



מבחן בקורס: "אלגוריתמים ויישומים ברשתות חברתיות"

סמסטר א' תשפ"ב - מועד א' (17.01.2022)
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 (10 נק'):

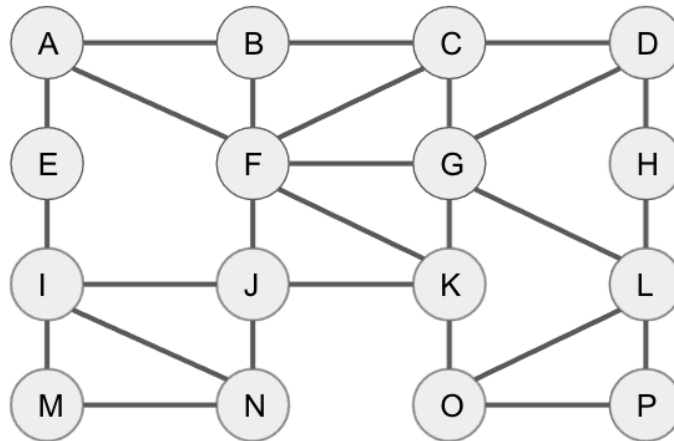
לקראת תחרות פינג-פונג באוניברסיטה, 3 נבחרות של בית הספר למדעי המחשב (נבחרות של שנה א', ב' וג') החליטו לעשות אימון משותף. ידוע שבכל נבחרת יש כמות זהה של שחקנים (פחות מ-7). כל שחקן בנבחרת של שנה ג' שיחק מול כל שחקן מהנבחרות של שנה א' ושנה ב' (לא היו משחקים נוספים). אחרי סיום האימון התברר שכמות הניצחונות שהשיגו שחקני נבחרת שנה ג' (ביחד) פי 9 יותר גדולה מכמות ההפסדים שלהם (ביחד).

א. כמה שחקנים יש בכל נבחרת?

ב. נתון שאף שחקן מנבחרות של שנים א' וב' לא ניצח יותר ממשחק אחד. השחקנים החליטו להמשיך את האימון לעוד סיבוב ולהתחלק לפי רמות ל2 קבוצות: ראשית ומשנית. קבוצת כל השחקים משנה ג' ועוד השחקנים מהשנים א' וב' שניצחו משחק אחד בסיבוב הראשון הוגדרו כקבוצה הראשית. שאר השחקנים מהשנים א' וב' הוגדרו כקבוצה המשנית. כל השחקנים בקבוצה הראשית שיחקו אחד נגד השני משחק אחד וגם כל השחקנים מהקבוצה המשנית שיחקו אחד נגד השני משחק אחד (לא היו משחקים נוספים). כמה משחקים התקיימו בסיבוב האימון השני?

שאלה 2 (15 נק')::

נתונה רשת חברתית בעלת 16 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. מוסיפים את הקשת (G, H) לרשת. איך זה משפיע על התשובות לסעיפים א' וב'?

שאלה 3 (10 נק')::

נתונה רשת חברתית בעלת 12 קודקודים ו 35 קשתות.

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

א. האם ייתכן שמדד ה Jaccard של זוג הקודקודים A, B יוצא גדול מ- 0.9?

ב. האם ייתכן שמדד ה Jaccard של זוג הקודקודים A, B יוצא קטן מ- 0.5?

שאלה 4 (15 נק'):

נתון גרף מאוזן עם N רכיבי קשירות וסימן על כל קשת. ברכיב ה- i יש n_i צמתים ו- m_i קשתות, מתוכן p_i קשתות עם סימן חיובי. כמה דרכים יש להשלים את הגרף לגרף מלא כך שהוא עדיין יישאר מאוזן. אין להניח שכל רכיב קשירות הוא גרף מלא.

א. עבור $N = 1$

ב. עבור $N = 3$

ג. עבור N כללי

שאלה 5 (15 נק'):

נתונה רשימה של 9 תיירים שביקרו בערים שונות. כל שורה ברשימה מייצגת תייר ורשימת הערים שהוא ביקר בהן.

