

## Το πρόβλημα

- Πώς μπορούμε να αξιολογήσουμε τους φοιτητές σε ανοικτά μαθήματα μεγάλης συμμετοχής;
- Το πλήθος των φοιτητών είναι τεράστιο, ενώ ο αριθμός των διαθέσιμων αξιολογητών είναι περιορισμένος
- **Προτεινόμενη λύση:** οι ίδιοι οι φοιτητές αναλαμβάνουν να βαθμολογήσουν ένα μικρό πλήθος εργασιών από άλλους φοιτητές, κατατάσσοντάς τις από την καλύτερη προς την χειρότερη
- Έπειτα, αυτές οι επί μέρους κατατάξεις **συναθροίζονται** σε μία τελική κατάταξη που αντιπροσωπεύει τη σχετική απόδοση των φοιτητών
- **Στόχος:** να μάθουμε όσο το δυνατόν μεγαλύτερο ποσοστό των συγκρίσεων μεταξύ όλων των ζευγαριών εργασιών

## Διαμοιρασμός εργασιών

- Κάθε φοιτητής αναλαμβάνει να αξιολογήσει  $k$  εργασίες
- Κάθε εργασία ανατίθεται σε  $k$  φοιτητές
- Περιορισμός: κανένας φοιτητής δεν αναλαμβάνει τη δική του εργασία
- Υλοποίηση: τυχαίος διαμοιρασμός

## Σενάρια βαθμολόγησης

### Τέλεια βαθμολόγηση

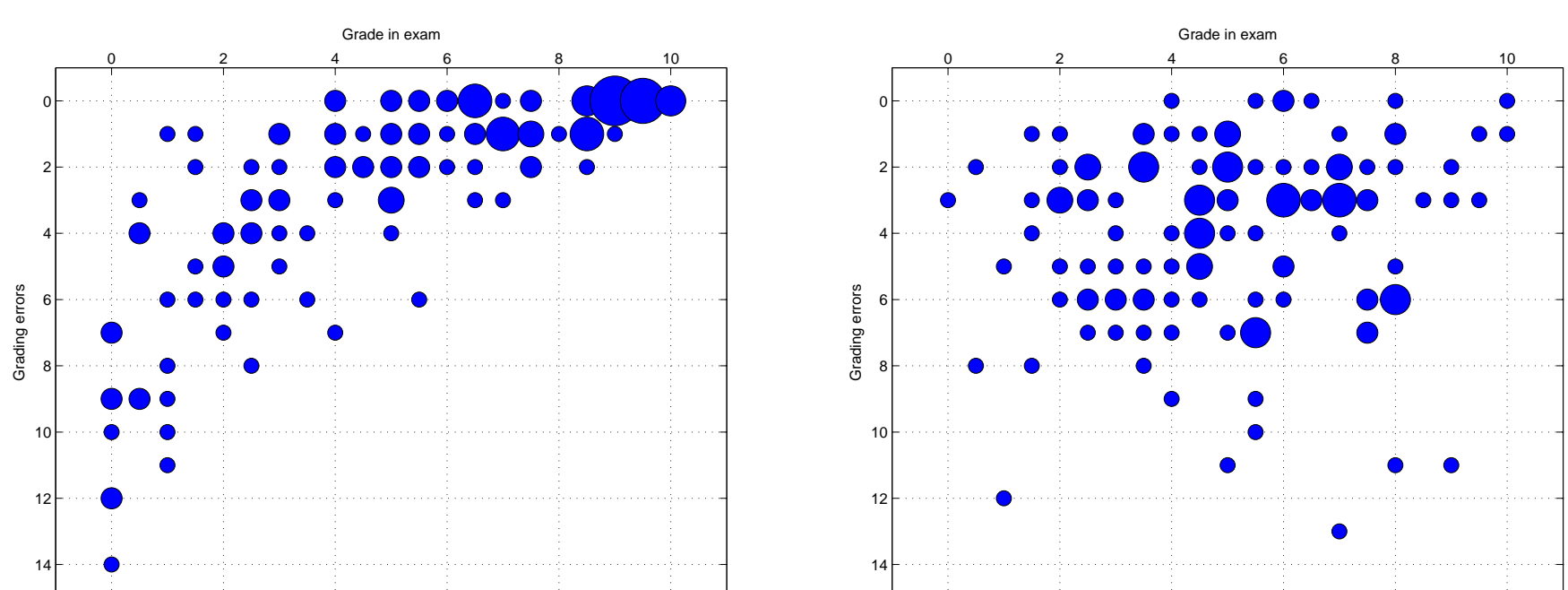
- Υποθέτουμε ότι οι φοιτητές κατατάσσουν σωστά τις εργασίες που αναλαμβάνουν

