

**ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ**  
Εργ. Συστημάτων Βάσεων Γνώσεων & Δεδομένων

**ΠΑΡΑΜΕΤΡΟΠΟΙΗΣΙΜΟ ΣΥΣΤΗΜΑ ΠΡΟΤΑΣΕΩΝ (RECOMMENDER SYSTEM)**

*ΠΛΗΡΟΦΟΡΙΕΣ:* Γιώργος Γιαννόπουλος, [giann@dblab.ece.ntua.gr](mailto:giann@dblab.ece.ntua.gr), Θανάσης Βεργούλης, [vergoulis@dblab.ece.ntua.g](mailto:vergoulis@dblab.ece.ntua.g), Παναγιώτης Μπούρος, [pbour@dblab.ece.ntua.gr](mailto:pbour@dblab.ece.ntua.gr)

**ΠΕΡΙΛΗΨΗ:** Στόχος της διπλωματικής εργασίας είναι η ανάπτυξη ενός πλήρως παραμετροποιήσιμου και προσαρμόσιμου συστήματος προτάσεων το οποίο θα παρέχει τη δυνατότητα δημιουργίας/τροποποίησης/διαγραφής των χαρακτηριστικών των προτεινομένων αντικειμένων καθώς και προσαρμογή των αλγορίθμων παραγωγής προτάσεων, ανάλογα με το είδος των αντικειμένων.

**ΑΤΟΜΑ:** 1-2

**ΠΛΑΤΦΟΡΜΑ ΕΡΓΑΣΙΑΣ:** Java, PHP, Javascript, MySQL.

**ΣΥΝΤΟΜΗ ΠΕΡΙΓΡΑΦΗ:** Τα συστήματα προτάσεων (Recommender Systems) χρησιμοποιούνται για να προβλέπουν ποια προϊόντα/υπηρεσίες/πληροφορίες ενδιαφέρουν τους χρήστες ενός πρωτεύοντος συστήματος και να τους τα προτείνουν.

Χαρακτηριστικά παραδείγματα συστημάτων προτάσεων αποτελούν οι ιστοσελίδες του Amazon (<http://www.amazon.com/>) και του imdb (<http://www.imdb.com/>). Η πρώτη είναι μία ιστοσελίδα για πώληση αγαθών, ενώ η δεύτερη μία δικτυακή βάση ταινιών. Στο μεν Amazon, όταν ο χρήστης επιλέγει ένα συγκεκριμένο προϊόν, κάτω από αυτό εμφανίζονται πεδία με προτάσεις της μορφής:

- *Πελάτες που αγόρασαν αυτό το προϊόν, αγόρασαν επίσης...*
- *Πελάτες που είδαν αυτό το προϊόν, αγόρασαν επίσης...*
- *Προϊόντα σχετικά με αυτό το προϊόν...*
- *Σχετικά προϊόντα από άλλες ιστοσελίδες...*

Στο imdb, όταν ο χρήστης κοιτάει πληροφορίες για μία ταινία, μπορεί να δει προτάσεις της μορφής:

- *Αν σας άρεσε αυτή η ταινία, η βάση μας σας προτείνει...*

Η αποτελεσματικότητα ενός συστήματος προτάσεων εξαρτάται από διάφορους παράγοντες:

