασία προτάσεων. Αυτό βασίζεται στο ότι οι άνθρωποι είναι πρόθυμοι να λάβουν προτάσεις από τους γνωστούς τους και μπορούν να επηρεαστούν από τις απόψεις τους, ακόμη και αν έχουν διαφορετικά συμφέροντα.
- Θεωρεί το ποσοστό των ενδοεπικοινωνιών μεταξύ τους ως δείκτη εμπιστοσύνης. Η δύναμη μιας σχέσης θα δείξει τη δύναμη της εμπιστοσύνης μεταξύ των συνομηλίκων, δηλαδή οι συχνές επικοινωνίες μεταξύ φίλων θα έχουν υψηλές τιμές εμπιστοσύνης

μεταξύ τους. Όσο μεγαλύτερη είναι η αξία της εμπιστοσύνης, τόσο περισσότερη δύναμη θα δοθεί στο υποψήφιο στοιχείο στη διαδικασία της σύστασης.

- Η πρόταση που παρέχει προσαρμόζεται με βάση τις απόψεις των φίλων. Προσπαθεί να μετατρέψει τις απόψεις των φίλων σε μια κλίμακα βαθμών, για παράδειγμα από 1 έως 5, με τις υψηλότερες βαθμολογίες να δείχνουν υψηλότερη θετική γνώμη προκειμένου να υιοθετήσει τις αξιολογήσεις των απόψεων στη διαδικασία των προτάσεων.
- Ο εντοπισμός των απόψεων των φίλων για αντικείμενα και η ανάθεση του επιπέδου εμπιστοσύνης μεταξύ των φίλων διαμορφώνουν τα τελικά εξατομικευμένα προτεινόμενα στοιχεία για έναν ενεργό χρήστη.
- Διαχειρίζεται τη λίστα φίλων ενός χρήστη ως τον προσωπικό του ιστότοπο εμπιστοσύνης (WOT, Web Of Trust). Ο χρήστης δεν θα κληθεί να αξιολογήσει τους ανθρώπους για να προετοιμάσει το WOT.
- Η μέτρηση εμπιστοσύνης και η ανάλυση γνώμης διερευνώνται στο πλαίσιο της υποκειμενικότητας και αξιολογούνται χρησιμοποιώντας διαφορετικούς τομείς προτάσεων.

Είναι σημαντικό να αναφέρουμε ότι η προσέγγιση αυτή είναι αρκετά γενική και μπορεί να εφαρμοστεί σε πολλές διαδικτυακές επιχειρήσεις μαζί με διαφορετικά κοινωνικά μέσα. Αυτά τα δίκτυα σχεδιάζουν διαφορετικούς τρόπους επικοινωνιακών συμπεριφορών, προκειμένου να επιτρέπουν στους χρήστες να δημοσιεύουν μηνύματα, να στέλνουν ξανά αναρτήσεις που τους αρέσουν ή ακόμη και να αποθηκεύουν συγκεκριμένες αναρτήσεις σε μια λίστα αγαπημένων. Φυσικά, υπάρχουν περισσότεροι τύποι αλληλεπίδρασης μεταξύ φίλων, όπως τα ιδιωτικά μηνύματα μεταξύ τους. Για παράδειγμα, στην πλατφόρμα του Twitter, η εκ νέου αποστολή είναι οι ενέργειες του τιτιβίσματος εκ' νέου και της εύνοιας των τιτιβισμάτων, ενώ στο Facebook αυτές οι δραστηριότητες μπορεί να είναι η πράξη του να αρέσουν τα σχόλια άλλων και η ενέργεια της ανταλλαγής φωτογραφιών με μια λίστα φίλων.

Το συνολικό πλαίσιο που προτείνεται απεικονίζεται στην Εικόνα 7, η οποία δείχνει τα βήματα που εμπλέκονται στο μοντέλο. Πρώτον, οι πληροφορίες εξάγονται από την υπηρεσία του μικρό-ιστολογίου. Δεύτερον, μετά τη συλλογή των απαιτούμενων δεδομένων, η εμπιστοσύνη υπολογίζεται από τη συμπεριφορά μεταξύ των χρηστών. Τρίτον, μια τεχνική ανάλυσης συναισθημάτων χρησιμοποιείται για να συμπεράνει τη γνώμη των φίλων με τη μορφή κλιμακούμενης αξιολόγησης για αντικείμενα.



Εικόνα 6. ISTS: Το προτεινόμενο πλαίσιο (Dimah et al, 2015)

Η παραπάνω προσέγγιση εφαρμόστηκε στο Twitter έχοντας υπόψη τα ακόλουθα χαρακτηριστικά του:

- Οι χρήστες μπορούν να διαχειριστούν τις σχέσεις με τη μορφή λιστών που ονομάζονται ακόλουθοι. Αυτές οι λίστες εμφανίζουν αναρτήσεις από άλλους που αντικατοπτρίζουν τις απόψεις και τις ιδέες τους.
- Οι άνθρωποι μπορούν να δημοσιεύσουν τις απόψεις τους για διάφορα θέματα, όπως πολιτικά θέματα, βιβλία, ταινίες, αγορές και διασημότητες κ.λπ. σε σύντομα τιτιβίσματα.
- Το Twitter διαθέτει ένα δημόσιο API που υποστηρίζεται από πολλές διαφορετικές γλώσσες προγραμματισμού. Η διαθεσιμότητα αυτού του API δίνει στους προγραμματιστές τη δυνατότητα να αναπτύξουν λογισμικό για πρόσβαση στον διακομιστή Twitter και να χρησιμοποιήσουν τις παρεχόμενες πληροφορίες.
- Η ευρεία εξάπλωση προηγμένων τεχνολογικών συστημάτων όπως τηλέφωνα και άλλες έξυπνες συσκευές επιτρέπει άμεση πρόσβαση στο Twitter. Αυτό εμπλουτίζει την υπηρεσία καθώς διαθέτει ένα ευρύ φάσμα τρεχουσών (online) πληροφοριών που διατίθενται από τους χρήστες της.

Στην προσέγγιση αυτή απαιτείται να μετρηθεί ο βαθμός της σιωπηρής εμπιστοσύνης μεταξύ φίλων. Στην πραγματικότητα αυτός ο τύπος κρυφής εμπιστοσύνης είναι μια υποκειμενική πτυχή που διαφέρει από τον ένα χρήστη στον άλλο. Ωστόσο, η συλλογή αυτού του τύπου

κοινωνικών πληροφοριών δεν είναι εύκολη υπόθεση. Επομένως, οι ερευνητές αναπτύξανε ένα εργαλείο για την αυτόματη συλλογή δεδομένων του κοινωνικού δικτύου χρησιμοποιώντας το Twitter API. Αυτό το εργαλείο εξάγει τις απαιτούμενες αλληλεπιδράσεις μεταξύ φίλων και ονομάζεται Twitter Interaction Extractor (TIE). Οι δυσκολίες που συναντώνται κατά τη συλλογή δεδομένων από μικρό-ιστολόγια είναι:

- Κανόνες: υπάρχουν διαφορετικοί νόμοι και κανονισμοί που εφαρμόζουν πολλοί πάροχοι μικρό-ιστολογίων για την προστασία των χρηστών τους.
- Δυναμικές αλλαγές: τα δεδομένα από τον ιστό σε πραγματικό χρόνο αλλάζουν δυναμικά όσον αφορά τις σχέσεις και το περιεχόμενο και αυτό είναι μια μεγάλη πρόκληση.
- Όριο τιμής: πολλοί πάροχοι μικρό-ιστολογίων εφαρμόζουν ένα όριο τιμών για να περιορίσουν την πρόσβαση στους προγραμματιστές και στα αιτήματά τους σε ένα σύντομο χρονικό διάστημα.
- Απόρρητο: ορισμένοι χρήστες προσθέτουν περισσότερους περιορισμούς απορρήτου για την προστασία των λογαριασμών τους και απαγορεύουν την πρόσβαση του κοινού.

Οι ερευνητές επισημάνανε ότι είχαν πρόσβαση μόνο σε λογαριασμούς που είναι διαθέσιμοι δημόσια και δεν είχαν πρόσβαση σε προστατευμένους λογαριασμούς για λόγους απορρήτου.

Σε μία επόμενη έρευνα αναλύθηκε η πτυχή της κοινωνικής εμπιστοσύνης καθώς σε πολλές περιπτώσεις η όποια μετρική δεν είναι ρεαλιστική και δεν ισχύει σε πολλές περιπτώσεις RS συστημάτων. Δηλαδή, οι διαδικτυακές σχέσεις ατόμων όπως εκπροσωπούνται στις RS δεν αποτελούν ακριβή αντανάκλαση των πραγματικών κοινωνικών φιλιών.

Αξίζει να σημειωθεί ότι η πτυχή της εμπιστοσύνης στην πλειονότητα της αναθεωρημένης βιβλιογραφίας βασίζεται είτε: στην ανάλυση κοινών αξιολογήσεων μεταξύ χρηστών προς συγκεκριμένα αντικείμενα, είτε στην άμεση αξιολόγηση της εμπιστοσύνης που παράγεται από τους χρήστες μεταξύ τους. Ωστόσο, αυτές οι υποθέσεις μπορεί να είναι σωστές για ορισμένα αντικείμενα, αλλά δεν ισχύουν για άλλα αντικείμενα.

Για παράδειγμα, στην περίπτωση δύο χρηστών  $a$  και  $b$ , εάν ο χρήστης  $a$  παρέχει βαθμολογία εμπιστοσύνης προς τον χρήστη  $b$  λόγω κοινών ενδιαφερόντων σε μια ταινία θρίλερ, αυτή η βαθμολογία δεν εγγυάται ότι οι επιλογές ταινιών δράσης του χρήστη  $b$  μπορούν επίσης να ενδιαφέρουν και τον χρήστη  $a$ . Τα εμπιστευτικά RS παρέχουν στους χρήστες νέες προτάσεις από τις προτιμήσεις του χρήστη  $b$ , αλλά η απόφαση και η βαθμολογία σχετικά με τις νέες προτάσεις μπορεί να είναι εντελώς διαφορετική.

Προκειμένου να παρουσιαστεί μια νέα προοπτική για μια πιο αποτελεσματική μέτρηση εμπιστοσύνης, η παραπάνω αρχιτεκτονική και προσέγγιση του ISTS δημιούργησε μια νέα μετρική που βασίζεται στον τρόπο με τον οποίο ένας ενεργός χρήστης μπορεί να αντιληφθεί τους ανθρώπους στα διαδικτυακά κοινωνικά δίκτυά του. Ο λόγος πίσω από αυτό είναι ότι οι άνθρωποι τείνουν να ρωτούν τους συγγενείς και τους φίλους τους για νέες εμπειρίες, ακόμη και αν έχουν διαφορετικά γούστα. Επομένως, οι άνθρωποι επηρεάζονται συνήθως από τις κοινωνικές τους σχέσεις παρά την ύπαρξη διαφορετικότητας στις προτιμήσεις. Χρησιμοποιώντας αυτήν την προοπτική εμπιστοσύνης, οι νέοι χρήστες RS μπορούν να αποκτήσουν πολλά οφέλη ως εξής. Τα αξιόπιστα άτομα βάσει της προτεινόμενης μέτρησης

εμπιστοσύνης μπορεί να ενεργοποιήσουν τις επιλογές των νέων χρηστών και να προσφέρουν ενδιαφέροντα αντικείμενα με καλύτερο τρόπο από ό, τι τα ανώνυμα άτομα στο σύστημα, κάτι που κάνουν συνήθως οι RS που βασίζονται στην εμπιστοσύνη. Αυτή η μέτρηση εμπιστοσύνης μπορεί να επιλύσει σημαντικά προβλήματα σε RS, όπως η υπερβολική εξειδίκευση που προέρχεται από κοινές αξιολογήσεις. Επίσης, μπορεί να προτείνει νέες μεθόδους στο μέλλον για την αντιμετώπιση των προκλήσεων ποικιλομορφίας στις RS. Ως εκ τούτου, η προσέγγιση μετράει την εμπιστοσύνη μεταξύ των ανθρώπων με έναν νέο τρόπο, διαφορετικό από τα RS που βασίζονται στην εμπιστοσύνη.

Όπως έχουμε αναφέρει προηγουμένως, οι ενδοεπικοινωνίες μεταξύ φίλων σε κοινωνικά δίκτυα μπορούν να χρησιμοποιηθούν για να εκφράσουν το ποσοστό εμπιστοσύνης που έχουν για τους συνομηλίκους τους λαμβάνοντας υπόψιν, το εκ νέου αποστολή τιτιβισμάτων, την αναφορά άλλων χρηστών, την προτίμηση αναρτήσεων άλλων και τον αριθμό των ακόλουθων χρηστών. Παραδείγματα συμπεριφορών ενδοεπικοινωνίας που μελετήθηκαν στο ISTS με βάση την πλατφόρμα Twitter περιλαμβάνουν:

- Επανάληψη του τιτιβίσματος, που σημαίνει ότι ένας χρήστης στέλνει ξανά ένα τιτιβισμό σε όλους τους φίλους του για να δείξει ενδιαφέρον και αυτό δηλώνεται ως RT.
- Υπάρχουν δύο τύποι λιστών φίλων στο Twitter: άτομα που ακολουθούν έναν χρήστη που ονομάζονται ακόλουθοι και άτομα που ακολουθούνται από έναν η περισσότερους χρήστες που ονομάζονται ακολουθίες. Οι άνθρωποι που έχουν μεγάλο αριθμό ακόλουθων σίγουρα ανησυχούν για την κοινή χρήση των απόψεών τους και την ποιότητα στα δημοσιευμένα σχόλιά τους. Ως αποτέλεσμα, η ύπαρξη ενός φίλου με μεγαλύτερο αριθμό ακόλουθων σε σύγκριση με τις ακολουθίες του, μπορεί να οδηγήσει σε μεγαλύτερη εμπιστοσύνη των tweets του φίλου.

