

עבודה 1 – ניתוח רשתות חברתיות

מגישים: מיכל לורברבום 311120133, גל מאיר 305382137

Implementation

מצורף בקובץ הזיפ בקובץ api.py.

הסבר קצר:

הגדרנו גולבלית רשימה של כל הקודקודים [inOrder] ו31 מפות:

- adjacency_list - רשימת שכנויות
- reverse_adjacency_list - רשימת שכנויות הפוכה (כדי לדעת מהן הצלעות הנכנסות של כל קודקוד – צריך את המידע הזה לחישוב הPageRank)
- byValue - מפה השומרת לכל קודקוד את ערך הPageRank שלו.

תחילה ביצענו קריאה של הקובץ והכנסתו לשתי מפות אחת לפי מקור [adjacency_list] ואחת לפי יעד [reverse_adjacency_list], ועדכנו את רשימת הקודקודים [inOrder].
חישוב הPageRank –

בתחילת החישוב עדכנו את המפה של הPageRank להיות בערך ההתחלתי שהוא $\frac{1}{N}$ (כאשר N זהו מספר הקודקודים) בנוסף הגדרנו 2 מפות נוספות:

- prev_value – מכילה את ערך הPageRank של כל קודקוד עבור האיטרציה הקודמת (בשביל חישוב הדלתא)
- tmp_value – מקבילה ל temporary PageRank מהנוסחה של הPageRank שראינו בהרצאה.

עבור כל קודקוד node אנו בודקים האם דרגת הכניסה שלו גדולה מאפס (אחרת ערך הPageRank שלו הוא 0) ועבורו מחשבים סכום של ערכי הPageRank של כל הקודקודים שנכנסים אליו חלקי דרגת היציאה של אותם קודקודים כפול בטא. בנוסף אנו מתחילים לצבור את הערכים שאנו מקבלים למשתנה זמני sum.

לאחר שלב מבצעים לכל קודקוד את שלב הPageRank leaked, מוסיפים לכל ערך שיצא לנו $\frac{1-sum}{N}$. את הערך הזה שומרים במפה של הPageRank.

לבסוף מחשבים את ההבדל בין התוצאות של האיטרציה הנוכחית עם האיטרציה הקודמת. ובודקים אם ההבדל גדול או קטן מהדלתא הנתונה לנו.

Evaluation

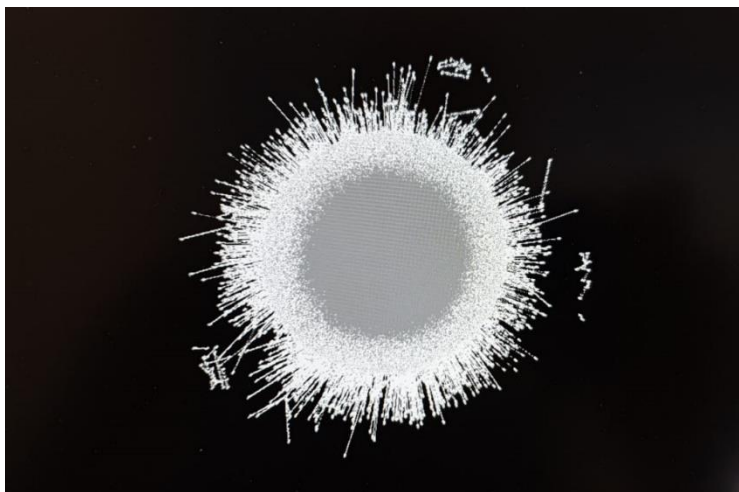
Wikipedia vote

1. סופק במודל של הקורס, זהו גרף שמתאר את הקשרים בין דפי הויקיפדיה השונים (איזה דף מצביע על איזה דף)
2. מספר קודקודים: 7115
3. מספר צלעות: 103689
- 4.

מספר קודקוד	ערך הPageRank
4037	0.00460748439537952
15	0.003680192055783484
6634	0.0035697588856128943

0.003284786278299254	2625
0.002608037581589704	2398
0.0025247548906776147	2470
0.0024978382264644605	2237
0.0022674014970163383	4191
0.002169363779832553	7553
0.0021500569410050634	5254

.5



facebook_combined

1. גרף זה מתאר מעגלים חברתיים (אנונימיים) בפייסבוק.

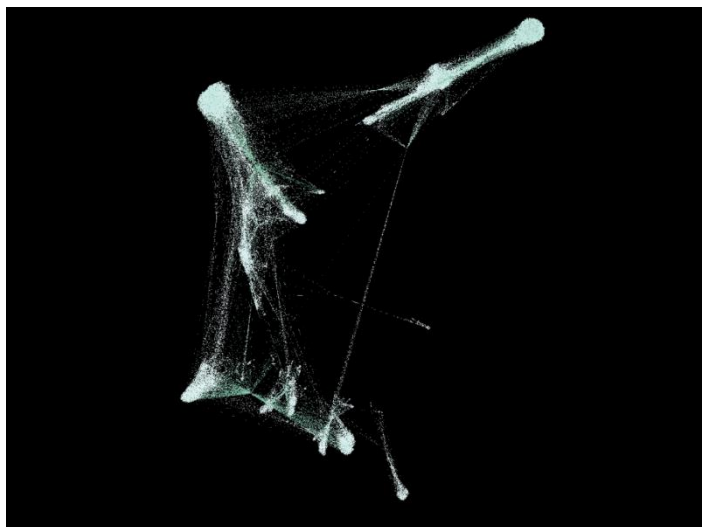
המקור: <http://snap.stanford.edu/data/index.html>

2. מספר קודקודים: 4,039

3. מספר צלעות: 88,234

.4

מספר קודקוד	ערך הPageRank
1911	0.009390175832560226
3434	0.009371212491909931
2655	0.009038053944431355
1902	0.008962911790390648
1888	0.006886442886055308
2649	0.006266898843730888
1907	0.005139001791838376
3971	0.00506553003250321
2654	0.004918068790617721
1910	0.004197650698068683



Extending the API:

על מנת לתמוך בגרף ממושקל יש להרחיב את האלגוריתם הנ"ל כך שכל קשת יש להכפיל במשקלה.