



# קורס רשתות מורכבות

## פרויקט מסכם -

## קורקינטים שיתופיים בשיקגו



מרצה: דר' אסף אלמוג

מגישים:

טל כהן 205985856

ציון גמליאל 203786587

יסכה שר שלום 206221764

## תוכן עניינים

3.....	תקציר
4.....	אפיון הנתונים ומקורות המידע
4.....	תיאור כללי:
4.....	תיאור הנתונים:
5.....	מטרות הפרויקט :
5.....	ניתוח תכונות הרשת המלאה
5.....	הרשת הכיוונית :
10.....	הרשת הלא כיוונית :
11.....	:Community Detection
12.....	: Centrality ניתוח מדדי
13.....	ניתוח תתי גרפים
13.....	ניתוח רשת לפי ימים ושעות :
15.....	ניתוח הרשת לפי אמצע שבוע וסוף שבוע:
16.....	ניתוח מסלולים ואורכי נסיעות ברשת :
17.....	ניתוח הרשת לאורך זמן :
19.....	סיכום :
21.....	נספחים :

## תקציר

בפרויקט זה הצגנו נתונים שנשאבו מאתר האינטרנט של עיריית שיקאגו בנוגע לפיילוט על נסיעות בקורקינטים שיתופיים שבוצעו ברחבי העיר. מידלנו 382,341 נסיעות שבוצעו לאורך 4 חודשים, ב-9,711 מסלולים שונים. בפיילוט השתתפו 21 אזורים מתוך 77 האזורים אליהם מחולקת העיר.

לאחר המידול, ביצענו ניתוחים מעולם ה-complex networks במטרה להבין את אופי השימוש בקורקינטים, את קהל היעד המרכזי שלהם, את המקומות בהם נרצה להקים עמדות הטענה וחניות של קורקינטים, את המסלולים המרכזיים בהם נוסעים בקורקינטים ואת השפעת הקורקינטים על מרקם חיי התחבורה בעיר.

את מרבית הניתוחים ביצענו על רשת כיוונית וממושקלת, בה כל צומת מציין מיקום אשר בו התחילה ו/או הסתיימה נסיעה. משקל כל קשר מייצג את כמות הנסיעות בין הצמתים ביניהם הוא מקשר, וכיוונו מייצג את כיוון הנסיעה. על רשת זו ביצענו ניתוחים שונים במטרה להבין את התנהגות הרשת. בניגוד לצפויותינו, מרבית הנסיעות היו בין אזורי המרכז, ולא בין הפריפריה למרכז.

כמו כן, ניתחנו את מדדי ה-centrality של הרשת. אחת התובנות המרכזיות שעלו מניתוח זה הוא מדד ה-Betweenness centrality אשר העלה כי מרבית הנקודות המרכזיות ברשת ממוקמות בסמוך לתחנות רכבת מרכזיות. תופעה זו מדגישה את השימוש בקורקינטים שיתופיים כתחבורה משלימה לרכבת. כך שנוסעים מאזורים שונים מבצעים נסיעות בקורקינטים אל ומתחנות הרכבת והתחנות מחברות את הרשת.

בהמשך, ביצענו ניתוחים שונים על הרשת המלאה ועל רשתות חלקיות. את הרשת חילקנו לפי ימים (סוף"ש ואמצ"ש), שעות (בוקר וערב), מרחקי נסיעות והתקדמות בפיילוט לאורך הזמן, במטרה לזהות מגמות משמעותיות.

המסקנות העיקריות שעלו מניתוחים אלו הן כי מרבית הנסיעות מתקיימות בשעות הצהריים-ערב או בסוף השבוע, ובקרבם של פארקים ומקומות משחק, דבר אשר מרמז לנו כי קהל היעד העיקרי של הקורקינטים הוא בני הנוער, וכי מטרתן של מרבית הנסיעות היא פנאי ומשחק. תובנה משמעותית נוספת אשר חזרה לכל אורך הפרויקט היא כי נסיעות רבות מתחילות ו/או מסתיימות בתחנות רכבת מרכזיות. מכך ניתן להסיק כי שימוש נוסף של הקורקינטים הוא לטובת "Last mile solution". על כן, יש להקים בתחנות רכבת חניות מותאמות לקורקינטים וכן עמדות הטענה.

בנוסף, ראינו כי מרבית הנסיעות מתבצעות בתוך אזורי מרכז העיר, ולא בין המרכז לפריפריה. לכן, אנו ממליצים לסלול באזורים אלו מסלולי נסיעה לרכבים דו-גלגליים על מנת לאפשר שילוב של הקורקינטים במרקם התחבורה בעיר, מבלי לפגוע בביטחון הרוכבים.

## אפיון הנתונים ומקורות המידע

### תיאור כללי:

קורקינטים שיתופיים הפכו לחלק ממרקם התחבורה העירוני, והינם חלק בלתי נפרד מפיתוח תחבורה בת קיימא בכל עיר מרכזית בעולם. השירות מבוסס על טכנולוגיה מתקדמת ומאפשר השכרה של קורקינטים באמצעות שימוש באפליקציה ייעודית, מכל מקום ובכל זמן נתון. כמו כן, הקורקינטים החשמליים כוללים נעילה אוטומטית ומחוברים למערכת GPS המאפשרת מעקב צמוד אחר תנועת המכשירים. השימוש בקורקינטים שיתופיים מאפשר לתושבי ואורחי העיר בה ממוקמים לעקוף את הפקקים, להימנע מחיפוש החנייה, לחסוך בעלויות הנסיעה וכמובן להתנייד בצורה יעילה ונעימה.

משתמשים שמעוניינים לשכור קורקינט חשמלי, צריכים להוריד אפליקציה לטלפון, לאחר התקנת האפליקציה, נכנסים אליה ומחפשים באמצעות מפה את הקורקינט הקרוב ביותר למקום האיסוף המבוקש. על מנת להתחיל לנסוע, יש לשחרר את הנעילה של הקורקינט על ידי כניסה לאפליקציה וסריקת קוד QR או באמצעות סיסמה. לאחר שחרור הנעילה, ניתן להתחיל לרכוב על הקורקינט החשמלי. לאחר סיום השימוש בקורקינט החשמלי יש לחנות את הקורקינט החשמלי בכל מקום מתאים לחניה. משתמשים יכולים גם לעבוד בטעינת הקורקינטים ולייצר הכנסה נוספת. התשלום בעבור השימוש בקורקינטים החשמליים אינו מצריך דמי מנוי ויועבר אוטומטית באמצעות כרטיס האשראי אותו הזין המשתמש באפליקציה, כאשר התשלום הוא לפי זמן הנסיעה (מרגע שחרור הנעילה ועד לרגע סימון סיום נסיעה).

ב-2019 בוצע פיילוט בעיר שיקגו על מנת להעריך את ההשפעה של הקורקינטים השיתופיים על הקהילות בשיקגו, על מערכות התחבורה הציבורית הקיימות ועל בטיחות התושבים והרוכבים בדרך. בפיילוט השתתפו עשר חברות שונות המפעילות שירותי קורקינט שיתופי בעיר.

### תיאור הנתונים:

בעבודת מחקר זו נעסוק בפיילוט שבוצע בעיר שיקגו על קורקינטים שיתופיים, הנתונים נאספו בין התאריכים 15 ביוני ועד 15 באוקטובר 2019. הנתונים שאיתם עבדנו לקוחים מאתר האינטרנט של עיריית שיקגו<sup>1</sup> ([www.chicago.gov](http://www.chicago.gov)).

בחרנו לעסוק בנושא קורקינטים שיתופיים כיוון שנושא זה נוגע גם אלינו בכלל הערים בישראל ובתל אביב בפרט. בנוסף, המגמה העולה של תחבורה שיתופית מביאה איתה אתגרים חדשים שנרצה לבחון בפרויקט זה.