Σε μια μελέτη οι συγγραφείς έδειξαν ότι οι χρήστες τείνουν να έχουν υψηλό επίπεδο ικανοποίησης όταν λαμβάνουν προτάσεις από φίλους, σε σύγκριση με εκείνες που προέρχονται από ξένους της γειτονιάς τους σε παραδοσιακά RS. Άλλα έργα προσδιορίζουν το προφίλ των χρηστών μέσω της αναγνώρισης της επιρροής. Για παράδειγμα, ερευνητές δημιούργησαν τρεις διαφορετικές μετρικές επιρροής στο Twitter: τον αριθμό των ακόλουθων, την βαθμολογία της σελίδας και τον αριθμό των επαναληπτικών τιτιβισμάτων. Διαπίστωσαν ότι διαφορετικά επίπεδα επιρροής μπορούν να καθοριστούν ανάλογα με τη μέτρηση που χρησιμοποιείται.

Σε μια άλλη μελέτη χρησιμοποιήθηκαν διαφορετικές μετρήσεις για τον καθορισμό της επιρροής: αριθμός ακόλουθων, αριθμός επαναληπτικών τιτιβισμάτων και αριθμός αναφορών. Ωστόσο, και πάλι αυτές οι μετρήσεις έδειξαν διαφορετική προοπτική επιρροής. Σε γενικές γραμμές, παρατηρείται ότι το Twitter διαθέτει ακόλουθους / ακολουθίες και οι πληροφορίες των εκ νέου απεσταλμένων τιτιβισμάτων χρησιμοποιούνται ως μετρήσεις δημοφιλούς επιρροής.

### Εξαγωγή Συναισθήματος

Ένα άλλο σημαντικό συστατικό του ISTS, είναι η διαχείριση συναισθημάτων πληροφορία που εξάγεται από τις σύντομες δημοσιεύσεις των φίλων, όπως φαίνεται στην Εικόνα 7. Στη

διαδικασία της επεξεργασίας κειμένου, το κομμάτι του κειμένου μετατρέπεται σε διανύσματα χαρακτηριστικών και αυτό το βήμα προορίζεται να καταστήσει διαθέσιμες για χρήση σημαντικές δυνατότητες.

### Πρόβλεψη Βαθμολογιών

Το επόμενο κρίσιμο βήμα στη διαδικασία είναι η δημιουργία προβλέψεων για τη βαθμολογία σε άγνωστα αντικείμενα. Η προσέγγιση αυτή προβλέπει βαθμολογίες για δύο τύπους χρηστών.

- Για νέους χρήστες χωρίς ιστορικό αξιολογήσεων: οι νέοι χρήστες θα λαμβάνουν προτάσεις βασισμένες κυρίως στις συνδέσεις τους σε κοινωνικά δίκτυα. Συνιστάται η χρησιμοποίηση από τους χρήστες εξατομικευμένων στοιχείων που προέρχονται από τις απόψεις των αξιόπιστων φίλων τους. Για αυτό το έργο αναπτύσσονται μέθοδοι εύρεσης και μοντέλα μηχανικής μάθησης. Επίσης, πραγματοποιούνται πειράματα για την αξιολόγηση της απόδοσής τους ως προς την ακρίβεια των συστάσεων.
- Για χρήστες που έχουν ήδη ιστορικό προφίλ αξιολόγησης: σε αυτήν την περίπτωση οι χρήστες έχουν γνωστές απόψεις και προτιμήσεις έναντι ορισμένων στοιχείων. Κατά συνέπεια, οι αξιολογήσεις στοιχείων χρηστών που είναι αποθηκευμένες στο σύστημα δεν πρέπει να αγνοούνται και θα πρέπει να χρησιμοποιούνται. Οι χρήστες θα λάβουν προτάσεις βάσει ενός συνδυασμού των γνωστών προτιμήσεών τους και των προτιμήσεων των αξιόπιστων φίλων τους. Με άλλα λόγια, πρέπει να συγχωνεύσουν τις προτιμήσεις των χρηστών και τις απόψεις των φίλων τους χρησιμοποιώντας συνεργατικές τεχνικές φιλτραρίσματος, συμπεριλαμβανομένων μεθόδων που βασίζονται σε χρήστες και βάσει αντικειμένων.

Σε ερευνητικό επίπεδο η διαδικασία της πρόβλεψη μίας βαθμολογίας προϊόντων εστιάζει στις παρακάτω προκλήσεις:

- Για την πρόβλεψη αξιολογήσεων για στοιχεία που λείπουν καλούνται να χρησιμοποιούν τα κοινωνικά δίκτυα ως νέα πηγή δεδομένων.
- Για τη βελτίωση της απόδοσης μίας πρόβλεψης θα πρέπει να χρησιμοποιηθούν διαφορετικοί αλγόριθμοι μηχανικής μάθησης.
- Η παροχή προς τους νέους χρήστες, χωρίς προφίλ προτιμήσεων, στοιχείων προβλέψεων με βάση τους φίλους τους στα κοινωνικά δίκτυα.
- Η ενσωμάτωση των κοινωνικών δικτύων στο προφίλ αξιολόγησης χρηστών και στη μέθοδο των συνεργατικών φίλτρων.



### 3.3.4 Η προσέγγιση H-ISTS

Μια βελτιωμένη έκδοση της μεθόδου ISTS για την υποστήριξη της βελτιστοποίησης των χαρακτηριστικών εμπιστοσύνης χρησιμοποιεί έναν γενετικό αλγόριθμο. Αυτή η προσέγγιση ονομάζεται Hybrid Implicit Social Trust and Sentiment (H-ISTS) που βασίζεται σε ένα πλαίσιο συνιστωσών και καλείται να ξεπεράσει τους περιορισμούς των ISTS. Σκοπός της μεθόδου είναι να διερευνήσει το αποτέλεσμα κάθε δυνατότητας εμπιστοσύνης κατά τη μέτρηση συνολικής αξιοπιστίας, βελτιστοποιώντας τις όποιες σχετικές παραμέτρους. Συνεχώς, οι νέες μετρήσεις αξιοπιστίας θα πρέπει να χρησιμοποιούνται για τη δημιουργία εξατομικευμένων προτάσεων.

Το H-ISTS βελτιώνει το ISTS με την προσθήκη των ακόλουθων δυνατοτήτων:

1. Βελτιστοποίηση παραμέτρων αξιοπιστίας: Το H-ISTS προσαρμόζει μια μέθοδο βελτιστοποίησης της μέτρησης εμπιστοσύνης. Ένας γενετικός αλγόριθμος (GA) χρησιμοποιείται για να συντονίσει τις παραμέτρους εμπιστοσύνης και να δημιουργήσει τα σωστά βάρη που περιγράφουν το σωστό εφέ ισχύος κάθε δυνατότητας εμπιστοσύνης.
2. Βελτίωση των χαρακτηριστικών εμπιστοσύνης: Αυτό περιλαμβάνει πρόσθετα διαδικτυακά χαρακτηριστικά κοινωνικής εμπιστοσύνης που ενισχύουν την οικοδόμηση εμπιστοσύνης. Κατά συνέπεια, χρησιμοποιούνται περισσότερες συμπεριφορές και δραστηριότητες μεταξύ των χρηστών και των φίλων τους σε ένα OSN για να δείξουν τον αντίκτυπό τους στο βαθμό της εμπιστοσύνης.
3. Συλλογή νέου συνόλου δεδομένων: Γίνεται ανίχνευση στο twitter για την εύρεση του κατάλληλου συνόλου δεδομένων που να τροφοδοτεί το παραπάνω.

Έχουν γίνει πολλές προσπάθειες για την εφαρμογή Γενετικών Αλγορίθμων (ΓΑ) σε RS με διαφορετικούς τρόπους: ομαδοποίηση, υβριδικά μοντέλα χρηστών και βελτιστοποίηση λειτουργιών ομοιότητας. Στην τεχνική ομαδοποίησης, για παράδειγμα στο μοντέλο K-μέσων που βασίζεται σε GA, οι χρήστες συγκεντρώνονται σε ομάδες που έχουν παρόμοιες προτιμήσεις, στη συνέχεια εφαρμόζεται η τεχνική συνεργατικού φιλτραρίσματος σε ομάδες και τα αποτελέσματα λαμβάνονται με το μόνο προνόμιο της μείωσης του χρόνου υπολογισμού.

Τα υβριδικά μοντέλα χρηστών βασίζονται στο συνδυασμό συνεργατικών τεχνικών φιλτραρίσματος και βασισμένων στο περιεχόμενο για να αξιοποιήσουν τα πλεονεκτήματα κάθε μεθόδου. Σε αυτές τις περιπτώσεις ο ΓΑ μπορεί να σχεδιαστεί για να διατηρεί πληροφορίες, όπως δημογραφικά χαρακτηριστικά εφαρμόζοντας δημογραφικά χαρακτηριστικά στη χρωμοσωμική δομή. Επιπλέον, ο ΓΑ χρησιμοποιείται για τη λήψη της λειτουργίας βέλτιστης ομοιότητας, αντί για τη μέτρηση ομοιότητας συντελεστή συσχέτισης που βασίζεται μόνο σε πληροφορίες βαθμολογίας.

Σε μια πρόσφατη έρευνα διευκρινίστηκε ότι τα συστήματα σύστασης που βασίζονται σε δεδομένα περιβάλλοντος μέσω του διαδικτύου που χρησιμοποιούν αλγόριθμους υπολογιστικής νοημοσύνης, όπως γενετικοί αλγόριθμοι, μπορούν να δείξουν αποτελεσματική απόδοση και να μετριάσουν ουσιαστικά τις προκλήσεις της ψυχρής εκκίνησης και της αδράνειας. Οι τρεις κύριοι τελεστές που χρησιμοποιούνται σε γενετικούς αλγόριθμους είναι η

επιλογή, η διασταύρωση και η μετάλλαξη για να βρουν τους απογόνους του υπάρχοντος πληθυσμού.

Στην προσέγγιση του H-ISTS ένας ΓΑ χρησιμοποιείται για τη βελτιστοποίηση των εξαγόμενων χαρακτηριστικών κοινωνικής εμπιστοσύνης ως εξής: χρησιμοποιείται μια μέθοδος επιλογής για την εκχώρηση των υποψηφίων χρωμοσωμάτων για την επόμενη γενιά. Τα επιζώντα χρωμοσώματα στην επόμενη γενιά περνούν από το στάδιο της διασταύρωσης. Τα επιλεγμένα χρωμοσώματα εκχωρούνται σε τυχαίες θέσεις ως πρόσφατα διασταυρωμένα χρωμοσώματα για τη δημιουργία του νέου πληθυσμού. Ο τελεστής μετάλλαξης ξεκινά μετά το βήμα crossover για να αποφασίσει το ενδεχόμενο αν ένα χρωμόσωμα μπορεί να μεταλλαχθεί στην επόμενη γενιά. Επίσης χρησιμοποιείται ομοιόμορφη μετάλλαξη λόγω της ιδιότητας της σταθερής πιθανότητας μεταβολής της γονιδιακής δομής σε γενιές και της επαναλαμβανόμενης ιδιότητας της αναστρέψιμης επιστροφής όταν συμβεί ενδεχόμενη σύγκλιση.

Ένας ΓΑ μπορεί να εφαρμοστεί αποτελεσματικά στη νέα προσέγγιση, ως μέθοδος βελτιστοποίησης για τρεις λόγους:

1. κάνει αξιολόγηση του βαθμού σφάλματος σε βαθμολογίες που προέκυψαν.
2. μπορεί να βοηθήσει στην εξεύρεση της καλύτερης δυνατής λύσης για τη βελτίωση της ακρίβειας των RS και στην εξεύρεση εξατομικευμένων προτάσεων που πλησιάζουν τις προτιμήσεις του χρήστη.
3. Η εμπιστοσύνη που είναι ο στόχος των συστημάτων για βελτιστοποίηση, μπορεί να χωριστεί σε δύο ανεξάρτητα στοιχεία RT και L για να αντιπροσωπεύουν γονίδια στον ΓΑ. Ο ΓΑ χρησιμοποιείται για τον προσδιορισμό των συνεισφορών των RT και L στο στοιχείο εμπιστοσύνης για την επίτευξη της πιο ακριβούς πρόβλεψης της βαθμολογίας.

