



מבחן בקורס: "אלגוריתמים ויישומים ברשתות חברתיות"

סמסטר א' תשפ"א - מועד ב' (25.03.2021)
מרצה: ד"ר סלבה נובגורודוב

משך המבחן: 3 שעות.
אין להשתמש בחומר עזר מלבד מחשבון כיס.
יש לתת תשובות קצרות ומנומקות היטב.

בהצלחה!

שאלה 1 (15 נק'):

בקבוצה של 120 חברי כנסת, כל אחד קיבל מספר סידורי בין 1 ל-120. כל חבר כנסת הכריז שהוא מוכן לשבת בקואליציה אך ורק עם חבר כנסת אחר שלמספרים המזהים שלהם יש לפחות 2 מחלקים ראשוניים שונים במשותף. לדוגמא, חבר כנסת מספר 45 מוכן לשבת עם חבר כנסת מספר 15 (שני המספרים מתחלקים ב-3 וב-5), אבל לא מוכן לשבת עם למשל 9, 11, 50, 64 וכו'.

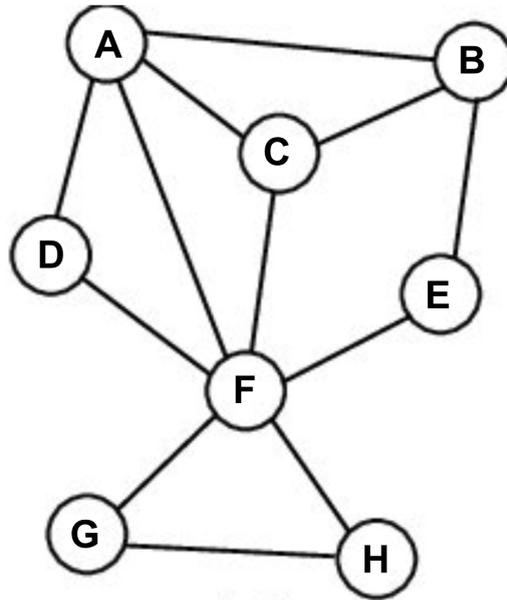
א. מה הקואליציה הכי גדולה שניתן לבנות עם חברי הכנסת האלה?

ב. כמה משולשים יש בקואליציה שהתקבלה בסעיף א'?

ג. כמה חברי כנסת שהמספר המזהה שלהם מתחת ל-60 לא מוכנים לשבת עם אף אחד אחר?

שאלה 2 (15 נק'):

נתונה רשת חברתית בעלת 9 קודקודים:



א. הריצו אלגוריתם לזיהוי קהילות K-Clique percolation method עם פרמטר $K = 3$ וחלקו אותה לקהילות.

ב. האם הקהילות שהתקבלו נחתכות? הסבירו את תשובתכם.

ג. מה כמות הקשתות המינימלית שניתן להוסיף לגרף על מנת שתתקבל רק קהילה אחת.

שאלה 3 (10 נק'):

נתונה רשת חברתית בעלת 9 קודקודים ו 35 קשתות.

נרצה לבצע עליה Link Prediction לפי מדדים שונים. נתון כי זוג הקודקודים A, B אינם מחוברים בגרף ו מדד ה Common Neighbors שלהם שווה ל 7.

א. לאיזה ערך יהיה שווה מדד ה Jaccard של זוג הקודקודים A, B.

ב. לאיזה ערך יהיה שווה מדד ה Adamic/Adar של זוג הקודקודים A, B? (הניחו שהבסיס של ה log בחישוב המדד הוא 2)

שאלה 4 (15 נק'):

נתון גרף מלא עם $N > 3$ קודקודים. מחלקים את הקודקודים לשתי קבוצות זרות - קבוצה בגודל L וקבוצה בגודל R (כך ש $L+R=N$). נתון שגודל של כל קבוצה הוא לפחות 2. כעת מחליטים על הסימנים על הקשתות באופן הבא: הקשתות שמחברות 2 צמתים מאותה קבוצה מקבלות סימן חיובי. כל קשת שמחברת 2 צמתים מקבוצות שונות מקבלת סימן חיובי בהסתברות p וסימן שלילי בהסתברות q (כך ש $p+q=1$). הסימן של כל קשת נקבע באופן בלתי תלוי מהסימנים של הקשתות האחרות.

א. מה ההסתברות לקבל גרף מאוזן (מבחינת ה Theory of Structural Balance)?

ב. מה היא התוחלת של כמות המשולשים הלא מאוזנים?