הנתונים שנאספו עבור כל נסיעה הם: זמן התחלה, זמן סיום, מרחק, משך, דיוק המיקום (תלוי במפעיל ובמיקום), אזור התחלה (מספר ושם), אזור סיום (מספר ושם) ונקודת ציון התחלתית, נקודת ציון סופית. כחלק מעיבוד מקדים של הנתונים הוסרו רשומות בעלות שדות חסרים. כתוצאה מכך איבדנו כמחצית מהנתונים אך עדיין נשארו עם כמות משמעותית של נתונים המאפשרת לבצע

<sup>1</sup> לינק להורדה של הדאטה:

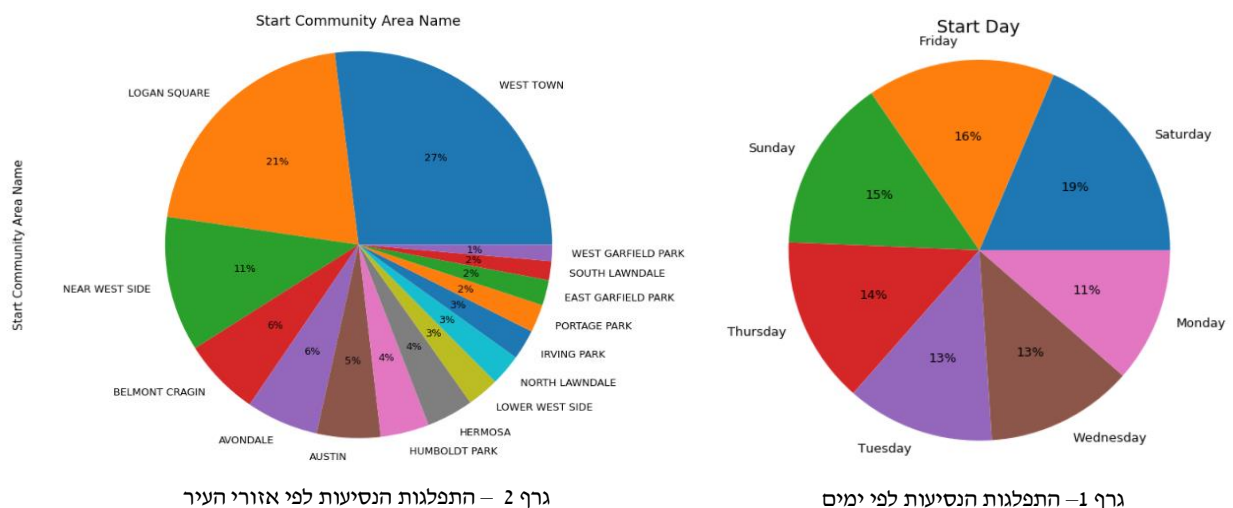
<https://drive.google.com/file/d/1rjKkl6Zae28lkHhrYYaVrEijIopMEJNt/view?usp=sharing>

ניתן מעמיק של הרשת. נציין, כי המפעילים בפילוט התחייבו בסוף כל יום להחזיר את הקורקינטים לנקודות מוצא מסוימות, דבר אשר עשוי להשפיע בהמשך על מבנה הרשת

הנתונים הסופים כוללים : 382,341 נסיעות ב-9,711 מסלולים שונים. העיר שיקגו מחולקת באופן מסורתי ל-77 אזורים, בפילוט השתתפו 21 אזורים. קו האורך וקו הגובה שנשמרו עבור המיקומים הם בדיוק של 6 ספרות אחרי הנקודה העשרונית.

בפרויקט זה נחקר את רשת המתארת את הנסיעות בקורקינטים ממקור ליעד. ברשת זו יש משמעות מה המקור ומה היעד ולכן היא כיוונית, בנוסף הוספנו משקלים כך שכל מסלול יופיע פעם אחת ומספר הנסיעות במסלול זה יהווה את המשקל שלו. בנוסף, נבנתה רשת שאינה כיוונית, המציגה המיקומים שבוצעו בהם נסיעות, על מנת לנתח את communities ולהסיק מסקנות נוספות.

במסגרת חקר הנתונים ביצענו וויזואליות שונות בכדי להבין את התפלגויות הנתונים-



גרף 2 – התפלגות הנסיעות לפי אזורי העיר

גרף 1 – התפלגות הנסיעות לפי ימים

### מטרות הפרויקט :

מטרתו העיקרית של הפרויקט הייתה להבין את אופי השימוש בקורקינטים על מנת לזהות מסלולים מרכזיים לסלילת נתיבים לכלים דו גלגליים. כמו כן, נרצה לאתר מיקומי מפתח בהם ניתן למצוא קורקינטים רבים בכדי להקים במיקומים אלו עמדות תשתית להטענת הקורקינטים וכן חניות מותאמות. בנוסף, נרצה לנסות ולזהות את מטרות השימוש בקורקינטים וקהל היעד המרכזי שלהם, זאת במטרה להבין איך משפיעים הקורקינטים על מרקם חיי התחבורה בעיר.

### ניתוח תכונות הרשת המלאה

#### הרשת הכיוונית :

עבור הרשת הכיוונית והממושקלת, בנינו גרף בו כל צומת מייצג נקודה (מיקום) וכל קשר מייצג מסלול נסיעה מנקודת התחלה לנקודת סיום. כאשר כל קשר שיוצא מנקודת ההתחלה (out degree) כנס לנקודת הסיום (in degree). מספר הנסיעות שבוצעו בכל מסלול מהווה את משקל הקשר.

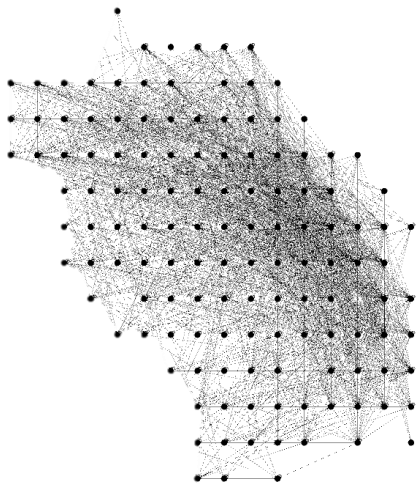




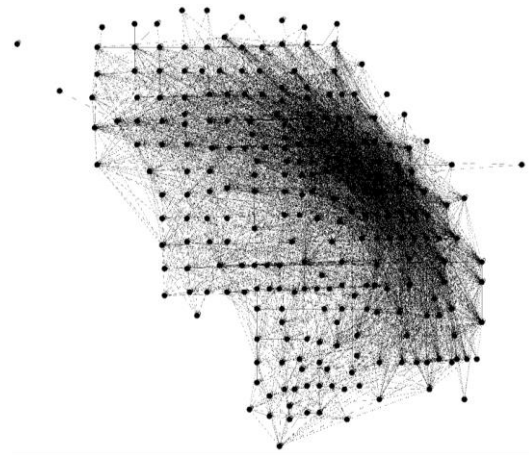
גרף 3 – הרשת המלאה על גבי מפה של העיר שיקגו

### עיגול מיקומי הצמתים ברשת:

המיקומים של הצמתים ברשת בנויים מנקודות אורך ורוחב, אשר מגיעות ממפעילים שונים ועל כן נתונות ברזולוציות שונות עקב דיוק שונה עבור כל מפעיל בפיילוט. בתחילת ניתוח הדאטה, שקלנו לבצע עיגול של המיקומים שיאלץ אחידות וכן ימנע מצב שבו שני מיקומים אשר כמעט זהים יקבלו צומת שונה בשל חוסר הדיוק בדאטה. לאחר בדיקה ראינו כי הרשת משתנה רק במידה ומיקומים יעוגלו עד 2 ספרות אחרי הנקודה, כאשר עיגול בסדר גודל זה מביא לאיחוד של נקודות במרחק של כמה מאות מטרים. ניתן לראות בגרפים למטה כי מצב זה הביא לאיבוד של מסלולים רבים. בנוסף, לא מעט מהנסיעות בקורקינטים הן קצרות באופן יחסי, ועל כן עיגול של כמה מאות מטרים עד כקילומטר מאבד חלק משמעותי מהנתונים ומטה את הדאטה. לכן, החלטנו להמשיך עם הרשת הלא מעוגלת עקב אופי השימוש בקורקינטים אשר משמשים בעיקר לנסיעות קצרות.



גרף 5 – הגרף המעוגל (2 ספרות)



גרף 4 – הגרף הלא מעוגל

החל משלב זה ועד לסוף הפרויקט, הרשת תוצג בפריסה של Geo Layout בגפי (כלומר על גבי מפה), עקב חשיבות המיקום בפרויקט.

אחרי ניתוח כללי של הרשת קיבלנו את מבנה הרשת הנ"ל:

Number of Nodes	Number of Edges	Average In Degree	Average Out Degree	Assortative In Degree	Assortative Out Degree
255	9711	38.0824	38.0824	0.233	0.144

טבלה 1 – מבנה הרשת הכיוונית

הגרף מכיוון ולכן אין משמעות לחישוב מדד Clustering Coefficient, על כן נבצע ניתוח למדד זה בהמשך, עבור רשת הלא כיוונית.

### : Degree Distribution

התפלגות הדרגות מוצגת בהמשך, בגרף בקנה מידה log-log. רשתות מסוג זה מאופיינות בדרגה ממוצעת גבוהה וקשרים מרובים בין קודקודי הרשת – מכל נקודה מסלולים לנקודות רבות אחרות.

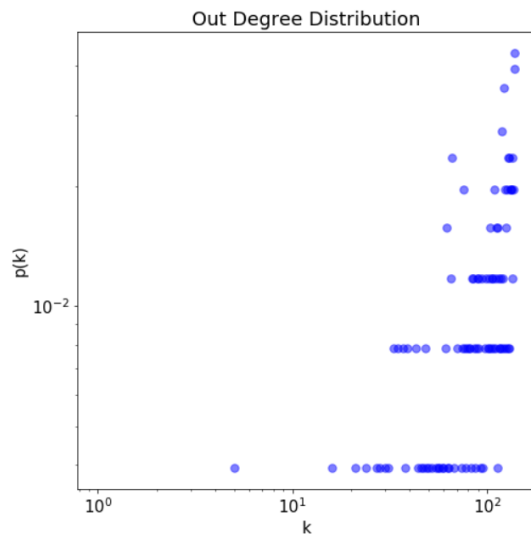
#### In Degree Distribution - גרף $P(K)$ :

הגרף מציג מה ההסתברות של מיקום להיות עם דרגה נכנסת  $k$ . דרגה נכנסת מייצגת את כמות הנסיעות שהסתיימו ביעד מסוים. ניתן לראות שקיימת הסתברות נמוכה לדרגה נמוכה בגרף, מנגד קיימת הסתברות גבוהה יותר להיות בדרגה גבוהה.