### Βελτιστοποίηση εμπιστοσύνης

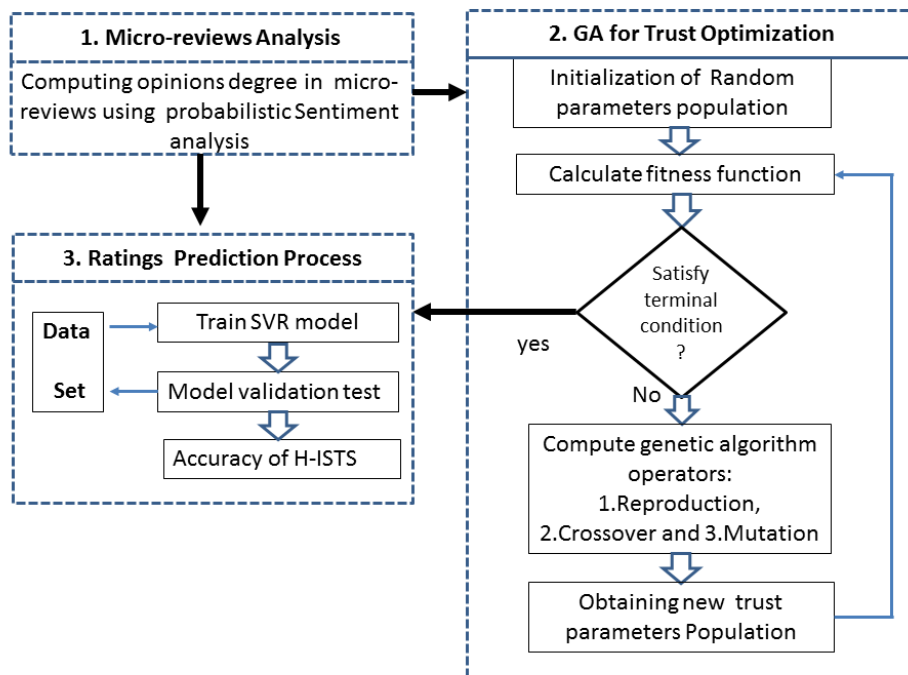
Ο βασικός στόχος της μεθόδου H-ISTS είναι ο συντονισμός της μέτρησης εμπιστοσύνης, προκειμένου να βελτιωθεί το επίπεδο ακρίβειας της πρόβλεψης. Απαντάει στο ερώτημα ποια χαρακτηριστικά εμπιστοσύνης μπορούν να παράγουν τη μεγαλύτερη επίδραση στην τελική εκτίμηση αξιολόγησης που να προσεγγίζουν τη γνώμη του νέου χρήστη.

Αρχικά, υποθέτει ότι οι νέοι χρήστες στο πλαίσιο H-ISTS έχουν παράσχει στο λογαριασμό τους τις λεπτομέρειες των λογαριασμών τους στο Twitter. Στη συνέχεια το H-ISTS ξεκινά αναλύοντας τους λογαριασμούς που ακολουθούν. Από την Εικόνα 8 φαίνεται ότι το πλαίσιο H-ISTS υιοθετεί τρία βασικά συστατικά:

- Πρώτον, μια τεχνική ανάλυσης πιθανολογικού συναισθήματος που εφαρμόζεται για την εξαγωγή του βαθμού ισχύος των απόψεων από μικρό-κριτικές.
- Δεύτερον, βελτιστοποίηση λειτουργιών κοινωνικής εμπιστοσύνης, όπως παρακολούθηση / ακόλουθοι και ενέργειες εκ νέου τιτιβίσματος χρησιμοποιώντας γενετικό αλγόριθμο.

- Τρίτον, χρησιμοποιώντας τόσο τις βελτιστοποιημένες παραμέτρους εμπιστοσύνης όσο και το βαθμό γνώμης των φίλων για να προβλέψει βαθμολογίες για νέους χρήστες που χρησιμοποιούν το μοντέλο παλινδρόμησης φορέα υποστήριξης (SVR).

Η συνολική διαδικασία της προτεινόμενης προσέγγισης απεικονίζεται στην Εικόνα 8. Η διαδικασία βελτιστοποίησης επαναλαμβάνεται και η λειτουργία φυσικής κατάστασης ενημερώνεται με νέες παραμέτρους εμπιστοσύνης σε κάθε επανάληψη.



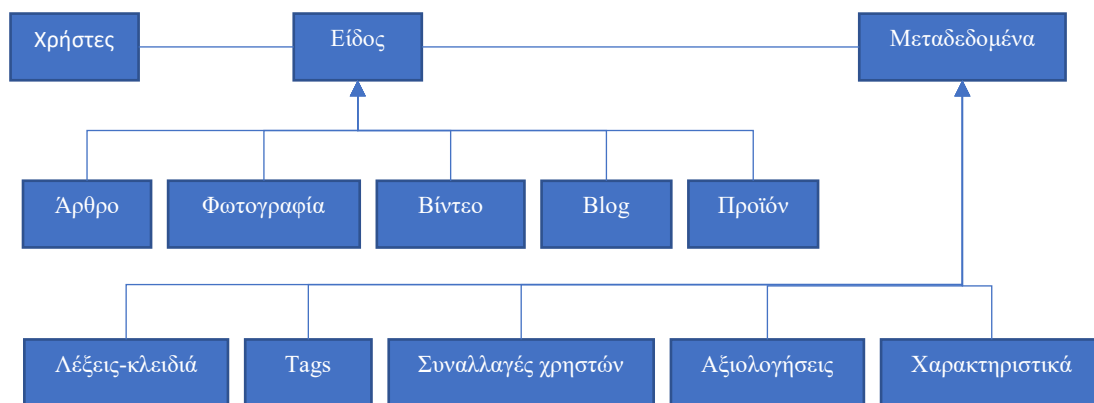
Εικόνα 7, Η διαδικασία βελτιστοποίησης του πλαισίου H-ISTSH διαδικασία βελτιστοποίησης του πλαισίου H-ISTS (Dimah et al, 2015)

Στο H-ISTS, οι τιμές της παραμέτρου αξιοπιστίας βελτιστοποιούνται δυναμικά χρησιμοποιώντας μια εξελικτική διαδικασία. Οι παράμετροι αξιοπιστίας ορίζονται τυχαία για τη δημιουργία ενός αρχικού πληθυσμού χρωμοσωμάτων. Όλες οι τιμές των παραμέτρων εφαρμόζονται σε χρωμοσώματα με τιμές πραγματικού αριθμού. Στη διαδικασία επιλογής χρωμοσωμάτων, το προτεινόμενο μοντέλο μας εφαρμόζει μια μέθοδο τουρνουά. Οι τελεστές crossover και μετάλλαξης χρησιμοποιούνται για την τροποποίηση των χρωμοσωμάτων και τη διασφάλιση ότι μόνο τα καλύτερα χρωμοσώματα σε κάθε γενιά επιβιώνουν και μετακινούνται στην επόμενη γενιά. Η λειτουργία φυσικής κατάστασης αξιολογεί τον υποψήφιο απόγονο. Αυτές οι διαδικασίες είναι συνεχείς και σε κάθε επανάληψη δοκιμάζονται τα κριτήρια τερματισμού για να σταματήσει η εκτέλεση του αλγορίθμου.

## 4. Μετρικές αξιολόγησης των Social Recommendation συστημάτων

### 4.1 Τεχνικές ανατροφοδότησης των SR συστημάτων

Όπως αναφέρθηκε παραπάνω τα SR συστήματα στοχεύουν στην εξαγωγή πληροφοριών σχετικά με τους χρήστες. Η παρακάτω εικόνα δείχνει τις διάφορες πιθανές ενέργειες και τα σχόλια των χρηστών για τα διάφορα στοιχεία (π.χ. άρθρα, προϊόντα, βίντεο κ.λπ.) σε ένα ιστότοπο ή εφαρμογή. Ουσιαστικά διαμορφώνονται μεταδεδομένα για όλους τους τύπους σχολίων που καταχωρεί ο χρήστης. Η άμεση ή η σιωπηρή ανατροφοδότηση απαιτεί εκτεταμένη ανάλυση για εξαγωγή πληροφοριών που βοηθούν στον προσδιορισμό ενός στοιχείου (Grasso and Convertino, 2004).



Εικόνα 8. Αλληλεπίδραση του χρήστη με διαφορετικά στοιχεία - μεταδεδομένα

Τυπικά, τα SR συστήματα αποθηκεύουν το ιστορικό των χρηστών, τις συναλλαγές τους με το σύστημα και τα σχόλια σχετικά με προϊόντα και άλλα στοιχεία για να εξαγάγουν αρκετές πληροφορίες σχετικά με τα ενδιαφέροντα των χρηστών. Για παράδειγμα, ένα αρχείο καταγραφής συναλλαγών μπορεί να καταγράψει μια αναφορά και μια περιγραφή του αντικειμένου που χρησιμοποιήθηκε από ορισμένους χρήστες, όπως το ιστορικό αγορών, τις αξιολογήσεις, τις κριτικές σε προϊόντα κ.λπ. Στην πραγματικότητα, τα πιο συχνά χρησιμοποιούμενα δεδομένα που αφορούν συναλλαγές και συλλέγονται από τα SR συστήματα είναι οι αξιολογήσεις των χρηστών.

Η αξιολόγηση ενός προϊόντος από τον χρήστη μπορεί να είναι σε δύο μορφές, ρητή και σιωπηρή. Στη ρητή περίπτωση, το σύστημα ζητά από το χρήστη να αξιολογήσει ένα προϊόν και του δίνεται μία κλίμακα για να δώσει τη δική του βαθμολογία και να εκφράσει επιπλέον απόψεις για το προϊόν. Ωστόσο, η σιωπηρή βαθμολογία χρησιμοποιείται σε περιπτώσεις όπου το σύστημα καταχωρεί το ίδιο τις αξιολογήσεις. Επομένως, η αυτόματη αξιολόγηση της ποιότητας ενός προϊόντος μπορεί να διευκολύνει καθώς κρίνει πως οι χρήστες αξιολογούν τα προϊόντα με απλό τρόπο. Σύμφωνα με τους Herlocker et al. (2004), η αναπαράσταση των

αξιολογήσεων μπορεί να γίνει σε διαφορετικά στυλ και κλίμακες, όπως περιγράφεται παρακάτω:

- Αριθμητική αξιολόγηση, όπως η κλίμακα 1 έως 5 με τη μορφή αστεριών.
- Αξιολόγηση με ένα σύνολο εύκολα αναγνωρισμένων εκφράσεων ή συμβόλων, για παράδειγμα («μου αρέσει αρκετά», «μου αρέσει», «ουδέτερος», «δεν μου αρέσει», «δεν μου αρέσει αρκετά»), οπότε το σύστημα ζητά από το χρήστη να επιλέξει μία από αυτές τις τιμές η οποία αντικατοπτρίζει κατά πολύ τη γνώμη του.
- Αξιολογήσεις δυαδικού τύπου όπου το σύστημα ζητά από το χρήστη να υποδείξει εάν ένα συγκεκριμένο προϊόν ή άλλο αντικείμενο είναι απλά καλό ή κακό.
- Αξιολογήσεις τύπου ναι/όχι για να μάθει το σύστημα αν ένας χρήστης αγόρασε ένα στοιχείο ή αν αξιολόγησε το στοιχείο θετικά.

Οι πληροφορίες που εξάγονται από τις αξιολογήσεις είναι παραδείγματα ρητών σχολίων που μπορούν να χρησιμοποιηθούν άμεσα από το σύστημα. Από την άλλη πλευρά, η έμμεση ανατροφοδότηση μπορεί επίσης να εξαχθεί από το σύστημα και μπορεί να λάβει διάφορες μορφές. Για παράδειγμα, γράφοντας σχόλια και αξιολογήσεις για προϊόντα ή αντικείμενα που έχει χρησιμοποιήσει ο χρήστης, τα clicks που γίνονται σε ορισμένα αντικείμενα, ο χρόνος που αφιερώνεται σε μια συγκεκριμένη ιστοσελίδα, που αγοράστηκε και επέστρεψε αντικείμενα, η λίστα παρακολούθησης σε ιστότοπους όπως το YouTube, η επεξεργασία λίστας επιθυμιών σε ένα ηλεκτρονικό κατάστημα και, τέλος, η επισήμανση στοιχείων.

Η χρήση σιωπηρών σχολίων απαιτεί από το σύστημα να αναλύσει τη συμπεριφορά των χρηστών και να παρατηρήσει το ιστορικό των αλληλεπιδράσεων. Και οι δύο μορφές ανατροφοδότησης από τη πλευρά του χρήστη μπορούν να αντικατοπτρίζουν μια σειρά διαφορετικών απόψεων είτε θετικών είτε αρνητικών. Ωστόσο, οι σιωπηρές πληροφορίες είναι θορυβώδεις και ενδέχεται να προκαλέσουν πολλά λάθη. Παρέχουν στο σύστημα την ευκαιρία να συλλέξουν ένα μεγάλο σύνολο δεδομένων για εκπαιδευτικούς λόγους. Σε αντίθεση, οι ρητές πληροφορίες είναι ξεκάθαρες αν και οι χρήστες τείνουν να βαθμολογούν μόνο έναν μικρό αριθμό στοιχείων, γεγονός που οδηγεί σε περιορισμένο αριθμό δεδομένων. Πράγματι, απαιτείται μια μέθοδος για να επαληθεύσει ότι έχουν εξαχθεί οι σωστές απόψεις και κριτικές των χρηστών και είναι διαθέσιμες προς ανάλυση (Grasso and Convertino, 2012).

Σε γενικές γραμμές, μια κριτική μπορεί να θεωρηθεί ως ένα σύνολο λέξεων, στις οποίες μόνο μερικοί από αυτούς τους όρους διατηρούν ένα συναίσθημα ή αντιπροσωπεύουν μία κριτική γνώμη. Το παραπάνω επηρεάζει πλέον τη ψηφιακή οικονομία και τις εκατομμύρια συναλλαγές μέσα σε αυτή ή τις αποφάσεις των χρηστών. Είναι προφανές ότι η απόκτηση καλής φήμης συνδέεται με τη συγκέντρωση πολλών θετικών κριτικών. Συνήθως, μια κριτική διαμορφώνεται μέσα από ρητές αξιολογήσεις που καθορίζονται από μεταδεδομένα όπως ο αριθμός των αστεριών.