ג. מה ההסתברות שקיים מסלול פשוט (שלא עובר באף קודקוד פעמיים) שמורכב אך ורק מקשתות חיוביות שעובר בכל הקודקודים?

שאלה 5 (10 נק'):

נתונה רשימה של 4 קבוצות ו 8 שחקנים ששיחקו בהן במהלך השנים. כל שורה ברשימה מייצגת קבוצה וכוללת את רשימת השחקנים ששיחקו שם.

[T - Team, P - Player]

T1: P2, P4, P6, P8

T2: P1, P3, P5, P7

T3: P1, P2

T4: P2, P5, P7

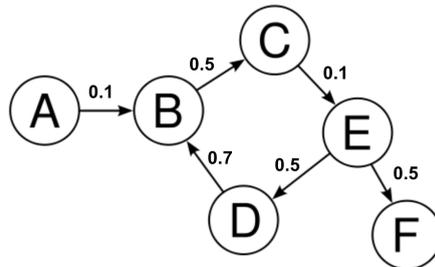
הציגו את הנתונים כגרף דו-צדדי ובצעו עליו folding, על מנת לקבל גרף שמייצג אילו שחקנים שחקו באותה הקבוצה (folded graph).

א. איזה קודקוד בעל Closeness centrality הכי גבוה ב folded graph?

ב. מצאו את כל נקודות ה periphery שקיימות ב folded graph.

שאלה 6 (10 נק')::

נתון גרף מכוון בעל 6 קודקודים המתאר השפעה של אנשים. כל קשת היא מכוונת, ועל הקשת מצויינת רמת ההשפעה. נרצה למצוא את הקודקודים הכי משפיעים (שיטת ה Inﬂuence Maximization), לפי המודל של Linear Threshold (בקיזור LT).
נניח שלכל קודקוד יש סף קבוע: A: 0.6, B: 0.6, C: 0.2, D: 0.8, E: 0.4, F: 0.2



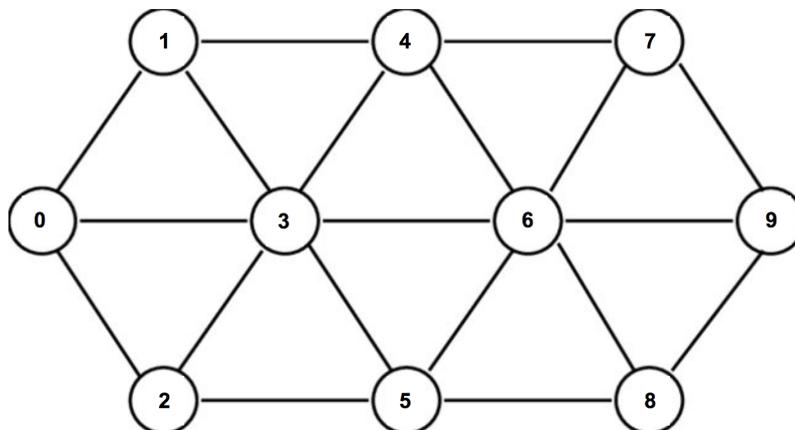
א. מי הקודקוד הכי משפיע בגרף?

ב. הריצו את האלגוריתם Greedy Hill Climbing ומצאו את זוג הקודקודים הכי משפיעים. תארו את השלבים של האלגוריתם.

שאלה 7 (10 נק')::

א. בחברה של 17 עובדים כל זוג עובדים מתקשר באופן קבוע בבדיוק אחת משלושת הדרכים הבאות: אימיילים, מסרונים ושיחות טלפון. הוכיחו שקיימים שלושה עובדים שמתקשרים אחד עם השני בבדיוק באותה הדרך.

ב. נתונה רשת של זרימת מידע. המידע עובר מקודקוד עם מספר קטן יותר לקודקוד עם מספר גדול יותר ורק לפי הקשתות הנתונות.



בכמה דרכים שונות הודעה היוצאת מקודקוד 0 יכולה להגיע לקודקוד 9?

שאלה 8 (15 נק'):

נתונה רשת חברתית שנוצרה בצורה אקראית באמצעות המודל של Erdős–Rényi, עם הפרמטרים $N=40$, $p=0.5$.

- א. מה היא התוחלת של הדרגה הממוצעת ברשת הזאת?
- ב. מה היא התוחלת של כמות הקליקות השונות בגודל 5 ברשת הזאת?
- ג. מה היא התוחלת של כמות המעגלים השונים באורך 5 ברשת הזאת?