- **Τρόπος αξιολόγησης αντικειμένων από τους χρήστες.** Η αξιολόγηση ενός αντικειμένου μπορεί να γίνεται με άμεσο τρόπο, δηλαδή να ζητείται από το χρήστη να το αξιολογήσει σε μία κλίμακα (π.χ. από 1-10), ή έμμεσα, θεωρώντας ότι ο χρήστης ενδιαφέρεται για ένα αντικείμενο αν επέλεξε να το δει. Ο πρώτος τρόπος δίνει πιο αξιόπιστες κρίσεις για τα αντικείμενα σε σχέση με τον δεύτερο, αλλά απαιτεί περισσότερο κόπο, κάτι που αποτελεί ανασταλτικό παράγοντα για το χρήστη.
- **Χαρακτηριστικά των αντικειμένων.** Ανάλογα με τις εφαρμογές στις οποίες χρησιμοποιείται ένα σύστημα προτάσεων, ενδέχεται να υπάρχουν διαφόρων ειδών χαρακτηριστικά που περιγράφουν τα αντικείμενα και χρησιμεύουν στην εξαγωγή των προτάσεων. Για παράδειγμα, όταν τα αντικείμενα μας είναι ταινίες, τα χαρακτηριστικά μπορεί να είναι οι ηθοποιοί, ο σκηνοθέτης, το είδος της ταινίας, ο προϋπολογισμός της κ.α. Αντίστοιχα, όταν τα αντικείμενα είναι κινητά τηλέφωνα, τα χαρακτηριστικά τους είναι η οθόνη, η κάμερα, η μπαταρία κ.α.
- **Οι αλγόριθμοι προτάσεων** οι οποίοι κατηγοριοποιούνται συνοπτικά ως εξής:
  - Ομοιότητα των προς πρόταση αντικειμένων με παλαιότερα αντικείμενα για τα οποία έχει ενδιαφερθεί ή αξιολογήσει ο χρήστης (content-based recommender systems)

**ΘΕΜΑΤΑ ΔΙΠΛΩΜΑΤΙΚΩΝ ΕΡΓΑΣΙΩΝ**  
Εργ. Συστημάτων Βάσεων Γνώσεων & Δεδομένων

- Ομοιότητα των ενδιαφερόντων ή αξιολογήσεων μεταξύ μίας ομάδας χρηστών (collaborative filtering recommender systems)
- Ιστορικό αναζητήσεων/αξιολογήσεων του χρήστη (personalized recommender systems)
- Συνδυασμός των παραπάνω μεθόδων (hybrid recommender systems)

Σκοπός της διπλωματικής είναι η σχεδίαση και υλοποίηση ενός συστήματος προτάσεων το οποίο θα δίνει στο διαχειριστή του τη δυνατότητα προσαρμογής των παραπάνω παραγόντων, έτσι ώστε να καλύπτονται όσο το δυνατόν περισσότερες περιπτώσεις προτεινόμενων αντικειμένων. Συγκεκριμένα, θα προσφέρει διαφορετικούς τρόπους αξιολόγησης των αντικειμένων, με επιπλέον δυνατότητα σταθμισμένου συνδυασμού των τρόπων αξιολόγησης. Επίσης, θα δίνει τη δυνατότητα κατασκευής διαφόρων ειδών χαρακτηριστικών για περιγραφή των αντικειμένων (αριθμητικές τιμές, συμβολοσειρές από ένα προκαθορισμένο πεδίο τιμών, ελεύθερο κείμενο κ.α.) και ρύθμισης της βαρύτητας κάθε χαρακτηριστικού στη διαδικασία παραγωγής προτάσεων. Τέλος, θα δίνει τη δυνατότητα επιλογής του αλγορίθμου (ή σταθμισμένου συνδυασμού αλγορίθμων) παραγωγής προτάσεων, ανάλογα με τη φύση της εκάστοτε εφαρμογής.

Στα πλαίσια της διπλωματικής, θα πραγματοποιηθούν οι παρακάτω εργασίες:

- Μελέτη της βιβλιογραφίας και εντοπισμός των πιο επίκαιρων προσεγγίσεων και αλγορίθμων για συστήματα προτάσεων.
- Αναζήτηση ανοιχτού κώδικα υλοποιήσεων συστημάτων προτάσεων στις οποίες θα βασιστεί η υλοποίηση των αλγορίθμων προτάσεων της εργασίας.
- Υλοποίηση τους συστήματος παραμετροποιημένης εισαγωγής τρόπου αξιολόγησης, παραμετροποιημένης κατασκευής χαρακτηριστικών αντικειμένων και παραμετροποιημένης επιλογής αλγορίθμων προτάσεων.
- Πειραματική αξιολόγηση του συστήματος σε μία ή περισσότερες πραγματικές εφαρμογές.

**ΣΧΕΤΙΚΕΣ ΠΛΗΡΟΦΟΡΙΕΣ:**

[http://www.readwriteweb.com/archives/recommender\\_systems.php](http://www.readwriteweb.com/archives/recommender_systems.php)

<http://www.deitel.com/ResourceCenters/Web20/RecommenderSystems/tabid/1229/Default.aspx>