Προκύπτει από έρευνες ότι η χρησιμότητα των μηχανισμών ανατροφοδότησης δεν συνίσταται κυρίως στις αξιολογήσεις με αριθμητικό τρόπο που κοινοποιούνται από άγνωστους μεταξύ τους χρήστες, αλλά μάλλον από τα σχόλια και το κείμενο που τα συνοδεύουν. Για αυτό το σκοπό έχει εφαρμοστεί και η ιδέα της μέσης πιθανότητας υποκειμενικότητας σε μια πρόταση και στα πλαίσια μιας κριτικής, όπως επίσης έχει εφαρμοστεί και η έννοια της τυπικής

απόκλισης σε μία συγκεκριμένη πρόταση και στα πλαίσια μιας κριτικής. Με αυτό τρόπο μπορεί να ελεγχθεί εάν η κριτική περιλαμβάνει μία ποικιλία υποκειμενικών και αντικειμενικών προτάσεων. Αυτό γιατί οι αξιολογήσεις των ανθρώπων περιέχουν διαφορετικά αποτελέσματα όταν είναι καθαρά υποκειμενικά, καθαρά αντικειμενικά ή αναμειγμένα μεταξύ τους.

#### 4.1.1 Τυπικοί μηχανισμοί ανατροφοδότησης

Έχουν γίνει διάφορες μελέτες για την εξαγωγή συμπερασμάτων για τα συναισθήματα των χρηστών όταν υποβάλλουν σχόλια για ένα προϊόν ή υπηρεσία. Για παράδειγμα, ερευνητές έχουν προτείνει τεχνικές βελτίωσης των SR συστημάτων με την αξιοποίηση πληροφοριών για θέματα και συναισθήματα στο περιεχόμενο των προτάσεων (Zou et al 2015). Μία από αυτές είναι η ανάλυση αξιολογήσεων από σχόλια κειμένου σε μία κλίμακα βαθμολογίας πολλαπλών σημείων και η εφαρμογή μεθόδων παλινδρόμησης για την ανάλυση των συναισθημάτων σε κλίμακες. Μία άλλη τεχνική, γνωστή ως Probabilistic Rating Inference Framework (Leung et al 2011) εφαρμόζει τεχνικές φυσικής γλώσσας για να αναλύσει τα συναισθήματα στα σχόλια των χρηστών. Η μέθοδος αυτή χρησιμοποιεί Naive Bayesian classifier για να ταξινομήσει τα αποτελέσματα των βαθμολογιών και, στη συνέχεια, εφαρμόζει ένα συνεργατικό αλγόριθμο φιλτραρίσματος για να βελτιώσει την ακρίβεια των προτεινόμενων στοιχείων στους πελάτες.

Μία άλλη τεχνική υλοποίησε έναν αλγόριθμο κατάταξης των αποτελεσμάτων με βάση την κατανομή Bernoulli. Αρχικά εφαρμόζει έναν πίνακα παραγοντοποίησης για την κανονικοποίηση των περιεχομένων των κριτικών πριν προχωρήσει σε υπολογισμό των πιθανοτήτων.

## 4.2 Αναγνώριση χρηστών με επιρροή

Το πρόβλημα του εντοπισμού των επιρροών έχει αποσπάσει πολλή μεγάλη προσοχή σε αρκετούς ερευνητικούς τομείς καθώς έχει εφαρμογές στο μάρκετινγκ (Lei et al, 2015), (Weng et al, 2010), τη πρόληψη και τη διάδοση ασθενειών (Kim et al, 2015) (Rossi et al, 2015), τη πολιτική (Huckfeldt and Sprague, 1987), (Dubois and Gaffney, 2014), τις έξυπνες πόλεις (Yang et al, 2016), πλατφόρμες κοινής χρήσης λογισμικού (Kaple et al, 2017). Έχουν καταγραφεί διάφοροι τρόποι προσέγγισης αυτού του προβλήματος και ορισμένες υποθέσεις που έχουν γίνει πηγάζουν από τη θεωρία των γραφημάτων. Σε ένα τέτοιο γράφημα τα επίπεδα διάδοσης μίας επιρροής ορίζονται οι κατευθύνσεις μεταξύ των κόμβων του γραφήματος αφού αυτό αρχικοποιηθεί. Ενώ οι συνθήκες ενεργοποίησης (δηλ. η "επιρροή") ορίζονται ως κόμβοι.

**Εμπειρική μέθοδος.** Μία τέτοια προσέγγιση ακολουθήθηκε από τους Domingos και Richardson (2001). Ο αλγόριθμος τους εφαρμόζει τεχνικές εξόρυξης δεδομένων στο πεδίο του μάρκετινγκ και μοντελοποιούν τις αγορές ως κοινωνικά δίκτυα. Προσδιόρισαν την εξάπλωση της επιρροής ως συνάρτηση πιθανολογικών μοντέλων αλληλεπιδράσεων. Κάθε κόμβος σχετίζεται με μια τιμή που μετράει το βαθμό στον οποίο μπορεί να επηρεάσει άλλες κορυφές

και χρησιμοποιείται για τον βέλτιστο προσδιορισμό όσων κόμβων επηρεάζονται από αυτόν. Σε δική τους έρευνα τα αποτελέσματα δείξαν ότι η προτεινόμενη στρατηγική μάρκετινγκ αποδίδει πολύ καλύτερα από δύο εναλλακτικές στρατηγικές.

**Ανάλυση βάση συσχετίσεων (correlations).** Σε μία τέτοια έρευνα (Badashian and Stroulia, 2016) οι επιρροές προσδιορίζονται από μία διαδικασία υπολογισμού των συσχετίσεων μεταξύ διαφορετικών μετρικών επιρροής. Η έρευνα πραγματοποιήθηκε στο Twitter και στη συσχέτιση των δεδομένων από Follow, Mention, Retweet. Σε

**Simulated Annealing (SA).** Οι Jiang et al. (2011) χρησιμοποίησαν τον αλγόριθμο αυτό για την εύρεση των πρώτων κ κόμβων με τη μέγιστη επιρροή στο γράφημα - κοινωνικό δίκτυο και λαμβάνοντας υπόψη τυχών επικαλύψεις μεταξύ των σχέσεων.

**NewGreedy.** Ο αλγόριθμος αυτός (Eirinaki et al, 2016) εισήχθη για την αναγνώριση επιρροών σε ένα μη κατευθυνόμενο γράφημα. Η προϋπόθεση ήταν ότι ένας κόμβος επηρεάζεται εάν ένας αριθμός άμεσων γειτόνων «ενεργοποιούνται» (δηλαδή επηρεάζονται / επηρεάζουν). Αν όχι τότε θεωρείται ότι δεν υφίσταται κάποια επιρροή.

**Linear Threshold Model and Independent Cascade Model.** Οι Kempe et al. (2015) όρισαν το πρόβλημα βελτιστοποίησης της επιλογής των πιο σημαντικών κόμβων σε ένα κοινωνικό δίκτυο ως πολυπλοκότητας NP. Με το linear threshold model, ένας κόμβος  $v$  επιλέγει τυχαία ένα κάτω άκρο μεταξύ 0 και 1 και ενεργοποιείται όταν το συνδυασμένο αποτέλεσμα από τους γείτονές του υπερβαίνει το κάτω άκρο, μόνο λαμβάνοντας υπόψη τους γείτονες που ήταν ενεργοί στην προηγούμενη επανάληψη. Με αυτόν τον τρόπο, το κάτω άκρο αλλάζει δυναμικά με κάθε επανάληψη. Στο δε cascade model, ένας κόμβος  $u$  μπορεί να ενεργοποιήσει ένα κόμβο  $v$  με μία πιθανότητα που έχει λάβει υπόψη τους γείτονες που έχουν ήδη δοκιμάσει και αποτύχει να ενεργοποιήσουν τον  $v$ .

**Dijkstra's algorithm.** Οι Hangal et al (2010) εστιάζουν αρκετά στη σημασία της διάδοσης επιρροής για μη κατευθυνόμενα γραφήματα και υπολογίζουν τη πιο σημαντική διαδρομή μεταξύ δύο κόμβων με μία έκδοση του αλγορίθμου Dijkstra. Έδειξαν πειραματικά ότι η διαδρομή με τη σημαντικότερη επιρροή είναι περισσότερο αποτελεσματική σε σύγκριση με τη συντομότερη διαδρομή σε ένα κοινωνικό γράφημα που ενσωματώνει κατευθυνόμενες και σταθμισμένες ακμές επιρροής. Η επιρροή ενός κόμβου ορίζεται ως το άθροισμα όλης της επιρροής που έχει ο κόμβος σε άλλους κόμβους.

#### 4.3 Μετρικές εμπιστοσύνης στα SR συστήματα

Το θέμα της εμπιστοσύνης στα SR συστήματα συσχετίζεται με την ομοιότητα μεταξύ των προτιμήσεων των χρηστών. Η εμπιστοσύνη σχετίζεται με την πεποίθηση ότι τρίτοι μπορούν να παρέχουν πολύτιμες βαθμολογίες και αξιολογήσεις αντικειμένων. Πιο συγκεκριμένα, η προοπτική της εμπιστοσύνης στα SR συστήματα έχει σημαντικό ρόλο στην επίλυση περιορισμών ή άλλων ζητημάτων. Η εμπιστοσύνη του καταναλωτικού κοινού αυξάνει την παραγωγικότητα των συστημάτων και λαμβάνεται υπόψη με μεγαλύτερη επιρροή η



διαδικτυακή συμπεριφορά των χρηστών για την παραγωγή συστάσεων (Golbeck, and Hendler, 2006).

Συγκεκριμένα, υπάρχουν δύο διαφορετικές χρήσεις του βαθμού εμπιστοσύνης στη δημιουργία προτάσεων:

1. εμπιστοσύνη στις σχέσεις μεταξύ χρηστών και
2. εμπιστοσύνη στην παραγωγή προτάσεων του συστήματος.

Τα συστήματα που λαμβάνουν υπόψη το βαθμό εμπιστοσύνης χρησιμοποιούν τη σχέση που υπάρχει μεταξύ των χρηστών για την εφαρμογή σχετικών αλγορίθμων. Οι ρόλοι που αποκτά το σύστημα στη περίπτωση αυτή στηρίζονται στην αποτελεσματικότητα (υποστήριξη των χρηστών να λαμβάνουν τις σωστές αποφάσεις) και τη δύναμη της πειθούς (δηλαδή τη δυνατότητα να πείσει τους πελάτες να δοκιμάσουν ή να αγοράσουν προϊόντα).

Υποθέτοντας τρεις χρήστες α, β και γ σε ένα πραγματικό σενάριο, ο α χρήστης μπορεί να ρωτήσει τον φίλο β για την εμπειρία του χρησιμοποιώντας το στοιχείο σ. Εάν ο φίλος β δεν έχει εμπειρία στη χρήση αυτού του αντικειμένου, τότε ο β μπορεί να ρωτήσει τον φίλο γ για τον / την γνώμη για το σημείο σ. Η πρόταση του φίλου γ μπορεί να είναι αξιόπιστη σε κάποιο βαθμό με βάση τις διαθέσιμες πληροφορίες εμπιστοσύνης. Μπορεί να παρατηρηθεί ότι με αυτόν τον τρόπο, το σύστημα μπορεί να εξατομικεύσει προτάσεις, όπου μπορεί ορισμένα προτεινόμενα αντικείμενα να προσεγγιστούν ως επιθυμητά αφού ήδη τα έχουν προσεγγίσει αξιόπιστοι χρήστες.

#### 4.3.1 Κατηγοριοποίηση των SR συστημάτων που βασίζονται στην εμπιστοσύνη

Αρχικά τα συστήματα μπορούν να χωριστούν σε αυτά που λειτουργούν βάσει πιθανοτήτων και αυτά που λειτουργούν σε στάδια. Οι προσεγγίσεις βάσει πιθανοτήτων μετρούν την εμπιστοσύνη σε ένα ασπρόμαυρο στίλ. Σε αυτήν την περίπτωση, μια υψηλή τιμή σε πιθανότητα δείχνει ότι α πηγή μπορεί να είναι αξιόπιστη. Από την άλλη πλευρά, οι προσεγγίσεις που λειτουργούν σε στάδια ασχολούνται με την πρόβλεψη του βαθμού εμπιστοσύνης όταν το αποτέλεσμα μιας ανάλυσης δεδομένων μπορεί να είναι θετικό σε κάποιο βαθμό. Για παράδειγμα, μερικές πληροφορίες δεν μπορεί να είναι απολύτως σωστές ή λάθος, αλλά μάλλον μπορεί να είναι σωστές σε κάποιο βαθμό ή λάθος σε κάποιο βαθμό (Golbeck, and Hendler, 2006). Πράγματι, στην πραγματική ζωή, διαχειριζόμαστε την εμπιστοσύνη ως ένα φαινόμενο που εκτυλίσσεται σε στάδια. Οι άνθρωποι τείνουν να εμπιστεύονται πολύ μερικούς ανθρώπους ή να τους εμπιστεύονται περισσότερο ή λιγότερο από ότι οι υπόλοιποι.