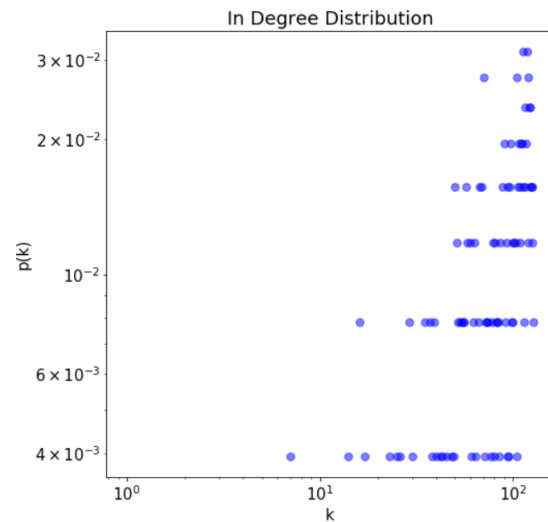
כפי שציפינו התפלגות הדרגות לא מקיימת את תכונת scale-free, זאת מכיוון שבדומה לרשתות גאוגרפיות אחרות שראינו בקורס, התלות בתנאים הטופוגרפיים משפיעים על מבנה הרשת. לעומת זאת, על אף שציפינו לקבל התפלגות דרגות הדומה לגרף של קווי החשמל או כל גרף אחר התלוי בגאוגרפיה, לא קיבלנו התפלגות כזאת. לדעתנו הדבר נובע מהעובדה כי למרות שקיימת תלות גאוגרפית ברשת, יש חשיבות גם לפופולריות של המיקום, זאת בשונה מקווי חשמל אשר חייבים להגיע לכל מקום בארץ (גם לפריפריות). כלומר, ברשת שלנו יש משמעות גם למקומות מרכזיים, אשר הינם בעלי דרגה גבוהה. חשוב גם לציין כי העובדה שהמפעילים החזירו את הקורקינטים לאותם מיקומים בסוף כל יום השפיעה גם היא על מבנה הרשת, ועל כן הרשת שאנו רואים אינה 100% "טבעית".

#### Out Degree Distribution - גרף $P(K)$ :

הגרף מציג מה ההסתברות של מיקום להיות עם דרגה יוצאת  $k$ . דרגה יוצאת מייצגת את כמות הנסיעות שהתחילו במקור מסוים. ניתן לראות כי הדרגה הנכנסת והדרגה היוצאת מתפלגות באופן דומה וכן גם הדרגה הנכנסת הממוצעת שווה לדרגה היוצאת הממוצעת. ממצא זה לא מפתיע, כי לרוב כאשר יוצאים ממקום מסוים, גם חוזרים אליו. נציין כי בהתפלגות זו הסיכוי לדרגה גבוהה מעט נמוכה יותר מאשר ב In Degree Distribution. להערכתנו שוני זה נובע מהעובדה שבמקומות מרכזיים בוצעו נסיעות פנימיות רבות וקצרות יחסית ואילו במקומות פריפריאליים מבוצעות יותר נסיעות אל מקומות מרכזיים.



גרף 7 – התפלגות הדרגות היוצאות



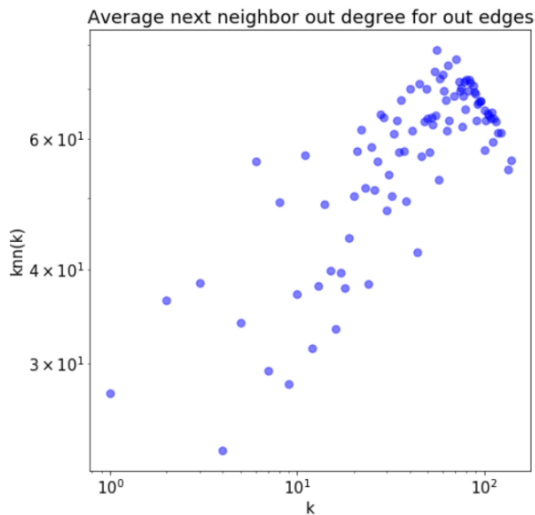
גרף 6 – התפלגות הדרגות הנכנסות

### Average next neighbor degree-Knn(k):

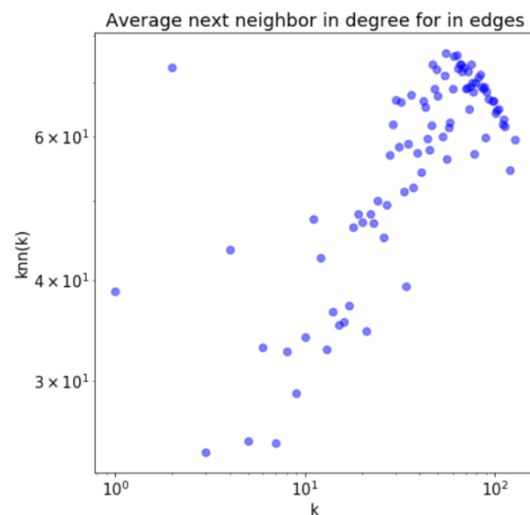
את ניתוח הדרגה הממוצעת ביצענו לפי 4 חלוקות בהתאם לסוג השכנים (מקושרים ע"י קשת יוצאת או ע"י קשת נכנסת) וכן בהתאם לשאלה על פי איזה סוג דרגה מחושב הממוצע לשכנים אלו (דרגה נכנסת או דרגה יוצאת). לאחר ביצוע הניתוחים ראינו כי הגרפים שהתקבלו עבור ממוצע הדרגות היוצאות זהים וכן, הגרפים שהתקבלו עבור ממוצע הדרגות הנכנסות זהים. כמו כן, בכל 4 הגרפים התקבלה מגמתיות דומה. על כן נתייחס רק ל-2 גרפים - גרף אחד עבור הדרגות היוצאות וגרף שני עבור הדרגות הנכנסות (שאר הגרפים נמצאים בנספח מספר 1). בנוסף, מדד האסורטטיביות שלהם ( $\alpha$ ) הוא נמוך וחיובי, ועל כן הרשת אסורטטיבית. כלומר קודקודים בעלי דרגה גבוהה נקשרים יותר זה לזה מאשר לקודקודים בעלי דרגה נמוכה וההפך.

גרפים אלו זה מתנהגים בצורה שונה מכפי שציפינו, מאחר וציפינו לראות כי אנשים נוסעים ממקומות פריפריאליים (עם דרגה נמוכה) למקומות מרכזיים (עם דרגה גבוהה) ובחזרה. אולם נראה כי הדבר אינו כך, בשל מדד האסורטטיביות הנמוך שקיבלנו. נציין כי כפי שניתן לראות בטבלה מספר 1, האסורטטיביות של הדרגה הנכנסת גבוהה משל הדרגה היוצאת, דבר אשר תואם את ההנחה הראשונית שלנו, שמהפריפריה נוסעים יותר למרכז. את תוצאות אלו ניתן להסביר ע"י כך שיותר נסיעות עוברות בין מקומות מרכזיים לעצמם בשל השימוש בקורקינטים לצרכי בילוי ופנאי אשר מתרחשים במקומות מרכזיים או בשל העובדה כי אנשים נוטים לבצע נסיעות למרחקים קצרים בקורקינטים (last mile solution).





