

Αν. Καθηγητής Π. Λουρίδας

Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας

Οικονομικό Πανεπιστήμιο Αθηνών

Πόσο Επαναστατική Είναι η Επιστήμη;

Μία πρόσφατη έρευνα αμερικάνων ερευνητών (Park, Leahey, and Funk 2023) δημιούργησε πληθώρα συζητήσεων σχετικά με το πόσο επαναστατική, ή ανατρεπτική, είναι η επιστημονική έρευνα. Μια έρευνα είναι επαναστατική όταν ανοίγει νέους ορίζοντες, και δεν χτίζει απλώς πάνω σε προηγούμενα ευρήματα, αλλά θέτει τα θεμέλια για διερεύνηση ερωτημάτων που ξεφεύγουν από το καθιερωμένο, μέχρι εκείνη τη στιγμή πλαίσιο. Οι ερευνητές βρήκαν ότι η επιστήμη τις τελευταίες δεκαετίες έχει γίνει λιγότερο ανατρεπτική.

Για να ποσοτικοποιηθεί η ανατρεπτικότητα της επιστημονικής έρευνας, οι ερευνητές χρησιμοποίησαν μία μετρική η οποία ονομάζεται *CD* (Consolidates or Destabilizes) index (Funk and Owen-Smith 2017). Η μετρική *CD* αποτιμά το κατά πόσο μία έρευνα εμπεδώνει και εμβαθύνει προηγούμενη έρευνα (consolidates) ή αντίθετα κουνά τα θεμέλια και αποσταθεροποιεί (destabilizes) την προηγούμενη έρευνα.

Ο τρόπος με τον οποίο προκύπτει η μετρική *CD* για μία έρευνα είναι η εξής. Κατ' αρχήν δημιουργούμε έναν γράφο με το σύνολο των ερευνών που έχουμε στη διάθεσή μας: οι κόμβοι του γράφου είναι οι έρευνες και οι σύνδεσμοι αντιστοιχούν στις αναφορές που γίνονται. Αν μία έρευνα *A* κάνει αναφορά σε μία έρευνα *B*, τότε προσθέτουμε έναν σύνδεσμο μεταξύ των κόμβων *A* και *B* του γράφου.

Στον γράφο αυτό, εντοπίζουμε τον κόμβο που αντιστοιχεί στην έρευνα που μας ενδιαφέρει. Από τον κόμβο αυτό, βρίσκουμε τους κόμβους με τους οποίους συνδέεται, μέσω των αναφορών όπως αναφέραμε παραπάνω. Ας ονομάσουμε το σύνολο αυτών των κόμβων *P*. Όλοι αυτοί οι κόμβοι του συνόλου αντιπροσωπεύουν εργασίες που έχουν δημοσιευθεί πριν από την εργασία που μας ενδιαφέρει και έχουν χρησιμοποιηθεί (άρα αναφερθεί) από αυτήν. Βρίσκουμε επίσης τους κόμβους οι οποίοι έχουν αναφορές στον κόμβο που μας ενδιαφέρει, ή σε οποιονδήποτε από τους κόμβους του συνόλου *P*. Ας ονομάσουμε το σύνολο αυτών των κόμβων *F*. Τότε η μετρική ορίζεται ως:

$$CD = \frac{\sum_{i \in F} -2f_i b_i + f_i}{|F|}$$

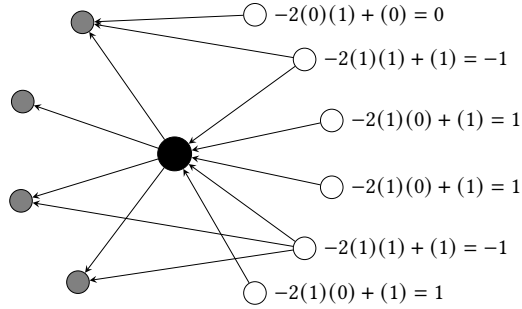
όπου:

$$f_i = \begin{cases} 1, & \text{αν η εργασία } i \text{ αναφέρει την εργασία που μας ενδιαφέρει} \\ 0, & \text{αν όχι} \end{cases}$$