Εμπιστοσύνη και δυσπιστία: Την τελευταία δεκαετία σημειώθηκε αύξηση της σταδιακής έρευνας στο επίπεδο εμπιστοσύνης. Ωστόσο, οι περισσότερες από αυτές τις μελέτες επικεντρώθηκαν μόνο στον υπολογισμό της εμπιστοσύνης και αγνόησαν τη δυσπιστία. Αυτό συμβαίνει επειδή η αναπαράσταση του βαθμού δυσπιστίας θεωρείται ως μία σχετικά ανεξερεύνητη περιοχή επειδή υπάρχει μια αυξανόμενη



άποψη ότι η δυσπιστία δεν μπορεί να ορίζεται ως η έλλειψη εμπιστοσύνης (Cofta 2006).

Καθολική και τοπική εμπιστοσύνη: Η εμπιστοσύνη μπορεί να εφαρμοστεί ως μία καθολική ή τοπική παράμετρος. Οι μετρήσεις τοπικής αξιοπιστίας υπολογίζουν την εμπιστοσύνη σύμφωνα με τις υποκειμενικές πεποιθήσεις ενός ενεργού χρήστη κατά τη γνώμη άλλων χρηστών. Ως εκ τούτου, η βαθμολογία τοπικής αξιοπιστίας θα διαφέρει μεταξύ των χρηστών καθώς αυτοί έχουν διαφορετικές απόψεις μεταξύ τους (Zejda, 2010). Αντίθετα, η καθολική εμπιστοσύνη μπορεί να αντιμετωπιστεί ως επίπεδο φήμης σε ολόκληρη την κοινότητα και τότε αυτός ο βαθμός μπορεί να είναι ταξινομημένος ως μία αντικειμενική αξία εμπιστοσύνης. Ένα τέτοιο παράδειγμα καθολικής αποδοχής είναι ο αλγόριθμος PageRank που χρησιμοποιείται από τη μηχανή αναζήτησης Google.com (Page et al, 1999).

Τα καθολικά αποτελέσματα εμπιστοσύνης υπολογίζονται ως το γενικό σημείο των απόψεων της κοινότητας σχετικά με έναν συγκεκριμένο χρήστη. Αξίζει να σημειωθεί ότι απαιτείται η τοπική εμπιστοσύνη να υπολογιστεί με ακρίβεια ώστε να ταιριάζει στις απόψεις κάθε χρήστη, κάτι που υπολογιστικά είναι ακριβό, ενώ οι καθολικοί αλγόριθμοι εμπιστοσύνης εκτελούνται μία φορά για όλους τους χρήστες. Δηλαδή, Τα μοντέλα καθολικής εμπιστοσύνης μπορεί να υπολογίζουν την εμπιστοσύνη στις σχέσεις μεταξύ όλων των χρηστών σε μια συγκεκριμένη κοινότητα ενώ οι τοπικές μετρήσεις εμπιστοσύνης εκτιμούν μια εξατομικευμένη βαθμολογία αξιοπιστίας με βάση τις απόψεις του χρήστη. Ωστόσο, η αντικειμενική εκτίμηση εμπιστοσύνης μπορεί να χρησιμοποιεί υποκειμενική εκτίμηση εμπιστοσύνης, κατά τη συγκέντρωση των σταθμισμένων υποκειμενικών τιμών εμπιστοσύνης. Μια άλλη έννοια, η αξιοπιστία, χρησιμοποιείται επίσης στη βιβλιογραφία για να δείξει την καθολική αποδοχή ή εμπιστοσύνη (Symeonidis et al, 2008).

#### 4.3.2 Μετρικές υπολογισμού της εμπιστοσύνης

Η εμπιστοσύνη μπορεί να υπολογιστεί ως εκτίμηση για το πόσο ένας χρήστης εμπιστεύεται έναν άλλο χρήστη κατανοώντας τις κοινές συνδέσεις και συμπεριφορές τους στα κοινωνικά δίκτυα. Στη βιβλιογραφία δύο στρατηγικές χρησιμοποιούνται για τη δημιουργία μετρήσεων του βαθμού εμπιστοσύνης σε SR συστήματα: η διάδοση και η συγκέντρωση (Victor et al, 2011):

- Διάδοση: Η υπόθεση πίσω από τη διάδοση είναι ότι η εμπιστοσύνη θεωρείται μεταβατική σε κάποιο βαθμό. Για παράδειγμα, όταν ο χρήστης  $u_1$  εμπιστεύεται τον χρήστη  $u_2$ , ο  $u_2$  ονομάζεται ένα τρίτο αξιόπιστο μέρος (TTP) και εάν το TTP  $u_2$  εμπιστεύεται τον χρήστη  $u_3$ , τότε είναι πιθανό να υποθέσουμε ότι ο  $u_1$  μπορεί να εμπιστευτεί τον  $u_3$  σε κάποιο βαθμό. Αυτή η βασική στρατηγική ονομάζεται ατομική άμεση διάδοση. Μπορεί να υπάρχουν

πολλές διαδρομές μεταξύ δύο χρηστών και φυσικά αυτές οι διαδρομές μπορεί να έχουν διαφορετικά μήκη. Για παράδειγμα, εάν υπάρχουν δύο διαδρομές μεταξύ του χρήστη προέλευσης  $u_1$  προς το στόχο χρήστη  $u_3$ , ενδέχεται να γίνουν διαφορετικές ενέργειες, οι οποίες θα βασίζονται στο συνολικό μήκος της διαδρομής, ή μπορεί να γίνει μια επιλογή που θα ακολουθήσει τους συντομότερους συνδέσμους διαδρομών μεταξύ δύο χρηστών.

- Συγκέντρωση. Μια μέτρηση εμπιστοσύνης μπορεί επίσης να χρησιμοποιεί μια στρατηγική συγκέντρωσης. Για να παρουσιάσουμε καλύτερα αυτήν την τεχνική, ας θεωρήσουμε ότι πολλές διαδρομές συνδέονται με έναν ενεργό χρήστη, για τον οποίο το σύστημα προσπαθεί να προβλέψει μια βαθμολογία αξιοπιστίας σε ένα μεγάλο δίκτυο. Σε αυτήν την περίπτωση, η πρόβλεψη για εμπιστοσύνη μπορεί να δημιουργηθεί μέσω διαφορετικών διαδρομών διάδοσης, οι οποίες πρέπει να ενσωματωθούν σε μια συγκεντρωτική εκτίμηση.

Συνδυάζεται συχνά η διάδοση και η συγκέντρωση στρατηγικών, και το τελικό αποτέλεσμα της αξιολόγησης της εμπιστοσύνης μπορεί να εξαρτάται από τον τρόπο που χρησιμοποιούνται από κοινού. Οι μετρικές συγκέντρωσης και βαθμού διάδοσης μπορεί να χρησιμοποιηθούν ως σταθμισμένοι τελεστές σε ένα σταθμισμένο άθροισμα, ένα μέσο όρο ή ένα σταθμισμένο μέσο όρος στη διαδικασία δημιουργίας συστάσεων.

Οι μετρικές του βαθμού εμπιστοσύνης στην έρευνα γύρω από τα SR συστήματα μπορούν να χωρισθούν επίσης σε δύο κύριες κατηγορίες:

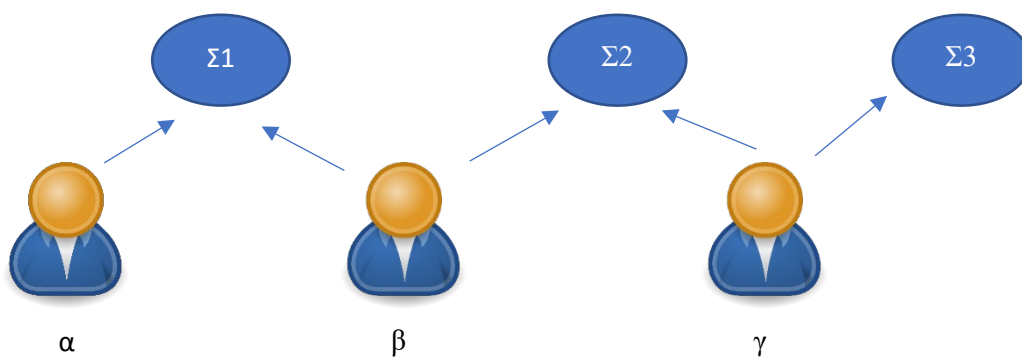
- Σαφή δεδομένα για τον βαθμό εμπιστοσύνης που χρησιμοποιούνται στη δημιουργία ενός δικτύου εμπιστοσύνης σε SR συστήματα.
- Μη σαφή δεδομένα που έχουν προκύψει κατά τη διαδικασία ταξινόμησης σε SR συστήματα.

Στη πρώτη περίπτωση, τα συστήματα συστάσεων με βάση την εμπιστοσύνη βασίζονται στην εκτίμηση των απόψεων από αξιόπιστους χρήστες αντί της ομοιότητας μεταξύ δύο χρηστών. Τέτοιες στρατηγικές στοχεύουν στη δημιουργία δικτύων εμπιστοσύνης, τα οποία δημιουργούνται ξεχωριστά από τους χρήστες εκχωρώντας ένα βαθμό εμπιστοσύνης σε άλλους χρήστες. Αυτά είναι τα λεγόμενα δίκτυα εμπιστοσύνης (Web of Trust - WOT) (Golbeck and Hendler, 2006).

Για παράδειγμα, η παρακάτω εικόνα δείχνει πώς μπορούν τα δίκτυα εμπιστοσύνης να δώσουν την ευκαιρία να προσεγγιστούν περισσότεροι χρήστες και αντικείμενα. Αν ο χρήστης  $\alpha$  και ο χρήστης  $\beta$  έδωσαν παρόμοια βαθμολογία στο στοιχείο  $\Sigma_1$  τότε συνδέθηκαν μεταξύ τους μέσω των αξιολογήσεών τους. Ως αποτέλεσμα μπορούμε επίσης να πούμε ότι ο χρήστης  $\alpha$  μπορεί να ενδιαφέρεται για το  $\Sigma_2$ . Ωστόσο, δεν υπάρχει τρόπος να το μάθουμε αν ο χρήστης θα ήθελε το στοιχείο  $\Sigma_3$ . Αυτή η

κατάσταση μπορεί να επιλυθεί όταν ένα δίκτυο εμπιστοσύνης προσαρμόζεται μεταξύ των χρηστών. Οι σταθερές γραμμές δείχνουν τις σχέσεις εμπιστοσύνης μεταξύ των χρηστών, όπου ο χρήστης εκφράζει ένα ορισμένο επίπεδο εμπιστοσύνης στον χρήστη β, και ο χρήστης β δείχνει μια παρόμοια τάση εμπιστοσύνης με τον χρήστη γ. Με την διαμόρφωση σχέσεων εμπιστοσύνης, χτίζεται ένα WOT και έτσι μπορούν να παρέχονται ακριβέστερες συστάσεις. Με βάση την ιδέα της διάδοσης της εμπιστοσύνης, η παρακάτω σχέση  $(\alpha \Rightarrow \beta \Rightarrow \gamma) \Rightarrow (\alpha \Rightarrow \gamma)$ ,

μπορεί να είναι αληθινή και μετά μπορούμε να προτείνουμε το στοιχείο Σ3 στο χρήστη α.



Εικόνα 9. Διαδικασία διαμόρφωσης πρότασης Σ3 προς τον χρήστη α

### Μετρικές έμμεσης εμπιστοσύνης

Για να ξεπεραστούν οι εντοπισμένες αδυναμίες στις προσεγγίσεις που βασίζονται στη ρητή εμπιστοσύνη, έχουν προταθεί SR συστήματα που βασίζονται στην έμμεση κοινωνική εμπιστοσύνη. Συνήθως, οι αλγόριθμοι που βασίζονται στον αυτόματο υπολογισμό κάποιου δείκτη εμπιστοσύνης χρησιμοποιούν ένα ιστορικό αξιολογήσεων των χρηστών στο σύστημα. Πιο συγκεκριμένα, η υψηλή τιμή ενός δείκτη εμπιστοσύνης για ένα συγκεκριμένο χρήστη σημαίνει ότι στο παρελθόν είχε καταχωρήσει συστάσεις που αποδείχθηκαν ακριβείς.

Σε άλλες έρευνες αναλύθηκαν απόψεις χρηστών και κατά πόσο υπήρχε συσχέτιση ανάμεσα στα ενδιαφέροντα ενός χρήστη και στο δείκτη εμπιστοσύνης προς τις προτάσεις του ίδιου χρήστη όπως επίσης συσχέτιση με απόψεις άλλων εταίρων που επίσης είχαν κριθεί αξιόπιστοι. Με άλλα λόγια ο συσχετισμός γίνεται με βάση το δείκτη εμπιστοσύνης και όχι με βάση το βαθμό ομοιότητας ανάμεσα στα ενδιαφέροντα χρηστών.

Σε ένα τέτοιο πλαίσιο που εστιάζει στην επιλογή προτάσεων χρηστών με τον υψηλότερο βαθμό αξιοπιστίας, ερευνητές προτείνουν δύο σχετικές μετρικές εμπιστοσύνης: σε επίπεδο προφίλ (μέσος όρος εμπιστοσύνης στο σύνολο των προφίλ χρηστών) και σε επίπεδο στοιχείου (μέσος όρος εμπιστοσύνης για ένα συγκεκριμένο προφίλ κατά τη δημιουργία συστάσεων για ένα συγκεκριμένο προϊόν ή υπηρεσία). Αυτά τα δύο μοντέλα μέτρησαν την ορθότητα των προτεινόμενων στοιχείων από τον χρήστη - δημιουργό στον χρήστη-στόχο. Δηλαδή, οι προτάσεις που προέρχονται από τον χρήστη  $\alpha$  θεωρούνται έγκυρες μόνο εάν βρίσκονται εντός του πραγματικού προφίλ με τις αξιολογήσεις του χρήστη. Η μετρική αυτή έχει μία δυαδική τιμή (ισχύει / δεν ισχύει).