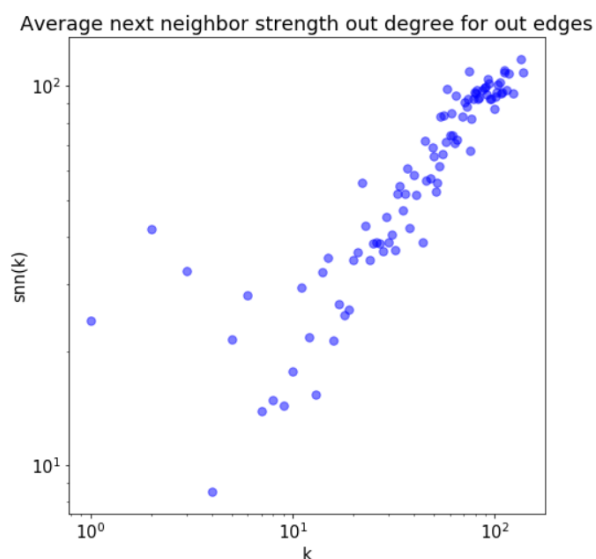
גרף 9 – ממוצע הדרגות היוצאות עבור שכנים המקושרים ע"י קשתות יוצאות



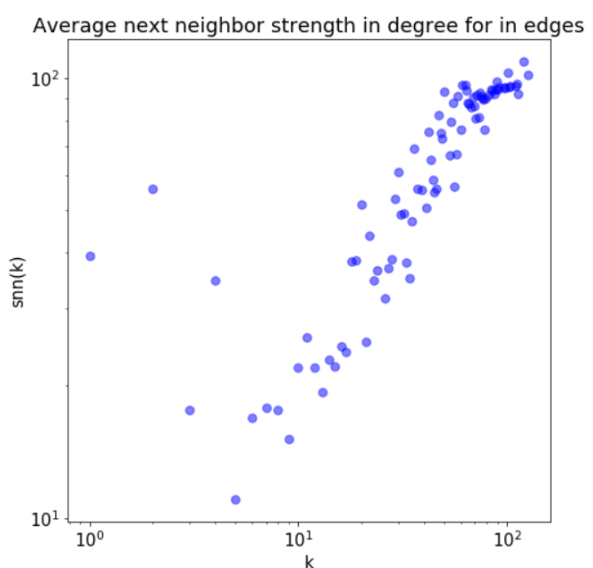
גרף 8 – ממוצע הדרגות הנכנסות עבור שכנים המקושרים ע"י קשתות נכנסות

### Average next neighbor strength-Snn(k)

גרף  $Snn(k)$  מראה את החוזק של השכנים של קודקוד בעל דרגה  $k$ . ככל שממוצע המשקולות המצטברת של הקשרים עם שכניו של קודקוד זה גבוה יותר, כך חוזקו גבוה יותר. גם את ניתוח זה ביצענו לפי 4 חלוקות בהתאם לסוג השכנים (מקושרים ע"י קשת יוצאת או ע"י קשת נכנסת) וכן בהתאם לשאלה על פי איזה סוג דרגה מחושב הממוצע לשכנים אלו (דרגה נכנסת או דרגה יוצאת). אולם גם בניתוח זה התקבלו ממצאים זהים עבור כל סוג משקל (קשת יוצאת וקשת נכנסת), על כן נציג גרף אחד עבור קשת יוצאת וגרף אחד עבור קשת נכנסת (ניתן לראות את שאר הגרפים בנספח מספר 2). בגרף זה ניתן לראות שככל שדרגת הקודקוד גבוהה יותר, כך עולה החוזק של שכניו. גרף זה מתנהג כמצופה, שהרי ככל שמקום הוא מרכזי יותר, כך יגיעו אליו יותר מסלולים ובכל מסלול יהיו נסיעות רבות יותר. כמו כן, מדד זה מתנהג באופן זהה למדד ה- $Knn$  וגם הוא מראה על הנסיעות רבות (יוצאות ונכנסות) בין קודקודים בעלי משקל גבוה.



גרף 11 – חוזק השכנים היוצאים הממוצע עבור שכנים המקושרים ע"י קשתות יוצאות



גרף 10 – חוזק השכנים הנכנסים הממוצע עבור שכנים המקושרים ע"י קשתות נכנסות

הרשת הלא כיוונית :

לאחר שבחנו את הגרף המכוון, ראינו כי המדדים שהתקבלו עבור ה-in degree ועבור ה-out degree דומים וההבדלים בניהם זניחים. לכן, כניתוח נוסף לרשת בנינו גרף לא מכוון, (וממושקל) שבאמצעותו נוכל לבחון אלגוריתמים שונים של community detection ולהגיע לתובנות נוספות. כמו בגרף המכוון, כל צומת מייצג נקודה (מיקום) בעוד כל נסיעה תהווה קשר ללא חשיבות לכיוון הנסיעה.

המדדים שהתקבלו עבור הרשת הלא מכוונת :

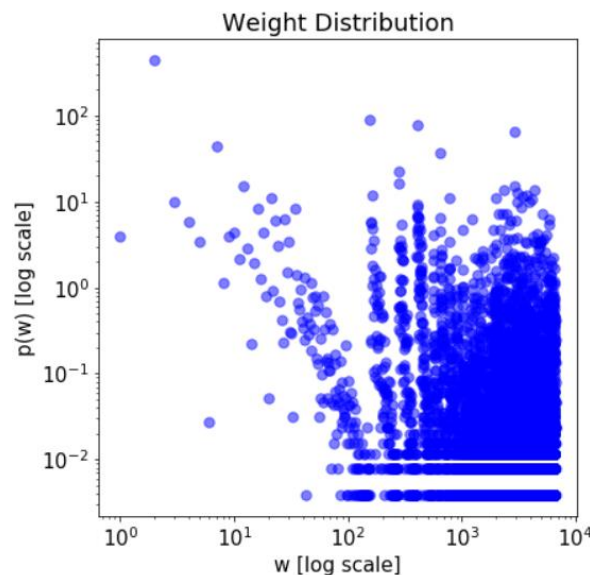
Number of Nodes	Number of Edges	Average Degree	Large component	Clustering Coefficient	Average Path Length
255	6579	51.6	254	0.58	inf

טבלה 2 – מבנה הרשת הלא כיוונית

Large component - בגרף הלא מכוון, גודל הקומפוננטה הגדולה הוא 254, כלומר כמעט 100% מהצמתים בגרף מחוברים זה לזה.

:P(w)

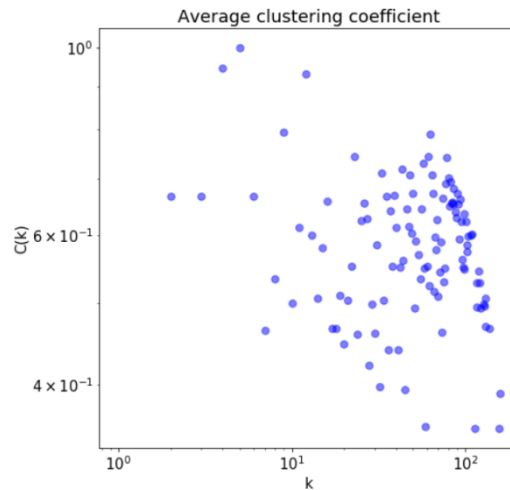
ניתן לראות את התפלגות המשקלים בקנה מידה log-log. לא ניתן לזהות התפלגות משקלים מוכרת.



גרף 12 - התפלגות המשקלים

: Clustering Coefficient

Clustering coefficient = 0.58, לעומת  $C(\text{rand}) = 0.299$ . ניתן לראות כי מרבית הקודקודים נמצאים בין 0.4 ל-0.6. כמו כן, מלבד מספר דרגות נמוכות, בהן מקדם ההתקבצות גבוה יחסית, לא ניכרת מגמה ברורה כתלות בדרגה של הקודקודים. מכך ניתן להסיק כי לא קיים קשר מובהק בין הדרגה לבין ממוצע מקדם ההתקבצות של כל דרגה.



גרף 13- מקדם ההתקבצות

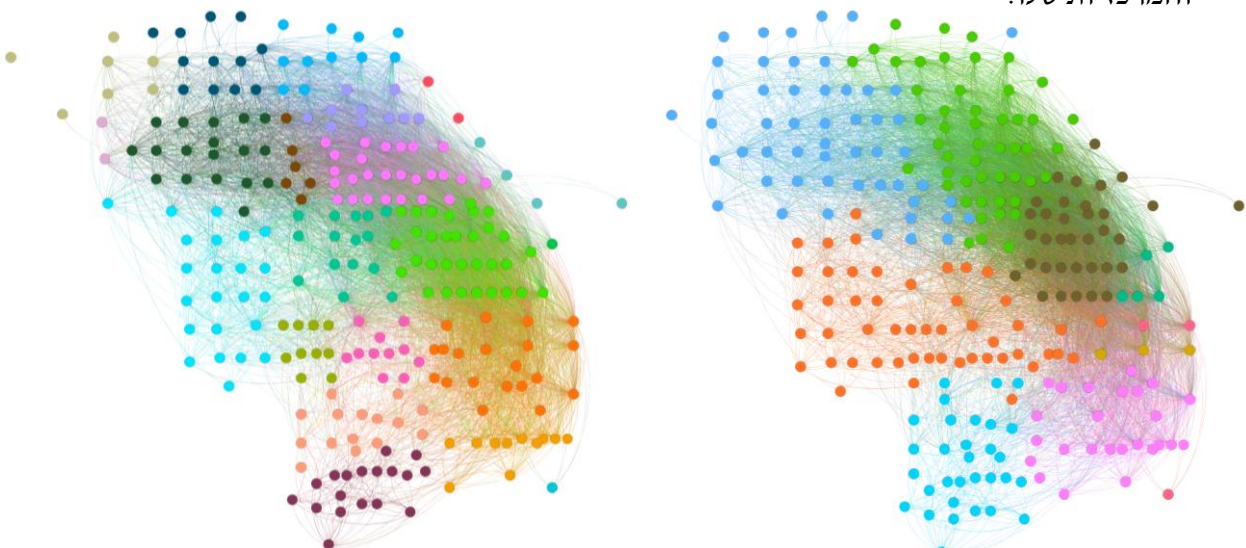
## Community Detection:

לאחר ניתוח מדדי הרשת, ביצענו מספר אלגוריתמים של community detection. להן מדדי ה-Modularity של האלגוריתמים:

KL communities generator	Greedy Modularity	Girvan Newman	Louvain
0.274	$2.293 \cdot 10^{-5}$	$9.93 \cdot 10^{-5}$	<b>0.565</b>

טבלה 3 – מדד הModularity של אלגוריתמי communities השונים

כפי שניתן לראות, מדד ה-Modularity של מרבית האלגוריתמים נמוך יחסית. כמו כן, מרביתם חילקו את הרשת לקהילה אחת גדולה וקהילה נוספת עם מספר צמתים בודדים, פרט לאלגוריתם Louvain שהחזיר חלוקה שונה למספר קהילות וכן מדד ה Modularity שלו היה הגבוה ביותר ביחס לשאר האלגוריתמים, ועל כן בחרנו להציג רק אותו (ניתן לראות את תוצאות שאר האלגוריתמים בנספח מספר 3). כפי שניתן לראות בגרפים 14 ו15, החלוקה של Louvain לקהילות דומה מאוד לחלוקת האזורים בעיר, כאשר הוא מחבר אזורים "קטנים" לקהילה אחת. ממצא זה אינו מפתיע, כיוון שלכל אזור יש את ההתנהגות שלו והנטיעות שלו וכן את רמת ה"פופולריות" והמרכזיות שלו.



גרף 15 – חלוקת הרשת לפי אזורי העיר שיקגו

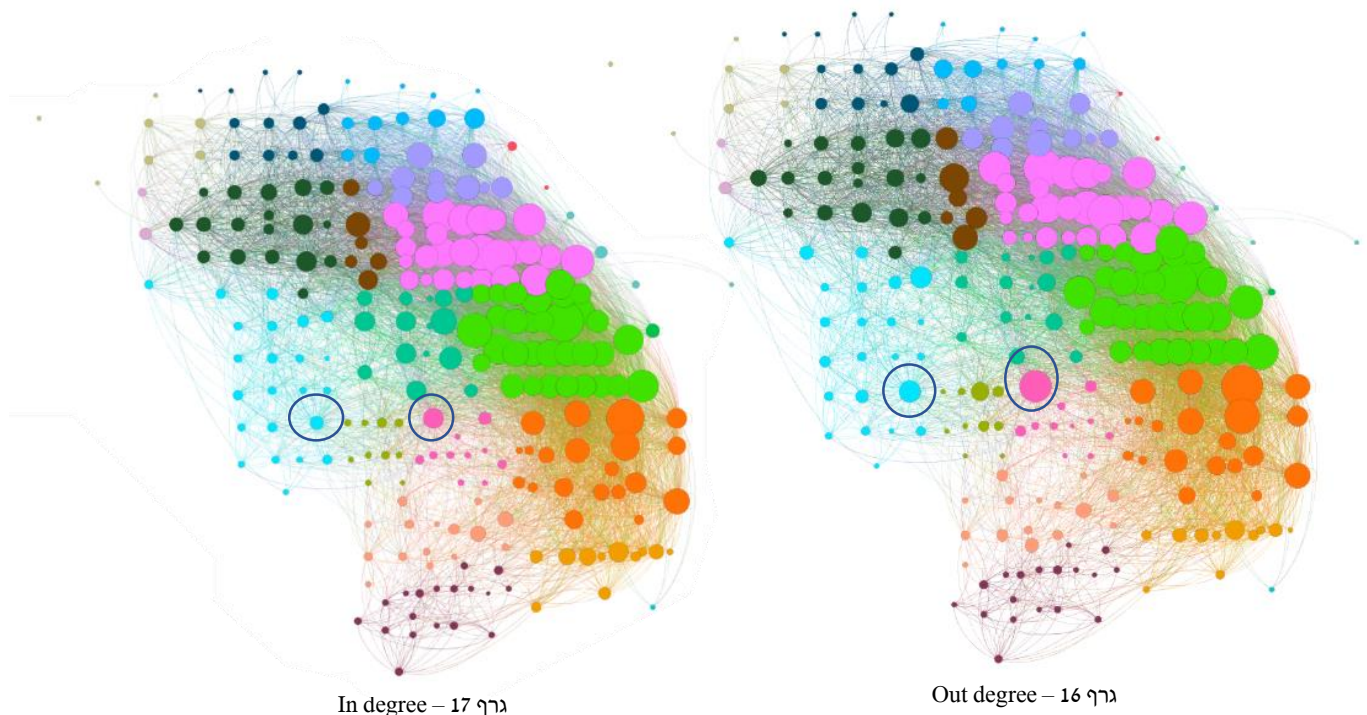
גרף 14 – חלוקת הרשת לקהילות לפי אלגוריתם Louvain

החל משלב זה, על מנת שנוכל לזהות בצורה ברורה את הקודקודים (בהתאם למיקומם הגאוגרפי), בכל הגרפים שיוצגו, כל קודקוד יסומן בצבע על פי האזור שלו בעיר, וכן כל אזור יקבל צבע קבוע בכל הגרפים. זאת בשל החשיבות שיש למיקום הגאוגרפי של כל קודקוד ברשת.

### ניתוח מדדי Centrality:

לאחר ניתוח communities ביצענו ניתוחי centrality, בהתאם למדדי centrality שונים.

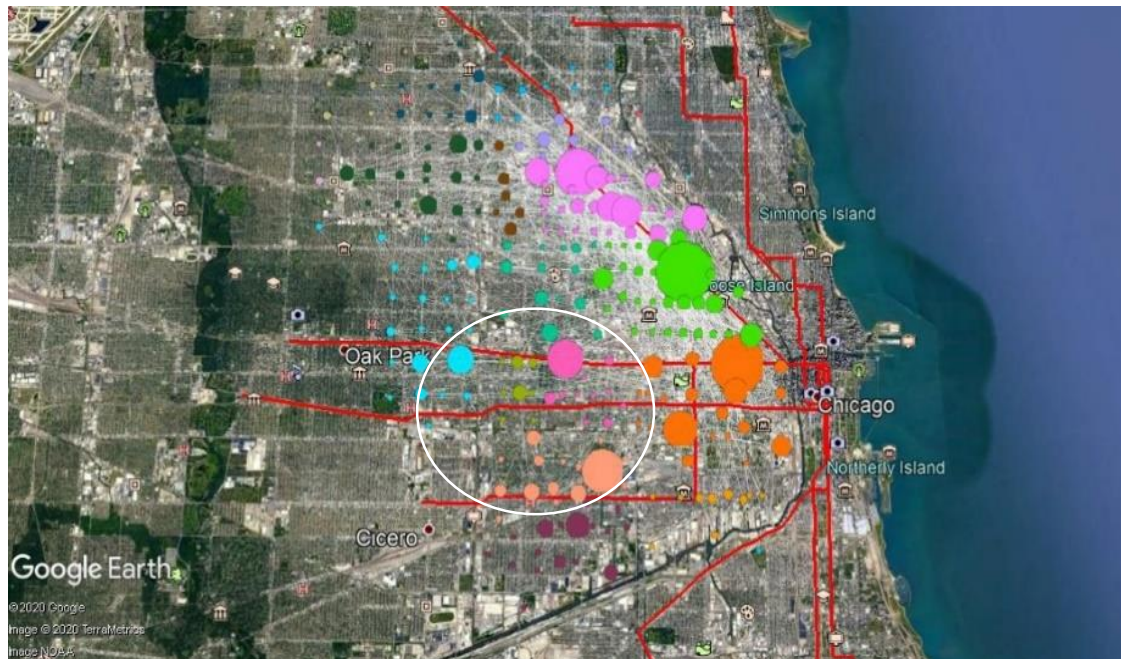
**Degree**- מדד זה, באופן לא מפתיע, מחלק את הגרף בהתאם לאזורים, לפי פריפריה ומרכז בעיר. כאשר באזורי המרכז הדרגה גבוהה יותר. על מנת להבין את התנהגות הרשת, ביצענו את הניתוח הן ביחס לדרגה היוצאת והן ביחס לדרגה הנכנסת. כפי שניתן לראות באיורים 16 ו-17, המדד כמעט זהה בין הדרגה הנכנסת לדרגה היוצאת. ממצא זה תואם את הממצא אשר ראינו כבר בעת ניתוח התפלגות הדרגות, אשר הראה התפלגות כמעט זהה הבין הדרגה הנכנסת לדרגה היוצאת. כלומר, לרוב, כאשר נוסעים ממקום מסוים עם קורקינט, גם נחזור אליו עם קורקינט. נציין כי ראינו 2 צמתים אשר ממוקמים יחסית בפריפריה, אך בעלי Out degree גבוהה יחסית (מוקפים בעיגול כחול), בהמשך ננסה להבין מה מייחד את הצמתים האלו ביחס לשאר הצמתים בפריפריה.