### Ατελής βαθμολόγηση (Mallows)

- Κάθε φοιτητής  $i$  έχει μια τυχαία ποιότητα  $q_i$
- Η υποκειμενική αλήθεια  $\succ$  προκύπτει διατάσσοντας τις εργασίες σε φθίνουσα σειρά με βάση τις ποιότητες των φοιτητών
- Ο φοιτητής  $i$  κατατάσσει τις εργασίες που έχει αναλάβει ως εξής:
  - Για κάθε ζευγάρι εργασιών  $x, y$  έτσι ώστε  $x \succ y$ , με πιθανότητα  $q_i$  θέτω  $x \succ_i y$
  - Αν δημιουργηθεί κύκλος, επαναλαμβάνω από την αρχή

### Πραγματική βαθμολόγηση

- Χρησιμοποιούμε δεδομένα που έχουμε συλλέξει μέσω ενός πειράματος πεδίου με πραγματικούς φοιτητές



## Πληροφορίες

▷ I. Caragiannis, G. A. Krimpas, and A. A. Voudouris. Aggregating partial rankings with applications to peer grading in massive online open courses. *AAMAS* 2015.

▷ I. Caragiannis, G. A. Krimpas, and A. A. Voudouris. How effective can simple ordinal peer grading be? *arXiv:1602.07985*