και

$$b_i = \begin{cases} 1, & \text{αν η εργασία } i \text{ αναφέρει κάποια εργασία } p \in P \\ 0, & \text{αν όχι} \end{cases}$$

Ας δούμε τη μετρική στην πράξη. Στην παρακάτω εικόνα, έχουμε σημειώσει με μαύρο την εργασία που μας ενδιαφέρει, με γκρίζο τις εργασίες στις οποίες κάνει αναφορά, δηλαδή το σύνολο P , και με λευκό τις εργασίες που έχουν αναφορές στις εργασίες του συνόλου P ή στην εργασία που μας ενδιαφέρει, δηλαδή το σύνολο F .



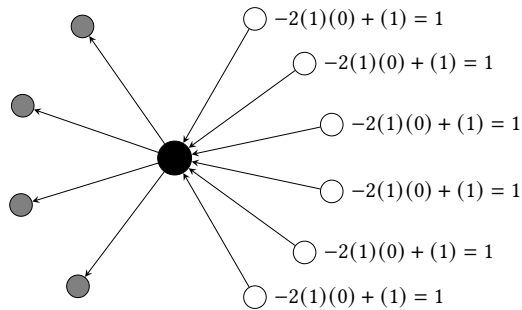
Στα δεξιά κάθε κόμβου του συνόλου F εμφανίζεται ο αντίστοιχος όρος του αθροίσματος που θα μπει στον αριθμητή της CD . Στη συνέχεια, σχηματίζουμε το άθροισμα:

$$\sum_{i \in F} -2f_i b_i + f_i = (0) + (-1) + (1) + (1) + (-1) + (1) = 1$$

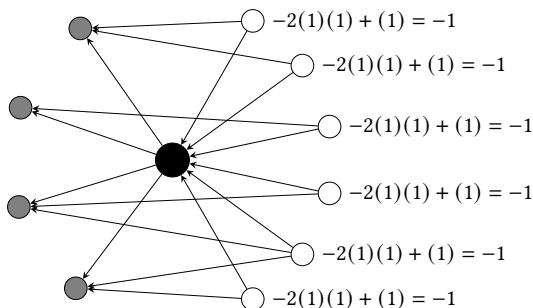
Τέλος, διαιρούμε με το πλήθος των στοιχείων του F για να πάρουμε το CD :

$$CD = \frac{1}{6} = 0.17$$

Οι δύο ακραίες περιπτώσεις της τιμής της CD είναι 0 και 1. Έχουμε $CD = 1$ όταν οι κόμβοι του συνόλου F δεν έχουν καμία αναφορά στους κόμβους του συνόλου P , όπως φαίνεται στο παρακάτω παράδειγμα.



Αντιστρόφως, έχουμε $CD = 0$ όταν όλοι οι κόμβοι του συνόλου F έχουν αναφορές στους κόμβους του συνόλου P , όπως φαίνεται στο παρακάτω παράδειγμα.



Ο σκοπός της εργασίας είναι, αν σας δίνεται ένας γράφος και ένας κόμβος, να υπολογίζετε τη μετρική CD με δύο τρόπους:

- Αναπαριστώντας το γράφο με λίστες γειτνίασης.
- Αναπαριστώντας το γράφο με πίνακα γειτνίασης.

Απαιτήσεις Προγράμματος

Κάθε φοιτητής θα εργαστεί σε αποθετήριο στο GitHub. Για να αξιολογηθεί μια εργασία θα πρέπει να πληροί τις παρακάτω προϋποθέσεις:

- Για την υποβολή της εργασίας θα χρησιμοποιηθεί το ιδιωτικό αποθετήριο του φοιτητή που δημιουργήθηκε για τις ανάγκες του μαθήματος και του έχει αποδοθεί. Το αποθετήριο αυτό έχει όνομα του τύπου `username-algo-assignments`, όπου `username` είναι το όνομα του φοιτητή στο GitHub. Για παράδειγμα, το σχετικό αποθετήριο του διδάσκοντα θα ονομαζόταν `louridas-algo-assignments` και θα ήταν προσβάσιμο στο <https://github.com/dmst-algorithms-course/louridas-algo-assignments>. Τυχόν άλλα αποθετήρια απλώς θα αγνοηθούν.
- Μέσα στο αποθετήριο αυτό θα πρέπει να δημιουργηθεί ένας κατάλογος `assignment-2022-5`.
- Μέσα στον παραπάνω κατάλογο το πρόγραμμα θα πρέπει να αποθηκευτεί με το όνομα `cd5.py`.
- Δεν επιτρέπεται η χρήση έτοιμων βιβλιοθηκών γράφων ή τυχόν έτοιμων υλοποιήσεων των αλγορίθμων, ή τμημάτων αυτών, εκτός αν αναφέρεται ρητά ότι επιτρέπεται.
- Επιτρέπεται η χρήση δομών δεδομένων της Python όπως στοιβές, λεξικά, σύνολα, κ.λπ.
- Επιτρέπεται η χρήση των παρακάτω βιβλιοθηκών ή τμημάτων τους:
 - `sys.argv`
 - `argparse`

- Το πρόγραμμα θα πρέπει να είναι γραμμένο σε Python 3.

Το πρόγραμμα θα καλείται ως εξής (όπου `python` η κατάλληλη εντολή στο εκάστοτε σύστημα):

```
python cd5.py [-m] input_file node
```

Η σημασία των παραμέτρων είναι η εξής:

- Η παράμετρος `input_file` μας δίνει το όνομα του αρχείου με το γράφο τον οποίο θα χρησιμοποιήσουμε. Το αρχείο περιέχει μία γραμμή για κάθε ακμή του γράφου. Η γραμμή `A B` σημαίνει ότι υπάρχει αναφορά από την εργασία `A` στην εργασία `B`.
- Η παράμετρος `node` μας δίνει τον κόμβο για τον οποίο θέλουμε να υπολογίσουμε τη μετρική `CD`.
- Αν δίνεται η παράμετρος `-m` τότε για την αναπαράσταση του γράφου θα πρέπει να χρησιμοποιήσετε πίνακα γειτνίασης.

Παραδείγματα

Παράδειγμα 1

Αν ο χρήστης του προγράμματος δώσει:

```
python3 cd5.py example_1.txt 0
```

τότε το πρόγραμμα θα διαβάσει τον γράφο από το αρχείο `example_1.txt` και στην έξοδο θα εμφανίσει:

```
0.16666666666666666
```

Παράδειγμα 2

Αν ο χρήστης του προγράμματος δώσει:

```
python3 cd5.py example_2.txt 0
```

τότε το πρόγραμμα θα διαβάσει τον γράφο από το αρχείο `example_2.txt` και στην έξοδο θα εμφανίσει:

```
1.0
```

Παράδειγμα 3

Αν ο χρήστης του προγράμματος δώσει:

```
python3 cd5.py example_3.txt 0
```

τότε το πρόγραμμα θα διαβάσει τον γράφο από το αρχείο `example_3.txt` και στην έξοδο θα εμφανίσει:

```
-1.0
```

Παράδειγμα 4

Αν ο χρήστης του προγράμματος δώσει:

```
python3 cd5.py example_4.txt 4
```

τότε το πρόγραμμα θα διαβάσει τον γράφο από το αρχείο [example_4.txt](#) και στην έξοδο θα εμφανίσει:

```
0.16666666666666666
```

Περισσότερες Πληροφορίες

Funk, Russell J., and Jason Owen-Smith. 2017. “A Dynamic Network Measure of Technological Change.” *Management Science* 63 (3): 791–817.

Park, Michael, Erin Leahey, and Russell J. Funk. 2023. “Papers and Patents Are Becoming Less Disruptive over Time.” *Nature* 613 (7942): 138–44.

Καλή Επιτυχία!