**Betweenness**- מדד Centrality נוסף אשר בחרנו להתמקד בו הוא מדד ה-Betweenness. ברשתות גיאוגרפיות ורשתות תחבורה קודקודים בעלי מדד Betweenness גבוה הם קודקודים המגשרים בדרך כלל בין קלסטרים שונים ו"מחברים" את הרשת. ציפינו לראות קודקודים באזורי מעבר בין הפריפריה למרכז אשר יקבלו ערך Betweenness גבוה. ניתן לראות כי אכן קיימים קודקודים כאלו (מוקפים בעיגול לבן בגרף למטה), כאשר שניים מהקודקודים הנ"ל כבר בלטו כבר בשלב ניתוחי הדרגה. אולם מרבית הקודקודים בעלי מדד ה-Betweenness נמצאים באזור המרכז. רצינו לבחון האם ישנו קשר בין הקודקודים בעלי מדד ה-Betweenness הגבוה, על כן ייצאנו את הרשת לתוכנת Google Earth אשר הציבה את הרשת על מפת העיר שיקגו. כפי שניתן לראות בגרף



למטה, הנקודות הניל ממוקמות קרוב לתחנות רכבת של קווי מטרו מרכזיים בשיקגו (מסומנים באדום בגרף) ו/או בתוך פארקים. נקודות אלו מעניינות וחשובות מכיוון שהן מרכזות כמות תעבורה משמעותית וקריטיות לחיבוריות הגרף ולהבנת השימוש בקורקינטים.

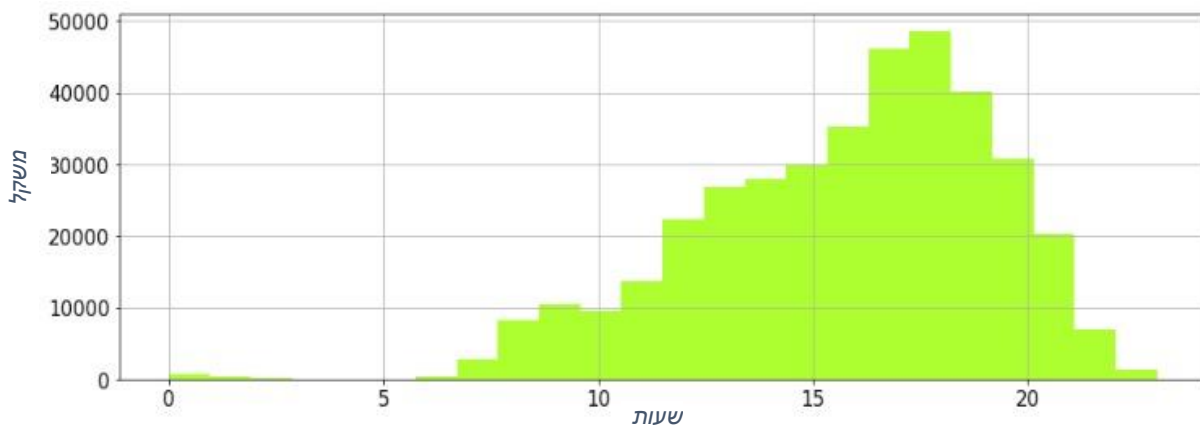


גרף 18 – מדד ה Betweenness על מפת העיר שיקגו

## ניתוח תתי גרפים

### ניתוח רשת לפי ימים ושעות :

לאחר ביצוע ניתוח זה רצינו לראות האם ניתן לזהות מגמות שונות בהתאם לזמנים שונים במהלך היום ובמהלך השבוע. על כן, סיננו את הנתונים כך שיציגו רק את אמצע השבוע (שני-רביעי) ובנינו 4 תתי רשתות מכוונות, כאשר כל 2 תתי רשתות מחלקות את הרשת המלאה לשעות הבוקר 7:00-15:00 ולשעות הערב 15:00-23:00 ולכל טווח שעות יצרנו גרף לדרגה הנכנסת וגרף לדרגה היוצאת. נציין כי בשעות הלילה כמעט ולא בוצעו נסיעות (כפי שניתן לראות בגרף מספר 19) ועל כן בחרנו להזניח את השעות האלו.

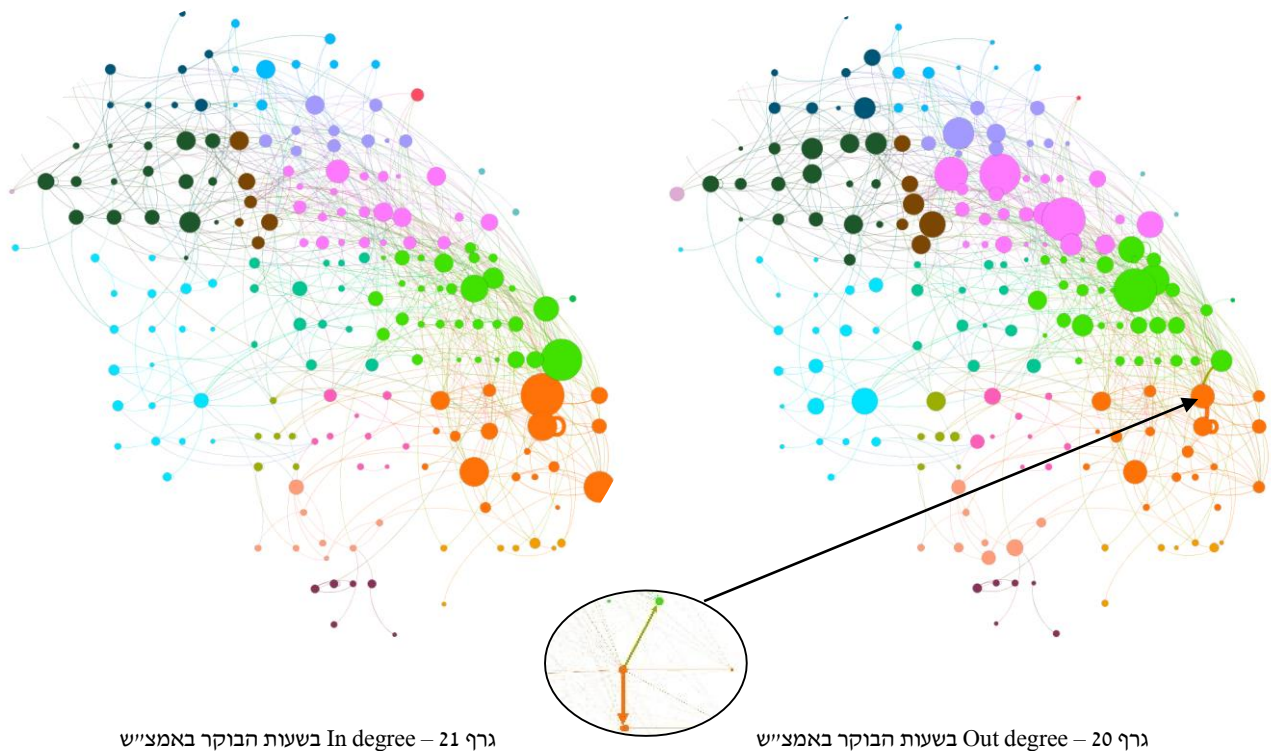


גרף 19 – התפלגות הנסיעות לפי שעות באמצע השבוע

מהתפלגות הנסיעות לפי שעות ניתן לראות שני דברים בולטים:

ראשית, ככול הנראה, הקורקינטים לא מהווים כלי תחבורה לטובת "יציאות" של בני ה-20-30, זאת מכיוון שכמעט ואין נסיעות בשעות הלילה. בנוסף, ניתן לראות כי חלק ניכר מהנסיעות מתבצעות בשעות הצהריים-ערב, בהמשך נחזור לממצא זה על מנת להבין מי הם המשתמשים העיקרים של הקורקינטים ומה מטרתם.

כאשר הצגנו את גודל הקודקודים בהתאם לדרגה, הנכנסת והיוצאת, ראינו כי בשעות הבוקר ניתן לראות באופן בולט כי הדרגה היוצאת בקודקודים אשר יושבים על תחנות המטרו גדולה מאוד<sup>2</sup>. ממצא זה תואם את הממצאים שראינו כבר בניתוחי מדדי ה-Centrality, וכן את השערותנו, כי פעמים רבות השימוש בקורקינטים הוא לצורך Last mile solution. כלומר, אנשים מהפריפריה לוקחים את המטרו בכדי להגיע למרכז העיר, אולם בתוך המרכז, על מנת לחסוך מעבר בין קווי מטרו או על מנת לחסוך הליכה, הם משתמשים בקורקינטים ל"מייל" האחרון בדרך לעבודה. לאור ממצא זה, בצירוף מדד ה-Betweenness אשר ראינו מקודם. ניתן להסיק כי תחנות הרכבת מהוות מיקומי מפתח, ועל כן לדעתנו כדאי להתקין בהן עמדות הטענה וכן חניות מוסדרות לקורקינטים. ניתוח נוסף שתאם את השערת ה Last mile solution היה ניתוח של גרף שעות הבוקר לפי מדד ה-Betweenness אשר הראה גם הוא ערך גבוה לקודקודים אשר נמצאים על קווי המטרו (ראה נספח 4).