Η μετρική σε επίπεδο προφίλ δηλώνει ότι όσο πιο επιτυχημένη είναι η διαμόρφωση συστάσεων στο παρελθόν με αρκετή ακρίβεια, τόσο υψηλότερος είναι ο δείκτης αξιοπιστίας για έναν χρήστη – δημιουργό και αποκτά μεγαλύτερη κατάταξη σε σχέση με ένα προφίλ που εμφανίζει κακό ιστορικό προτάσεων.

Ο Kim et al (2015), παρουσίασε ένα υπολογιστικό πλαίσιο για την εκτίμηση του επιπέδου εμπιστοσύνης μεταξύ ενός ζεύγους χρηστών σε διαδικτυακές κοινότητες με βάση την αξιολόγηση (π.χ. κριτικές) χωρίς να δίνεται ένα σαφές WOT. Στο πλαίσιο τους, συνδύασαν την έννοια της φήμης έχοντας ως πηγή τη συλλογή απόψεων από δημιουργούς περιεχομένου και την έννοια της εμπιστοσύνης κατά τον ορισμό των προτιμήσεων τους. Χρησιμοποίησαν κυρίως τις αξιολογήσεις στα σχόλια πάνω στο περιεχόμενο και όχι το WOT γιατί έκριναν ότι οι χρήστες αξιολογούν πιο συχνά προϊόντα / υπηρεσίες παρά κάνουν σχόλια πάνω σε σχόλια άλλων χρηστών. Ένα μειονέκτημα αυτής της μελέτης τους είναι ότι βασίστηκε μόνο στις αξιολογήσεις (βαθμολογίες) και όχι στα σχόλια μέσα στο κείμενο του χρήστη.

Ωστόσο, σε άλλες περιπτώσεις η έρευνα ανέδειξε ότι η αξιολόγηση βάση εμπιστοσύνης έχει τρεις βασικές αδυναμίες:

1. Οι σχέσεις εμπιστοσύνης διαφέρουν από τις κοινωνικές φιλίες σε πολλές πτυχές. Συνήθως, στον Ιστό, ένας χρήστης μπορεί να προσθέσει έναν χρήστη  $\beta$  σε μια λίστα εμπιστοσύνης, όταν σε ένα χρήστη  $\alpha$  αρέσει και δείχνει κάποιο ενδιαφέρον για τη γνώμη του χρήστη  $\beta$ . Για το σκοπό αυτό, η διαδικασία δημιουργίας εμπιστοσύνης είναι μια μονομερής δράση. Αυτό σημαίνει ότι αυτή η διαδικασία δεν απαιτεί ο χρήστης  $\beta$  να επιβεβαιώσει τη σχέση προσθέτοντας τον χρήστη  $\alpha$  στη λίστα εμπιστοσύνης του. Επιπλέον, αυτό δεν σημαίνει ότι ο χρήστης γνωρίζει τον χρήστη  $\beta$  στην πραγματική ζωή. Από την άλλη πλευρά, οι «κοινωνικές φιλίες» υποδηλώνουν τις σχέσεις συνεργασίας και τις συνδέσεις γύρω μας. Για παράδειγμα, συμμαθητές, συγγενείς ή αγαπημένοι μας συγγραφείς ή καλλιτέχνες.

Πολλοί από τους ιστότοπους κοινωνικής δικτύωσης, όπως το Facebook και το Twitter, έχουν σχεδιαστεί για τη δημιουργία διαδικτυακών συνδέσεων μεταξύ φίλων στον πραγματικό κόσμο. Με βάση αυτή τη μελέτη, μπορούμε να καταλήξουμε στο συμπέρασμα ότι τα συστήματα που βασίζονται στην εμπιστοσύνη δεν περιλαμβάνουν την έννοια της «κοινωνικής πρότασης» επειδή αυτή η έννοια συνεπάγεται ένα κοινωνικό δίκτυο φίλων που ενσωματώνεται σε ένα σύστημα δημιουργίας συστάσεων. Αυτή η διαφορά μπορεί να προκαλέσει μια αξιοσημείωτη διαφορά και στο σχεδιασμό των μοντέλων εμπιστοσύνης.

2. Η υπόθεση των συστημάτων σύστασης που βασίζονται στην εμπιστοσύνη είναι ότι μπορεί να οικοδομηθεί η εμπιστοσύνη όταν δύο χρήστες  $\alpha$  και  $\beta$  έχουν παρόμοια γούστα. Στην πραγματικότητα, αυτό δεν συμβαίνει πάντα στις κοινωνικές προτάσεις, καθώς οι προτιμήσεις ενός χρήστη δεν κοινοποιούνται απαραίτητα σε όλους τους φίλους τους. Με άλλα λόγια, ένα άτομο μπορεί να έχει κοινά γούστα με κάποιους φίλους, ενώ οι άλλοι φίλοι του μπορούν να έχουν εντελώς διαφορετικές προτιμήσεις. Ως εκ τούτου, οι αλγόριθμοι των παραδοσιακών προτάσεων που βασίζονται στην εμπιστοσύνη δεν μπορούν να είναι άμεσα εφαρμόσιμοι για την παροχή συστάσεων σε συστήματα κοινωνικής πρότασης λόγω της διαφορετική υπόθεσης σχετικά με το βαθμό ομοιότητας και εμπιστοσύνης.
3. Η τεράστια ανάπτυξη των web εφαρμογών έχει τεράστιο αντίκτυπο στους ανθρώπους. Άνθρωποι αφιερώνουν όλο και περισσότερο χρόνο σε διαδικτυακά κοινωνικά δίκτυα που είναι μία περισσότερο ελκυστική δραστηριότητα από την απλή πλοήγηση στο διαδίκτυο. Από την άλλη πλευρά, υπάρχουν λίγα διαδικτυακά συστήματα που λαμβάνουν υπόψη την έννοια της εμπιστοσύνης. Έτσι, για την παροχή εξατομικευμένων συστάσεων στους διαδικτυακούς χρήστες, οι ερευνητές θα πρέπει να διεκδικούν τις δικές τους μελέτες που να περιλαμβάνουν κοινωνική χροιά.

Επιπλέον, ερευνητές (Jamali and Ester, 2010). επέλεξαν τη μέθοδο της παραγοντοποίησης με πίνακες χαμηλού βαθμού για να παραγοντοποιήσουν τον πίνακα με τις αξιολογήσεις των χρηστών (χρήστης – αντικείμενο – αξιολόγηση). Τα δεδομένα στοιχεία που είχαν παραγοντοποιηθεί χρησιμοποιήθηκαν κατόπιν για την πρόβλεψη περισσότερων τιμών που λείπουν. Το προτεινόμενο μοντέλο βασίστηκε στην υπόθεση ότι οι άνθρωποι λαμβάνουν υπόψη τις προτάσεις των φίλων τους για να πάρουν μια απόφαση. Το μοντέλο ορίζει δύο σύνολα φίλων: το F+ σημαίνει το σύνολο των χρηστών που ο χρήστης θεωρεί φίλους και F- είναι το σύνολο των χρηστών που θεωρούν τον χρήστη ως φίλο. Ενδεχομένως αυτά τα δύο σύνολα να είναι ίδια σε ορισμένες περιπτώσεις. Για παράδειγμα, στο facebook, τα F+ και F- είναι όμοια, εφόσον ο χρήστης  $\alpha$  είναι στη λίστα των φίλων του  $\beta$  και αντίστοιχα ο  $\beta$  είναι

στη λίστα φίλων του χρήστη  $\alpha$ . Σε άλλο μέσο κοινωνικής δικτύωσης το  $F +$  δεν είναι ίσο με το  $F-$ , πράγμα που σημαίνει ότι δεν υπάρχει αμφίδρομη σχέση απαραίτητα.

Με το μοντέλο αυτό οι ερευνητές δοκίμασαν να ελαχιστοποιήσουν τη διαφορά στις προτιμήσεις μεταξύ ενός συγκεκριμένου χρήστη και των φίλων του. Μια συνάρτηση ομοιότητας χρησιμοποιείται για να δείξει εάν η απόσταση μεταξύ των χαρακτηριστικών του ενός χρήστη και των χαρακτηριστικών των φίλων τους είναι μικρή ή μεγάλη για να αντιμετωπιστούν ανάλογα οι φίλοι των χρηστών. Η συνάρτηση υπολογισμού της ομοιότητας χρησιμοποιεί τεχνικές όπως Vector Space Similarity (VCC) και Pearson Correlation Coefficient (PCC).

Ωστόσο, υπάρχουν ορισμένα μειονεκτήματα που έχουν εντοπιστεί σε αυτή τη μελέτη. Πρώτον, το μοντέλο βασίστηκε στον πίνακα αξιολόγησης χρήστη-στοιχείου και επομένως αυτά τα μοντέλα μπορούν να λειτουργήσουν μόνο όταν δίνονται ικανοποιητικά δεδομένα με αξιολογήσεις. Επιπλέον, το μοντέλο υποθέτει ότι οι φίλοι στα κοινωνικά δίκτυα βαθμολογούν επίσης προϊόντα ή υπηρεσίες, και αυτό δεν ισχύει πάντα. Σε μεγαλύτερο ποσοστό οι χρήστες σε αυτά τα δίκτυα εκφράζουν τη γνώμη τους προς τους φίλους τους γράφοντας διάφορα σχόλια. Με άλλα λόγια, το μοντέλο δεν χρησιμοποιεί εξωτερικές πηγές δεδομένων που σχετίζονται με συνδέσεις φίλων, όπως το Twitter. Επιπλέον, επικεντρώθηκε στη βελτίωση των προβλέψεων για περισσότερη ακρίβεια, ωστόσο αγνόησε τα προβλήματα της έλλειψης δεδομένων κατά την δημιουργία νέων προφίλ («κρύα δεδομένα») και τη πρόκληση του ζητήματος της ποικιλομορφίας.

Με βάση τα παραπάνω εξάγονται σημαντικές παρατηρήσεις σχετικά με τα κενά και τους περιορισμούς στα παραπάνω μοντέλα που χρησιμοποιούν τεχνικές implicit trust (εμπιστοσύνη με έμμεσο τρόπο) σε συστήματα συστάσεων. Τα περισσότερα από τα αναφερόμενα μοντέλα περιόρισαν τον όρο «κοινωνική πρόταση» μόνο στους συνδρομητές των συστημάτων και αγνοήθηκαν οι συνδέσεις με φίλους σε μέσα κοινωνικής δικτύωσης. Τα περισσότερα από αυτά τα μοντέλα χρησιμοποίησαν δεδομένα από αξιολογήσεις για την παροχή προτάσεων προς άλλους χρήστες και δεν περιελάμβανε τη χρήση άλλης πηγής δεδομένων. Επιπλέον, η εμπιστοσύνη ερμηνεύτηκε ως ομοιότητα μεταξύ χρηστών και δεν αντικατοπτρίζει τη συμπεριφορά των χρηστών ως μια άλλη ένδειξη σχέσεων εμπιστοσύνης.

#### 4.3.3 Προκλήσεις στα μοντέλα εμπιστοσύνης

Αρκετά ζητήματα και περιορισμοί θα πρέπει να αντιμετωπιστούν προκειμένου να ενσωματωθεί κάποιο μοντέλο εμπιστοσύνης σε συστήματα δημιουργίας αυτοματοποιημένων προτάσεων. Αυτές οι δυσκολίες αναλύονται ως εξής:

1. Όσον αφορά το πρόβλημα της ψυχρής εκκίνησης, ο Massa και ο Avesani (2005) θεωρούν ότι όταν αυτό το ζήτημα παρουσιάζεται σε τυπικά συστήματα δημιουργίας προτάσεων, μπορεί να μετριαστεί με την καθιέρωση ενός δικτύου εμπιστοσύνης μεταξύ των χρηστών. Για να γίνει αυτό, οι νέοι χρήστες πρέπει να εκφράσουν με διάφορους τρόπους ποιους χρήστες ή στοιχεία εμπιστεύονται για να λάβουν τις πρώτες προτάσεις από το σύστημα.

Ωστόσο οι Victor and Cornelis (2011) θεωρούν ότι αυτό δεν είναι πάντα σύνηθες για τους νέους χρήστες και ότι πολύ συχνά αντιμετωπίζουν την πρόκληση του κρύου ξεκινήματος ειδικά σε WOT συστήματα. Ως εκ τούτου, προκειμένου να υπάρχει ένα δίκτυο εμπιστοσύνης οι νέοι χρήστες πρέπει να δημιουργήσουν συνδέσεις με άλλους χρήστες και να επεκτείνουν την εμπιστοσύνη τους στο δίκτυο μόλις μπουν σε αυτό.