[C - City, T - Tourist]

T1: C1;	T2: C1, C2, C3;
T3: C3;	T4: C1, C4, C5;
T5: C5;	T6: C1, C6, C7;
T7: C7;	T8: C1, C8, C9;
T9: C9;	

הציגו את הנתונים כגרף דו-צדדי ובצעו עליו folding, על מנת לקבל גרף שמייצג אילו אנשים ביקרו באותה העיר (folded graph).

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

ב. חשבו את ה Clustering Coefficient של הקודקוד מסעיף א'.

ג. מצאו את כל קליקות ה- Maximum וה- Maximal ב folded graph (יש להתייחס רק לקליקות בעלות 3 קודקודים או יותר).

שאלה 6 (10 נק'):

נתון גרף מכוון בעל N קודקודים המתאר השפעה של אנשים. כל קשת היא מכוונת, ועל הקשת מצויינת רמת ההשפעה. נרצה למצוא את קבוצת הקודקודים הכי משפיעה (שיטת ה-Influence Maximization), לפי ההרחבה הבאה של מודל ה-Independent Cascade.

נשנה את המודל כך שבמקום ניסיון הדבקה אחד, כל קודקוד שנהיה אקטיבי, מקבל 2 ניסיונות בלתי תלויים להדביק כל אחד מהשכנים שלו. הניסיונות מתבצעים ברצף, אחד אחרי השני. הסיכוי של כל ניסיון להצליח הוא המשקל על הקשת (כמו במודל המקורי).

א. האם הפתרון (על אותו הקלט) משתנה כאשר אנו משתמשים במודל המורחב החדש במקום במודל ה-Independent Cascade המקורי? בהכרח כן, בהכרח לא, או שיש קלטים שעליהם כן וקלטים אחרים שעליהם לא?

ב. נתון גרף "שרשרת" בעל 36 קודקודים, הממוספרים מ-1 עד 36, כך שמכל קודקוד i יוצאת קשת מכוונת לקודקוד $i+1$ (סה"כ 35 קשתות). המשקל על כל קשת הוא $p < 1$ (אותו p לכל קשת), וכמו בסעיף א', לכל קודקוד יש 2 ניסיונות הדבקה רצופים. עבור $k=2$ מה הפתרון האופטימלי?

שאלה 7 (10 נק'):

א. בצוות של 5 רופאות שעושות סטאז' באותו בית החולים, כל שתיים מהן מתכתבות במסגרת העבודה. כל זוג רופאות הסכימו על דרך תקשורת מסוימת ויחידה מתוך 2 האפשרויות הבאות: להתכתב באמצעות הודעות מייל (שיטה א') או להתכתב באמצעות הודעות במערכת הפנימית של בית החולים (שיטה ב'). אחרי החלוקה הזאת התברר שאין שלישיית רופאות כך ששלושתן מתכתבות אחת עם השניה באותה השיטה. הוכיחו כי כל רופאה מתכתבת עם 2 מהקולגות שלה בשיטה א' ועם 2 האחרות בשיטה ב'.

ב. אחרי סיום תקופת הסטאז' כל הרופאות עברו לבתי חולים אחרים, ותקשורת באמצעות המערכת הפנימית נפסקה (שיטה ב'). הרופאות שהיו בקשר באמצעות הודעות מייל (שיטה א') ממשיכות להיות בקשר. האם הגרף החדש שמייצג את התקשורת בין חמשת הרופאות עדיין קשיר?

שאלה 8 (15 נק'):

נרצה לבנות גרף עם 5 קודקודים באופן אקראי באמצעות המודל של Erdős–Rényi, עם הסתברות ליצירת קשת $p = 0.5$.

א. מה הסיכוי שיצא גרף שדרגת כל הקודקודים שלו היא 2?

ב. מה הסיכוי שיצא גרף שדרגת כל הקודקודים שלו היא 3?

ג. מה הסיכוי שיצא גרף שדרגת כל הקודקודים שלו היא 4?