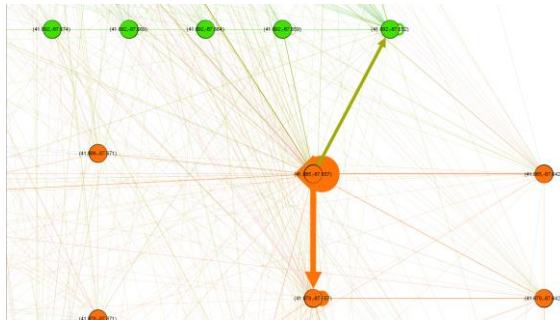


כאשר הסתכלנו על הרשת של שעות הבוקר ביחס לשעות הערב, ראינו כי כמות הנסיעות בשעות הערב גדולה באופן משמעותי, זאת בהתאם להתפלגות הנסיעות לפי שעות שצגנו בגרף מספר 19. בנוסף, כאשר השווינו בין הדרגה הנכנסת לדרגה היוצאת בשעות הערב, לא ראינו הבדלים משמעותיים (ניתן לראות את הגרפים בנספח מספר 5). ממצא זה הפתיע אותנו, מכיוון שציפינו

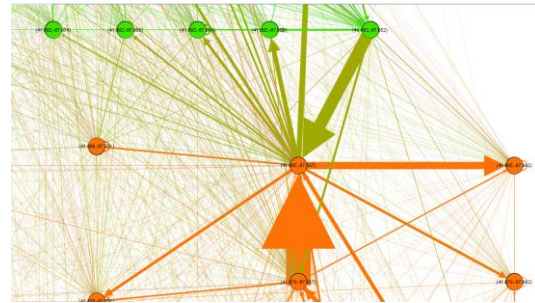
<sup>2</sup> את קווי המטרו ניתן לראות בגרף מספר 18



לראות הרבה נסיעות לכיוון תחנת הרכבת (בכיוון ההפוך משעות הבוקר). אולם, כאשר הסתכלנו על הרשת יותר "מקרוב", ראינו כי אכן קיימים מסלולים כאלו, כפי שניתן לראות בגרפים 22 ו-23, אך הם כנראה "נבלעים" בשל הכמות הגדולה של הנסיעות בשעות הערב.



גרף 23 – מסלול נסיעה בבוקר באמצ"ש

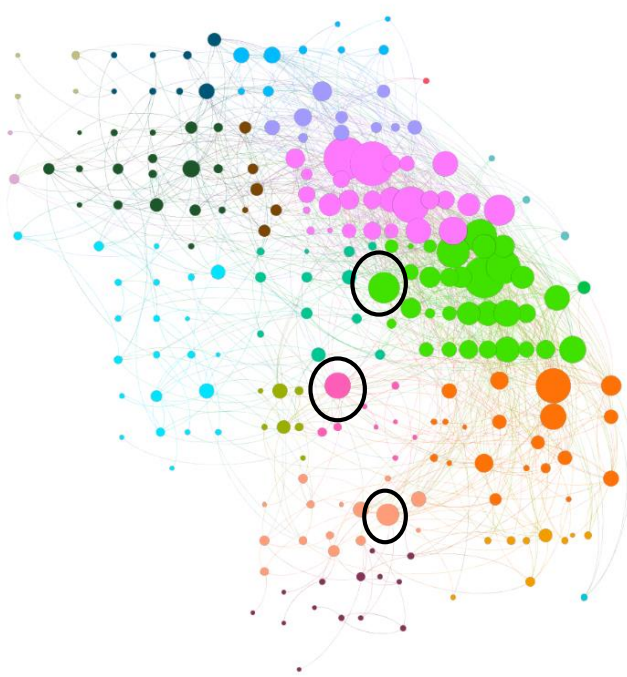


גרף 22 – מסלול נסיעה בערב באמצ"ש

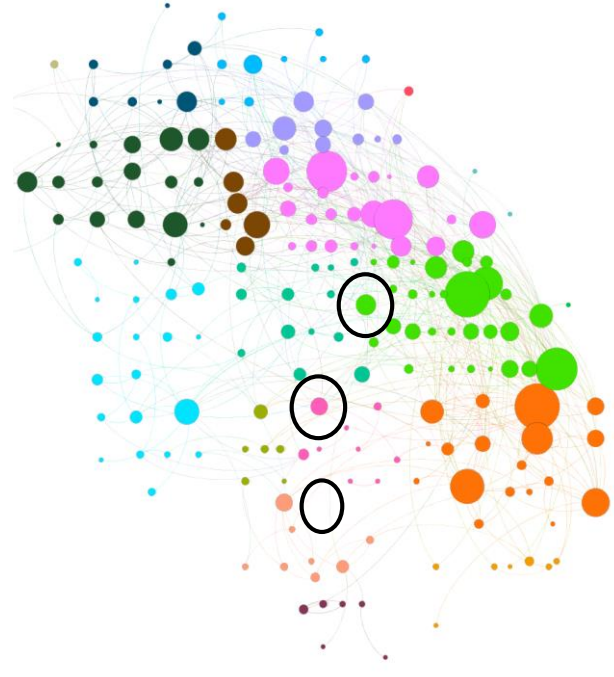
### ניתוח הרשת לפי אמצע שבוע וסוף שבוע:

על מנת להבין את מטרות הנסיעות בצורה טובה יותר, החלטנו לפצל את הרשת הלא מכוונת לשני תתי גרפים, כאשר תת גרף אחד מכיל את הנסיעות בסוף השבוע (שישי-ראשון) והשני את הנסיעות באמצע השבוע (שני-רביעי). את יום חמישי בחרנו להזניח גם בכדי לקבל כמות "שווה" של ימים וגם מכיוון שחמישי הוא לא ממש סופ"ש אך גם לא ממש יום עבודה רגיל. בהתחלה לא ראינו הבדלים משמעותיים (ניתן לראות את הגרפים בנספח 6). לאחר שביצענו סינון נוסף, לפי שעות, ראינו כי אכן בשעות הערב לא קיים הבדל מובהק בין אמצע השבוע לסוף השבוע (ניתן לראות את הגרפים בנספח 7), אולם בשעות הבוקר ראינו שני הבדלים משמעותיים. ראשית, ניתן לראות כי כמות הנסיעות בשעות הבוקר גדולה משמעותית בסוף השבוע לעומת שעות אלו באמצע השבוע. בנוסף, ראינו כי ישנם מספר מיקומים אשר מקבלים דרגה גבוהה בסוף השבוע, אך באמצע השבוע הדרגה שלהם יחסית נמוכה. בבדיקה על מפה, ראינו כי מיקומים אלו הינם פארקים גדולים בשיקגו (מוקפים בעיגול בגרף), אשר מהווים מקום משחק ובילוי ביחוד בסוף השבוע (ראה סרטון מצורף<sup>3</sup>). ממצא זה תואם את להתפלגות שעות הנסיעות הכללית, אשר מראה כמות גדולה של נסיעות בשעות הצהריים – ערב, שעות המשחק של בני הנוער. לדעתנו, ממצא זה, ניתן להסיק כי אחד השימושים העיקריים של הקורקינטים הוא לצורכי פנאי ומשחק של בני הנוער.

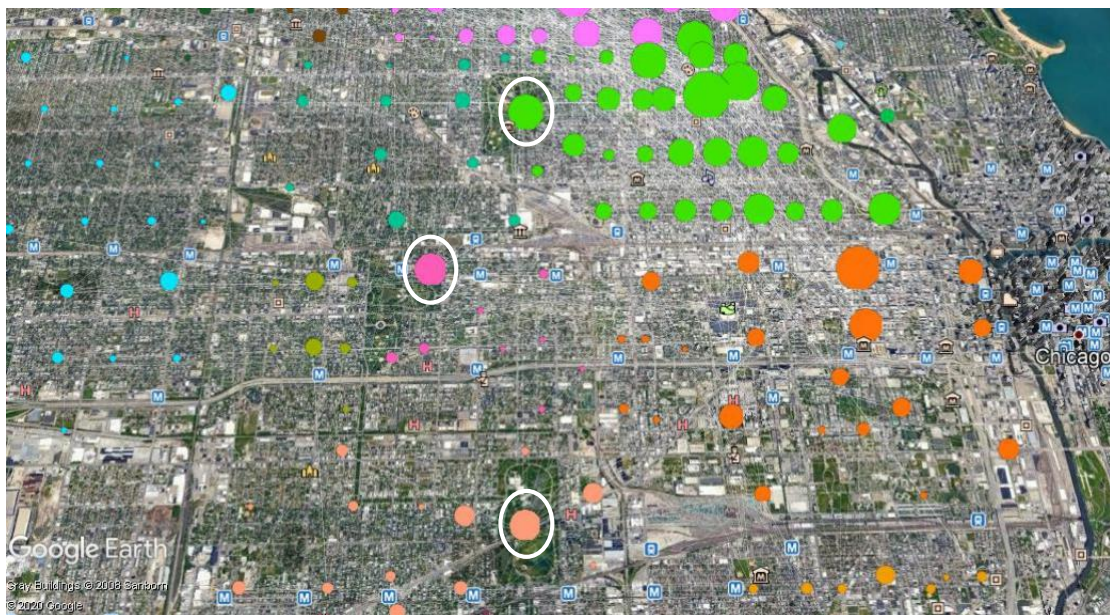
<sup>3</sup> לינק לסרטון: <https://drive.google.com/file/d/1Q4Oh-A2jCkoAJHQv5BVDQWtP4VV8Q1ZK/view?usp=sharing>