2. Περισσότερη οπτικοποίηση των κοινωνικών σχέσεων: το ξεκίνημα των σχέσεων σε ένα δίκτυο εμπιστοσύνης φαίνεται να είναι ένα σημαντικό πρόβλημα. Η χρήση περισσότερων μέσων οπτικοποίησης μπορεί να αντιμετωπίσει αυτό το πρόβλημα με την εισαγωγή εργαλείων που επιτρέπουν ένα τύπο συνεργατικού φιλτραρίσματος που βασίζεται στην εμπιστοσύνη. Το σύστημα απεικονίζει τις πληροφορίες που προέρχονται από τον υπολογισμό των όποιων κοινών στοιχείων μεταξύ χρηστών - στοιχείων (πχ με τη τεχνική PCC) και πληροφορίες που περιέχουν κάποιο δείκτη εμπιστοσύνης σύμφωνα με τα τωρινά δεδομένα αξιολόγησης. Έχοντας στη διάθεση τους μία διαδραστική διεπαφή για την εκπροσώπηση των χρηστών στο σύστημα, αυτή η προσέγγιση επιτρέπει στους νέους χρήστες να καταγράψουν τις προτιμήσεις τους και να λάβουν πληροφορίες με δείκτες εμπιστοσύνης σε πραγματικό χρόνο.
3. Διερεύνηση της πιθανότητας δυσπιστίας: λίγες προσπάθειες έχουν γίνει μοντελοποίησης της έλλειψης εμπιστοσύνης (Cofta, 2006) και αυτό οφείλεται στην περιορισμένη διαθεσιμότητα συνόλων δεδομένων που περιλαμβάνουν πληροφορίες δυσπιστίας, και που είναι ένα σημαντικό ζήτημα. Επίσης, δεν υπάρχει μία γενική μέθοδο σχετικά με τον τρόπο ενσωμάτωσης αυτού του τύπου δεδομένων σε χρήστες. Για παράδειγμα, ορισμένες στρατηγικές προσπάθησαν να περιλαμβάνουν την έλλειψη εμπιστοσύνης ως φίλτρο επιλογής χρηστών.

Η μελέτη που παρουσιάστηκε στο (Ma et al 2009), έδειξε μια ισχυρή θετική σχέση μεταξύ εμπιστοσύνης και ενδιαφέροντος για τις ομοιότητες μεταξύ των χαρακτηριστικών των χρηστών για ομοιότητα μεταξύ χρηστών. Ωστόσο, για να διευρευνηθεί η χρήση εμπιστοσύνης σε RS

4. και προσομοίωση πραγματικών αλληλεπιδράσεων, πρέπει να εφαρμοστούν νέες μέθοδοι για τον υπολογισμό
5. έμμεσες αξίες εμπιστοσύνης. Για το σκοπό αυτό, η πρακτική εμπιστοσύνη πρέπει να αποκτηθεί σιωπηρά και από
6. άτομα που έχουν σημαντικές σχέσεις με τους χρήστες.



## 5. Συμπεράσματα

Σε αυτή τη πτυχιακή εργασία περιγράψαμε τις κύριες ιδέες που υπογραμμίζουν και υποστηρίζουν τις προτεινόμενες ερευνητικές προσεγγίσεις στον τομέα των συστημάτων δημιουργίας προτάσεων για χρήστες υπό το πρίσμα των μέσων κοινωνικής δικτύωσης. Παρουσιάσαμε το θεωρητικό υπόβαθρο αυτών των συστημάτων και τις σύγχρονες προκλήσεις και τα τρέχοντα σενάρια που εισάγουν ένα νέο ερευνητικό πλαίσιο με επίκεντρο την αλληλεπίδραση των ανθρώπων στα μέσα κοινωνικής δικτύωσης.

Στη συνέχεια περιγράψαμε τα γενικά χαρακτηριστικά των διάφορων προσεγγίσεων περιγράφοντας τα κύρια χαρακτηριστικά τους, τα αποτελέσματα των ερευνών τους γύρω από τις προσπάθειες μοντελοποίησης εννοιών όπως η εμπιστοσύνη, τα συναισθήματα και η ακρίβεια στις αξιολογήσεις των χρηστών (πχ καταναλωτών). Οι ιδέες που χρησιμοποιούνται για τη δημιουργία κάθε συνιστώσας στο ερευνητικό πλαίσιο εντοπίστηκαν και παρουσιάστηκαν.

Σχετικά με το στοιχείο εμπιστοσύνης στα διάφορα προτεινόμενα συστήματα παρατηρήθηκε η συσχέτιση με τους παρακάτω παράγοντες:

- Απαιτεί συλλογή ενός συνόλου δεδομένων σχετικά με τις συμπεριφορές που αναδεικνύονται μέσα από την επικοινωνία μεταξύ των ενεργών χρηστών και των φίλων τους.
- Απαιτεί ανάλυση μοτίβων συμπεριφορών στα κοινωνικά δίκτυα που να υποδεικνύουν εμπιστοσύνη.
- Ο ορισμός της κατάλληλης μεθόδου για τη διαμόρφωση μίας μετρικής εμπιστοσύνης είναι ακόμη σε ερευνητικό στάδιο.
- Η μελέτη του κατά πόσο αυτή η μετρική εμπιστοσύνης μπορεί να συμβάλει στην τελική πρόβλεψη των προτάσεων και των αξιολογήσεων θα μπορούσε να είναι η συνέχεια αυτής της πτυχιακής εργασίας.

Πράγματι, υπάρχει σε εξέλιξη μια εκτεταμένη έρευνα που επιλύει το πρόβλημα της επιλογής χαρακτηριστικών με βάση λύσεις μηχανικής μάθησης ειδικά στην ανάλυση συναισθημάτων. Αυτό γίνεται με διάφορες τεχνικές όπως:

- Εμφάνιση λέξεων κλειδιών έναντι συχνότητας: Η συχνότητα ενός όρου έχει τονιστεί κλασικά στην τυπική ανάκτηση πληροφοριών, όπως η δημοφιλής τεχνική στάθμιση TF-IDF, η οποία παρουσιάζει λογικά αποτελέσματα. Ωστόσο, διάφορες έρευνες έχουν παρουσιάσει καλύτερη απόδοση χρησιμοποιώντας την εμφάνιση ενός όρου σε σύγκριση με τη συχνότητα του, όπου οι τιμές εισαγωγής αυξάνονται με τη συχνότητα εμφάνισης τέτοιων λέξεων – κλειδιών. Αυτό σημαίνει ότι τα διανύσματα χαρακτηριστικών με δυαδική αξία στα οποία η είσοδος υποδεικνύει εάν μια λέξη-κλειδί εμφανίζεται (τιμή 1) ή δεν εμφανίζεται (τιμή 0) έδειξαν μεγαλύτερη αποτελεσματικότητα στον τομέα της ταξινόμησης πολικότητας κριτικής από ότι τα διανύσματα δυνατοτήτων πραγματικής αξίας. Αυτό το εύρημα ρίχνει φως στη διαφορά μεταξύ κατηγοριοποίησης εγγράφων βάσει θέματος και ταξινόμησης πολικότητας. Ο προσδιορισμός ενός θέματος είναι πιο πιθανό να επηρεαστεί από τις

συχνές εμφανίσεις ορισμένων όρων. Ωστόσο, ενδέχεται το συνολικό συναίσθημα να μην ανιχνεύεται συχνά μέσω του αριθμού των εμφανίσεων των ίδιων όρων.

- **Θέσεις όρων:** Η θέση ενός διακριτικού όρου μέσα σε ένα έγγραφο μπορεί να επηρεάσει το συνολικό συναίσθημα. Ως εκ τούτου, οι πληροφορίες σχετικά με τη θέση εμφάνισης μερικές φορές κωδικοποιούνται στα διανύσματα χαρακτηριστικών. Μέσα στη συζήτηση σχετικά με τη χρησιμότητα των n-gram υψηλότερης τάξης, έχει διαπιστωθεί ότι σε ορισμένες περιπτώσεις, τα bigrams και τα trigrams αποδίδουν καλύτερα στην ταξινόμηση πολικότητας της αναθεώρησης προϊόντων. Παρόλα αυτά, έχει ανακαλυφθεί ότι τα unigrams ξεπερνούν τα bigrams στην ταξινόμηση των κριτικών ταινιών, ως προς την πολικότητα του συναισθήματος.
- **Μέρος του λόγου (POS, Part Of Speech):** Πρόκειται για έναν τύπο πληροφοριών που μπορούν να είναι χρήσιμες στον τομέα της ανάλυσης συναισθημάτων και της εξόρυξης απόψεων. Τα επίθετα έχουν χρησιμοποιηθεί ως χαρακτηριστικό από αρκετούς ερευνητές. Για παράδειγμα, μελέτες για τον προσδιορισμό της υποκειμενικότητας έδειξαν μια ισχυρή συσχέτιση μεταξύ της παρουσίας συγκεκριμένων επίθετων και της υποκειμενικότητας των προτάσεων. Αυτό το αποτέλεσμα εξήγησε ότι ορισμένα επίθετα είναι αποτελεσματικά στον καθορισμό του συναισθήματος. Άλλες εργασίες συνέκριναν την απόδοση των επίθετων, των ρημάτων και των επιρρημάτων όπου η υπό-κατηγοριοποίηση μπορεί να διαδραματίσει σημαντικό ρόλο.
- **Αρνητικότητα:** Η εξέταση της άρνησης αποτελεί μείζον μέλημα της ανάλυσης γνώμης και συναισθημάτων. Είναι σαφές ότι τα διακριτικά άρνησης όπως το όχι, όταν προστίθενται σε μια φράση μπορούν να αλλάξουν το συνολικό συναίσθημα στην αντίθετη τάξη. Ορισμένες εργασίες κωδικοποίησαν απευθείας τις αρνήσεις στα διανύσματα χαρακτηριστικών. Για παράδειγμα, στην πρόταση «δεν μου αρέσει να μαγειρεύω», το λεκτικό «αρέσει» θα αλλάξει σε νέο λεκτικό «αρέσει-όχι». Ένα άλλο πρόβλημα που μπορεί να εμφανιστεί στην μοντελοποίηση της άρνησης, είναι ότι μερικές φορές η άρνηση εκφράζεται σε ένα πλαίσιο ειρωνείας και σαρκασμού. Αυτές οι εκφράσεις δεν είναι εύκολο να εντοπιστούν.

Σε μία επόμενη μελλοντική έρευνα θα πρέπει να εξετάσουμε επίσης την χρήση γλωσσικών χαρακτηριστικών μέσα σε διαλόγους μεταξύ χρηστών στα κοινωνικά δίκτυα, όπως τα emoticon σύμβολα. Θεωρούμε ότι περιέχουν μια πολύ σαφή εικόνα σχετικά με τις απόψεις των χρηστών στο κείμενο. Ενδεχομένως με αυτόν τον τρόπο θα μπορεί ένα σύστημα να αναπαραστήσει τις απόψεις των χρηστών με μεγαλύτερη ακρίβεια. Θα μπορούν να μετρήσουν τον βαθμό του συναισθήματος και των απόψεων με κλίμακα βαθμολογίας πολλαπλών σημείων. Η αξιολόγηση των συναισθημάτων είναι μία σημαντική πρόκληση για τη συνέχιση της έρευνας σε αυτό το πεδίο θέτοντας στόχους όπως οι παρακάτω:

- Κατά πόσο η χρήση μικρό-ιστολογίων είναι χρήσιμη όπως και η συλλογή κατάλληλων τιτλισμάτων σχετικά με την εξαγωγή συμπερασμάτων για τα συναισθήματα των χρηστών και τη προσαρμογή των προτάσεων του συστήματος.
- Εξαγωγή χαρακτηριστικών όπως εμπιστοσύνη και συναίσθημα μέσα από σύντομες δημοσιεύσεις.
- Εφαρμογή μια κατάλληλης μεθόδου που να μετατρέπει αυτές τις δημοσιεύσεις σε βαθμίδες βαθμολογίας και η οποία θα αντιπροσωπεύει τις δημοσιεύσεις των φίλων με τη μορφή αριθμητικής αξιολόγησης.
- Χρησιμοποίηση αυτή της συμπερασματικής αριθμητικής αξιολόγησης για την υποστήριξη της τελικής απόφασης λαμβάνοντας υπόψη ότι για νέους χρήστες θα υπάρχουν ελλιπή στοιχεία στο προφίλ τους ενώ για παλιούς χρήστες θα υπάρχει ένα πλούσιο ιστορικό με αξιολογήσεις σε προϊόντα, αντικείμενα και υπηρεσίες.