## Ο κανόνας Borda

### Ορισμός

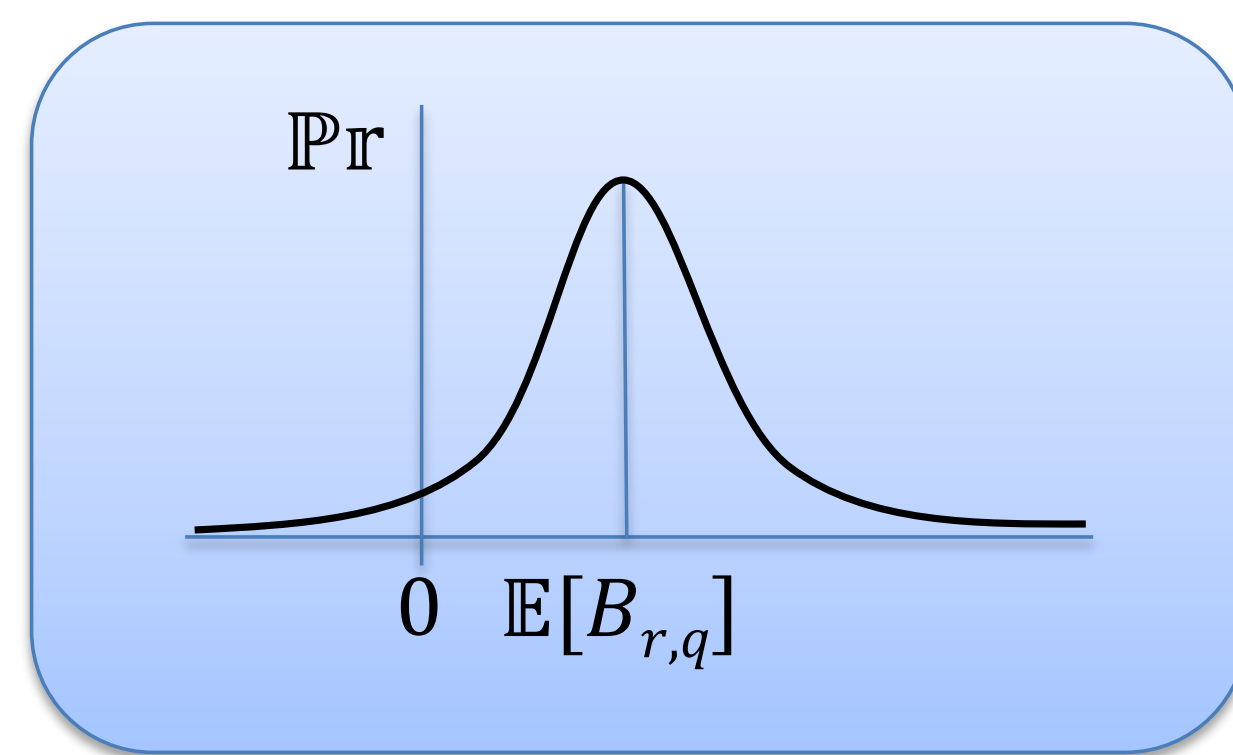
- Κάθε εργασία παίρνει  $k$  βαθμούς για κάθε πρώτη θέση,  $k - 1$  βαθμούς για κάθε δεύτερη θέση, κ.ο.κ.
- Η τελική κατάταξη προκύπτει διατάσσοντας τις εργασίες κατά φθίνουσα σειρά ως προς τους συνολικούς βαθμούς τους

### Θεώρημα

Στο μοντέλο τέλει βαθμολόγησης, η αναμενόμενη απόδοση του κανόνα Borda είναι τουλάχιστον  $1 - \mathcal{O}(1/k)$ .

### Αποδεικτική ιδέα

- Εργασίες με πραγματικές κατατάξεις  $r < q$
- $B_{r,q}$  = διαφορά βαθμών Borda των  $r$  και  $q$
- $\mathbb{E}[B_{r,q}]$  = ανάλογη της διαφοράς  $q - r$
- Ισχυρή συγκέντρωση των τιμών της  $B_{r,q}$  γύρω από τη μέση τιμή  $\mathbb{E}[B_{r,q}]$
- $\Pr[\text{η σχέση των } r \text{ και } q \text{ ανακτάται σωστά}] = \Pr[B_{r,q} > 0]$



- Αθροίζουμε για όλα τα ζευγάρια εργασιών

### Άλλες ιδιότητες

- Βέλτιστος κανόνας κατάταξης τύπων στο σενάριο τέλει βαθμολόγησης
- Σχεδόν βέλτιστος σε άλλα σενάρια

## RSD

- Αρχικά, η τελική κατάταξη είναι κενή

### Σειριακή φάση

- Εξετάζουμε τις επί μέρους κατατάξεις με τυχαία σειρά
- Για κάθε επί μέρους κατάταξη, αντιγράφουμε όλες τις σχέσεις μεταξύ εργασιών, εφόσον δεν διαφωνούν με τις σχέσεις που έχουμε μέχρι τώρα, και
- συμπληρώνουμε τις σχέσεις που προκύπτουν λόγω μεταβατικότητας

### Φάση τυχαίας συμπλήρωσης

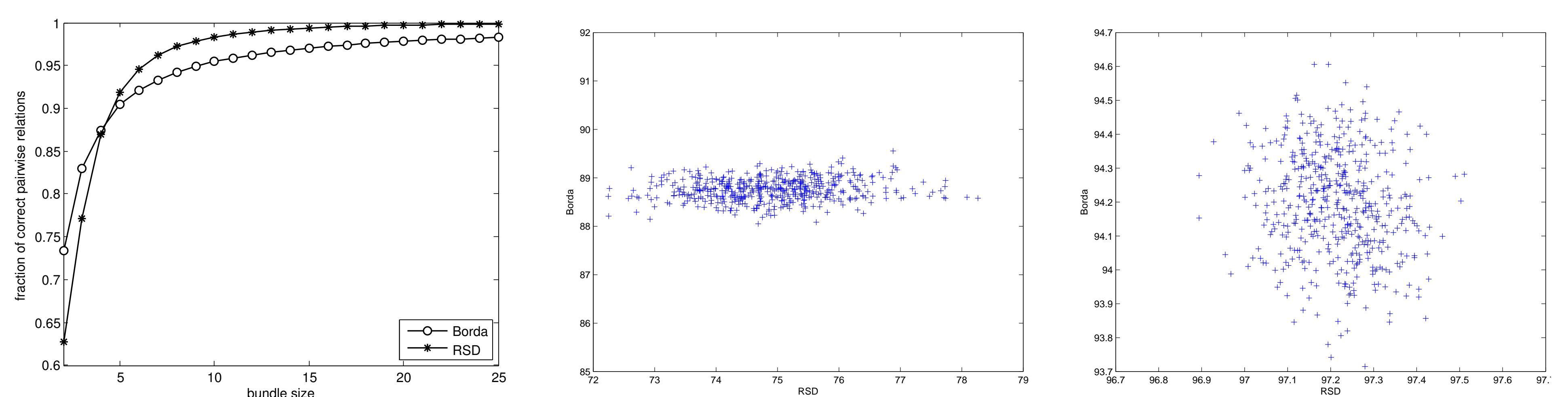
- Επαναληπτικά, διαλέγουμε τυχαία ένα ζευγάρι εργασιών για το οποίο δεν γνωρίζουμε τη σχέση του,
- την αποφασίζουμε τυχαία, και
- συμπληρώνουμε όλες τις σχέσεις που προκύπτουν λόγω μεταβατικότητας

## Κανόνες κατάταξης τύπων

- Κάθε εργασία έχει έναν *τύπο* ο οποίος είναι ο συνδυασμός των θέσεων που έχει πάρει στις επί μέρους κατατάξεις που εμφανίζεται
- Παράδειγμα: μια εργασία με τύπο  $(1, 3, 3)$  εμφανίζεται σε τρεις επί μέρους κατατάξεις, σε μία από αυτές είναι στη πρώτη θέση και στις άλλες δύο είναι στη τρίτη θέση
- Ένας τέτοιος κανόνας ορίζει μια κατάταξη επί όλων των δυνατών τύπων και η τελική συνολική κατάταξη των εργασιών προκύπτει από αυτή την κατάταξη τύπων, επιλύοντας ισοπαλίες τυχαία
- Ειδική περίπτωση: ο κανόνας Borda

## Σύγκριση Borda και RSD

- Ο RSD έχει εξαιρετική απόδοση (πολύ κοντά στο 100%) στο σενάριο τέλει βαθμολόγησης, αλλά δεν είναι καθόλου αποδοτικός σε πιο ρεαλιστικά σενάρια
- Ο Borda υπερέρχει στο σενάριο ατελούς βαθμολόγησης και έχει πολύ καλή απόδοση στο σενάριο τέλει βαθμολόγησης



## Βέλτιστοι κανόνες κατάταξης τύπων

- Δεδομένου ενός μοντέλου θορύβου που αναπαριστά ένα σενάριο βαθμολόγησης, έχουμε αναπτύξει μια μεθοδολογία που υπολογίζει το βέλτιστο κανόνα κατάταξης τύπων
- Πως; Λύνοντας ένα στιγμιότυπο του προβλήματος FEEDBACK ARC SET
- Εύρωστη μέθοδος ακόμη και αν το μοντέλο θορύβου δεν είναι ακριβές