## Βιβλιογραφία

- Felfernig A., Jeran M., Ninaus G., Reinfrank F., Reiterer S. (2013) Toward the Next Generation of Recommender Systems: Applications and Research Challenges. In: Tsihrintzis G., Virvou M., Jain L. (eds) Multimedia Services in Intelligent Environments. Smart Innovation, Systems and Technologies, vol 24. Springer, Heidelberg.  
[https://doi.org/10.1007/978-3-319-00372-6\\_5](https://doi.org/10.1007/978-3-319-00372-6_5)
- Grasso, A., Convertino, G. (2012) Collective Intelligence in Organizations: Tools and Studies. *Comput Supported Coop Work* 21, 357–369 (2012).  
<https://doi.org/10.1007/s10606-012-9165-3>
- Dimah H. Alahmadi, Xiao-Jun Zeng (2015) ISTS: Implicit social trust and sentiment based approach to recommender systems, *Expert Systems with Applications*, Volume 42, Issue 22, 2015, Pages 8840-8849, ISSN 0957-4174,  
<https://doi.org/10.1016/j.eswa.2015.07.036>.
- Devashish Shankar and Sujay Narumanchi and H A Ananya and Pramod Kompalli and Krishnendu Chaudhury (2017) Deep Learning based Large Scale Visual Recommendation and Search for E-Commerce, arXiv, 1703.02344
- P. Symeonidis, A. Nanopoulos, and Y. Manolopoulos (2008), “Providing justifications in recommender systems,” *Systems, Man and Cybernetics, Part A: Systems and Humans*, IEEE Transactions on, vol. 38, no. 6, pp. 1262–1272, 2008.
- Lu, Z., Dou, Z., Lian, J., Xie, X., & Yang, Q. (2015). Content-Based Collaborative Filtering for News Topic Recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(1). Retrieved from  
<https://ojs.aaai.org/index.php/AAAI/article/view/9183>
- X. Yang, H. Steck, et Y. Liu (2012) Circle-based recommendation in online social networks , in *Proceedings of the 18<sup>th</sup> ACM SIGKDD international conference on Knowledge discovery and data mining*, 2012, p. 1267–1275.
- H. Ma, H. Yang, M. R. Lyu, et I. King (2008), Sorec: social recommendation using probabilistic matrix factorization, in *Proceedings of the 17th ACM conference on Information and knowledge management*, 2008, p. 931–940.
- H. Ma, I. King, et M. R. Lyu (2009), “Learning to recommend with social trust ensemble “, in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, 2009, p. 203–210.
- P. Cofta (2006), “Distrust,” in *Proceedings of the 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet*,

pp. 250–258, ACM, 2006.

J. H. Kietzmann, K. Hermkens, I. P. McCarthy, et B. S. Silvestre (2011), « Social media? Get serious! Understanding the functional building blocks of social media », *Bus. Horiz.*, vol. 54, no 3, p. 241–251, 2011.

D. Zejda (2010), “From subjective trust to objective trustworthiness in on-line social networks: Overview and challenges,” *Journal of Systems Integration*, vol. 1, no. 1-2, pp. 16–22, 2010.

Golbeck, J., & Hendler, J. (2006). FilmTrust: movie recommendations using trust in web-based social networks. *CCNC 2006. 2006 3rd IEEE Consumer Communications and Networking Conference, 2006.*, 1, 282-286.

K. Miyahara and M. J. Pazzani (2002), “Improvement of collaborative filtering with the simple bayesian classifier,” *IPSI Journal*, vol. 43, no. 11, pp. 3429–3437, 2002.

L. Page, S. Brin, R. Motwani, and T. Winograd (1999), “The pagerank citation ranking: bringing order to the web.,” 1999.

Melguizo, M.C. & Boves, Lou & Ramos, Olga. (2009). A Proactive Recommendation System for Writing: Helping without Disrupting. *International Journal of Industrial Ergonomics*. 39. 516-523. 10.1016/j.ergon.2008.10.004.

Rennie, Jason & Srebro, Nathan. (2005). Fast maximum margin matrix factorization for collaborative prediction. 713-719. 10.1145/1102351.1102441.

P. Victor, M. De Cock, and C. Cornelis (2011), “Trust and recommendations,” in *Recommender Systems Handbook*, pp. 645–675, Springer, 2011.

S. Vucetic and Z. Obradovic (2005), “Collaborative filtering using a regression-based approach,” *Knowledge and Information Systems*, vol. 7, no. 1, pp. 1–22, 2005

R. M. Bell and Y. Koren (2007), “Lessons from the netflix prize challenge,” *ACM SIGKDD Explorations Newsletter*, vol. 9, no. 2, pp. 75–79, 2007.

John S. Breese, David Heckerman, Carl Kadie (2013) *Empirical Analysis of Predictive Algorithms for Collaborative Filtering*, arXiv:1301.736

GG Chowdhury (2010). *Introduction to modern information retrieval*. Facet publishing.

Yue Liu; Hua Yang; Gengxin Sun; Sheng Bin (2020) Collaborative Filtering Recommendation Algorithm Based on Multi-relationship Social Network. *Ingénierie des Systèmes d'Information* . Jun2020, Vol. 25 Issue 3, p359-364. 6p.

G.-R. Xue, C. Lin, Q. Yang, W. Xi, H.-J. Zeng, Y. Yu, and Z. Chen (2005), “Scalable collaborative filtering using cluster-based smoothing,” in *Proceedings of the 28th*

annual international ACM SIGIR conference on Research and development in information retrieval, pp. 114–121, ACM, 2005.

Schafer J.B., Frankowski D., Herlocker J., Sen S. (2007) Collaborative Filtering Recommender Systems. In: Brusilovsky P., Kobsa A., Nejdl W. (eds) *The Adaptive Web. Lecture Notes in Computer Science*, vol 4321. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-72079-9\\_9](https://doi.org/10.1007/978-3-540-72079-9_9)

J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl (2004), “Evaluating collaborative filtering recommender systems,” *ACM Transactions on Information Systems (TOIS)*, vol. 22, no. 1, pp. 5–53, 2004.

Pelin Atahan, Sumit Sarkar (2010) Accelerated Learning of User Profiles. *Management Science* 57 (2) 215-239 <https://doi.org/10.1287/mnsc.1100.1266>

H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King (2011), “Recommender systems with social regularization,” in *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 287–296, ACM, 2011.

J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, *Recommender systems survey, Knowledge-Based Systems, Volume 46, 2013, Pages 109-132, ISSN 0950-7051, <https://doi.org/10.1016/j.knosys.2013.03.012>*.

Jesus Bobadilla, Antonio Hernando, Fernando Ortega, Jesus Bernal (2011) A framework for collaborative filtering recommender systems, *Expert Systems with Applications, Volume 38, Issue 12, 2011, Pages 14609-14623, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2011.05.021>*.

H. Zou, X. Tang, B. Xie and B. Liu (2015), "Sentiment Classification Using Machine Learning Techniques with Syntax Features," *2015 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2015, pp. 175-179, doi: 10.1109/CSCI.2015.44.

Y. Kim and R. Phalak (2012) “A trust prediction framework in rating-based experience sharing social networks without a web of trust,” *Information Sciences*, 2012.

C. Aggarwal (2016). *Recommender Systems: The Textbook*. Springer Publishing Company, Incorporated, 1st ed., 2016.

I. Guy (2015). *Social Recommender Systems*, pp. 511–543. Boston, MA: Springer US, 2015.

M. Eirinaki, N. Moniz, and K. Potika (2016), “Threshold-bounded influence dominating sets for recommendations in social networks,” in *2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking*

(SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), pp. 408–415, Oct 2016.

S. Lei, S. Maniu, L. Mo, R. Cheng, and P. Senellart (2015), “Online influence maximization,” in Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 645–654, ACM, 2015.

J. Weng, E.-P. Lim, J. Jiang, and Q. He (2010), “Twitterrank: Finding topic-sensitive influential twitterers,” in Proc. of WSDM’10, pp. 261–270, 2010.

D. A. Kim, A. R. Hwang, D. Stafford, D. A. Hughes, A. J. O’Malley, J. H. Fowler, and N. A. Christakis (2015), “Social network targeting to maximise population behaviour change: a cluster randomised controlled trial,” *The Lancet*, vol. 386, pp. 145–153, 5 2015.

M.-E. G. Rossi, F. D. Malliaros, and M. Vazirgiannis (2015), “Spread it good, spread it fast: Identification of influential nodes in social networks,” in Proc. of WWW’15 Companion, pp. 101–102, ACM, 2015.

R. Huckfeldt and J. Sprague (1987), “Networks in context: The social flow of political information,” *American Political Science Review*, vol. 81, pp. 1197–1216, 12 1987.

E. Dubois and D. Gaffney (2014), “The multiple facets of influence: Identifying political influentials and opinion leaders on twitter,” *American Behavioral Scientist*, vol. 58, no. 10, pp. 1260–1277, 2014.

Y. Yang, X. Mao, J. Pei, and X. He (2016), “Continuous influence maximization: What discounts should we offer to social network users?,” in Proceedings of the 2016 International Conference on Management of Data, SIGMOD ’16, (New York, NY, USA), pp. 727–741, ACM, 2016.

M. Kaple, K. Kulkarni, and K. Potika (2017), “Viral marketing for smart cities: Influencers in social network communities,” in 9th IEEE International Workshop on Big Data Applications in Smart City Development, 2017.

P. Domingos and M. Richardson (2001), “Mining the network value of customers,” in Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 57–66, ACM, 2001.

A. S. Badashian and E. Stroulia (2016), “Measuring user influence in GitHub: The million follower fallacy,” in 2016 IEEE/ACM 3rd International Workshop on CrowdSourcing in Software Engineering (CSI-SE), pp. 15–21, May 2016.

Q. Jiang, G. Song, G. Cong, Y. Wang, W. Si, and K. Xie (2011), “Simulated annealing based influence maximization in social networks,” in *AAAI*, vol. 11, pp. 127–132, 2011.

D. Kempe, J. Kleinberg, and Eva Tardos (2015), “Maximizing the spread of influence through a social network,” *Theory of Computing*, vol. 11, no. 4, pp. 105–147, 2015.

- S. Hangal, D. Maclean, M. S. Lam, and J. Heer (2010), "All friends are not equal: Using weights in social graphs to improve search."
- J. M. C. Miller McPherson, Lynn Smith-Lovin (2001), "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, vol. 27, pp. 415–444, 2001.
- J. Tang, X. Hu, and H. Liu (2013). "Social recommendation: a review.," *Social Network Analysis Mining*, vol. 3, no. 4, pp. 1113–1133, 2013.
- J. Tang, H. Gao, X. Hu, and H. Liu (2013), "Exploiting homophily effect for trust prediction," in *Proc. of WSDM '13*, pp. 53–62, ACM, 2013.
- H. Ma, I. King, and M. R. Lyu (2009), "Learning to recommend with social trust ensemble," in *Proc. of SIGIR '09*, pp. 203–210, ACM, 2009.
- J. Tang, H. Gao, and H. Liu (2012). "mtrust: Discerning multi-faceted trust in a connected world," in *Proc. of WSDM '12*, pp. 93–102, ACM, 2012.
- M. Jamali and M. Ester (2010). "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proc. of RecSys '10*, pp. 135–142, ACM, 2010.
- H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King (2011), "Recommender systems with social regularization," in *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 287–296, ACM, 2011.
- P. Massa and P. Avesani (2005), "Controversial users demand local trust metrics: An experimental study on epinions.com community," in *Proc. of AAI-05*, 2005.
- I. Varlamis, M. Eirinaki, and M. Louta (2013). "Application of social network metrics to a trust-aware collaborative model for generating personalized user recommendations," in *The Influence of Technology on Social Network Analysis and Mining*, pp. 49–74, Springer Vienna, 2013.
- S. Deng, L. Huang, G. Xu, X. Wu, and Z. Wu (2017). "On deep learning for trust-aware recommendations in social networks," *IEEE transactions on neural networks and learning systems*, vol. 28, no. 5, pp. 1164–1177, 2017.
- L. Page, S. Brin, R. Motwani, and T. Winograd (1999). "The pagerank citation ranking: Bringing order to the web.," *Technical Report 1999-66*, Stanford InfoLab, November 1999.
- Q. Liu, B. Xiang, N. J. Yuan, E. Chen, H. Xiong, Y. Zheng, and Y. Yang (2017). "An influence propagation view of pagerank," *ACM Trans. Knowl. Discov. Data*, vol. 11, pp. 30:1–30:30, Mar. 2017.



- G. Adomavicius and A. Tuzhilin (2005), "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 17, no. 6, pp. 734–749, 2005.
- F. Ricci, L. Rokach, and B. Shapira (2011), "Introduction to recommender systems handbook," *Recommender Systems Handbook*, pp. 1–35, 2011.
- P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl (1994), "Grouplens: an open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp. 175–186, ACM, 1994.
- B. Sarwar, G. Karypis, J. Konstan, and J. Riedl (2001), "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, pp. 285–295, ACM, 2001.
- M. Pazzani and D. Billsus (1997), "Learning and revising user profiles: The identification of interesting web sites," *Machine learning*, vol. 27, no. 3, pp. 313–331, 1997.
- L. Si and R. Jin (2003), "Flexible mixture model for collaborative filtering," in *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*, vol. 20, p. 704, 2003.
- H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King (2011), "Recommender systems with social regularization," in *Proceedings of the fourth ACM international conference on Web search and data mining*, pp. 287–296, ACM, 2011.
- Y. Kim and R. Phalak (2012), "A trust prediction framework in rating-based experience sharing social networks without a web of trust," *Information Sciences*, 2012.
- C. Birtolo and D. Ronca (2013), "Advances in clustering collaborative filtering by means of fuzzy c-means and trust," *Expert Systems with Applications*, vol. 40, no. 17, pp. 6997–7009, 2013.
- B. Pang, L. Lee, and S. Vaithyanathan (2002), "Thumbs up?: sentiment classification using machine learning techniques," in *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pp. 79–86, Association for Computational Linguistics, 2002.
- G. Ganu, N. Elhadad, and A. Marian (2009), "Beyond the stars: Improving rating predictions using review text content," in *12th International Workshop on the Web and Databases*, 2009.
- C. W.-K. Leung, S. C.-F. Chan, F.-L. Chung, and G. Ngai (2011), "A probabilistic rating inference framework for mining user preferences from reviews," *World Wide Web*, vol. 14, no. 2, pp. 187–215, 2011.
- Gulati, Avni & Eirinaki, Magdalini. (2018). Influence Propagation for Social Graph-based Recommendations. 2180-2189. 10.1109/BigData.2018.8622213.