גרף 25 – סופ"ש בשעות הבוקר



גרף 24 – אמצ"ש בשעות הבוקר



גרף 26- סופ"ש, שעות בוקר על מפת העיר שיקגו

### ניתוח מסלולים ואורכי נסיעות ברשת:

דבר נוסף שרצינו לבדוק הוא את מאפייני מסלולי הנסיעות, זאת במטרה למצוא מסלולים משמעותיים בעלי כמות גדולה של נסיעות, וכן את האזורים בהם מרוכזים מסלולים אלו. לכן, הסרנו מסלולים בעלי משקל נמוך מ-50. כפי שניתן לראות בגרף מספר 27, המסלולים המשמעותיים ביותר מרוכזים במרכז העיר. על כן, לדעתנו, יש לסלול באזור זה נתיבים לרכבים דו גלגליים, על מנת להפוך את הנסיעה לבטוחה יותר ולאפשר זרימה של קורקינטים ואופניים מבלי לסכן את הולכי הרגל.

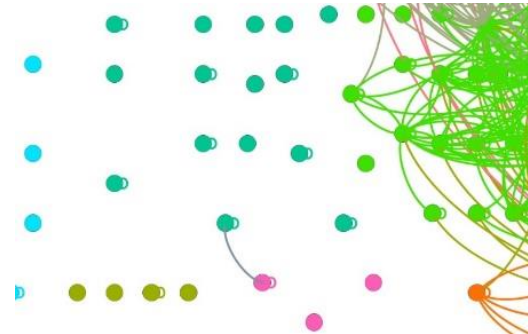


גרף 27 – מסלולים בעלי משקל גבוה מ-50

ממצא נוסף שראינו הוא כי הקשרים בעלי המשקלים הגבוהים ביותר, ביחוד באזורי הפריפריה, הם קשרים מקדקוד לעצמו. כאשר סיננו את הדאטה לפי מרחקי נסיעות (מעל למייל) מרבית מהקשרים הללו נעלמו. לדעתנו, בשל העובדה כי עלות השימוש בקורקינט הינה לפי זמן ולא לפי מרחק, ניתן להניח כי אין סיבה לא "לסיים נסיעה" במידה והקורקינט אינו בשימוש. על כן, אנו מסיקים כי נסיעות אלו הינן נסיעות למטרות פנאי ומשחק.



איור 2- אחרי סינון מרחקי נסיעות



איור 1- לפני סינון מרחקי נסיעות

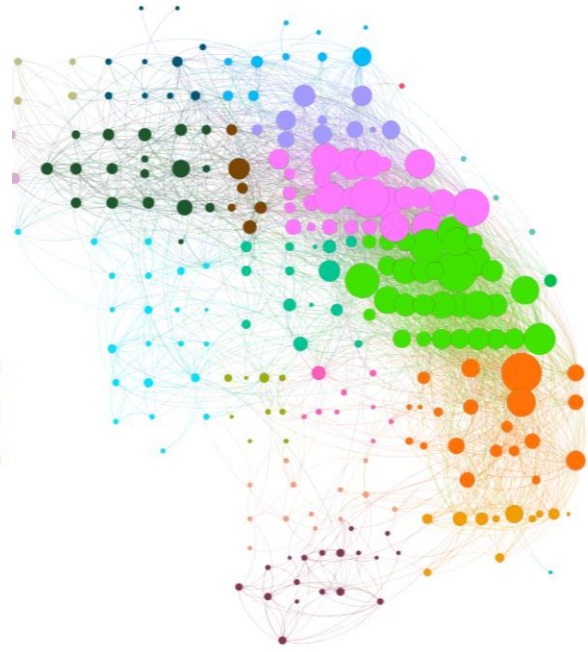
### ניתוח הרשת לאורך זמן:

הניתוח האחרון שביצענו לנתוני הפרויקט הוא התנהגות הרשת לאורך זמן. הפרויקט ארך 4 חודשים, ועל כן יצרנו 4 גרפים בהפרשי זמן יחסית קבועים, כאשר כל גרף מייצג שבוע מהפילוט. כפי שניתן לראות בגרפים 28-31, כמות הצמתים הלכה והתמעטה ככול שהפרויקט התקדם, אולם ניתן לראות כי מבנה הרשת עצמה לא השתנה, אלא פשוט כמות הנסיעות פחתה. חשוב לשים לב כי מרבית הצמתים ש"נעלמו" עם הזמן נמצאים בפרפריה. נציין, כי מכיוון שמפעילי הקורקינטים שהשתתפו בפילוט החזירו את הקורקינטים לאותן נקודות מוצא בכל יום, הדבר לא קרה בשל חוסר זמינות של הקורקינטים בנקודה זו או אחרת.





גרף 29 – שבוע 2, 218 צמתים



גרף 28 – שבוע 1, 240 צמתים



גרף 31 – שבוע 4, 104 צמתים



גרף 30 – שבוע 3, 204 צמתים

להערכתנו, ניתן לשייך את התופעה למספר גורמים שונים:

1. קלקול בקורקינטים- יכול להיות שבמהלך הפרויקט התקלקלו חלק מהכלים ולא כולם תוקנו, כתוצאה מכך התמעטו הקורקינטים אשר היו זמנים למשתמשים מה שהוביל לירידה בכמות הנסיעות.
  2. חוסר עניין – אחת המסקנות שלנו מהניתוחים שביצענו היא שחלק גדול מהנסיעות בוצעו לצורכי פנאי ולא לצורכי עבודה וכדו. כתוצאה מכך, כמו כל צעצוע חדש, לאחר מספר חודשים ההתלהבות מהקורקינטים פחתה, (ביחוד כאשר השימוש עולה כסף), מה שהוביל לירידה בכמות הנסיעות. השערה זו נתמכת על ידי העובדה כי מרבית הקודקודים שנעלמו נמצאים באזורי הפריפריה, ולפי הניתוח אשר ביצענו מקודם, מרבית הנסיעות בצמתים אלו הן נסיעות מקודקוד לעצמו, כאשר לדעתנו נסיעות אלו הן נסיעות לצורכי פנאי ומשחק.
  3. יכול להיות שבמהלך הפרויקט (ביחוד בין החודש שלישי לרביעי) החלו המפעילים ל"קפל" את הכלים וכתוצאה מכך, פחתו כמות הכלים הזמינים. השערה זו נתמכת על ידי הירידה המשמעותית בכמות הנסיעות בין החודש השלישי לרביעי (לקראת זמן סיום הפרויקט).
- על מנת לרדת לשורש העניין, עלינו לקבל עוד מידע על מאפייני המשתמשים בכל נסיעה – כגון גיל המשתמשים, וכן מידע אודות כמות הקורקינטים שהיו תקינים וזמינים לשימוש בכל שלב בפרויקט.

## סיכום:

בעבודה זו חקרנו בכלים של ניתוח רשתות את השימוש בכלי תחבורה שיתופיים בסביבה ללא רציפים. מטרת העבודה היא להבין את אופי השימוש בקורקינטים את התנהגות הרשת והמקומות בהם נדרשת התערבות חיצונית. לצורך כך השתמשנו בנתוני פיילוט על קורקינטים שיתופיים שנערך בעיר שיקאגו בשנת 2019. הנתונים הראשוניים הכילו מידע על כ-700 אלף נסיעות שהתבצעו בשטח מוגדר במשך תקופה של ארבעה חודשים. במהלך העבודה ביצענו ניתוחים שונים על רשתות מלאות או חלקיות על בסיס נתונים אלו.

הניתוח בוצע באופן דינאמי כך שהעמקנו בכל פעם שזיהינו תופעות מעניינות העולות מהנתונים. ראינו שהרשת המלאה היא מקושרת מאוד ואסורטטיבית. ישנם יותר קשרים בין הנקודות באזור המרכז שהן בעלות תעבורה גדולה יותר ומעט מאוד נסיעות ארוכות מהפריפריה עד המרכז.

מניתוח הדרגות רואים כי בניגוד למה שציפינו, אין מקומות שהם מקורות ומקומות שהם יעדים באופן מובהק. למרות ההתערבות בהחזרת הכלים למקומם, הרשת מאוזנת ואין הרעבה ו/או הצטברות כלים במקומות בהם יש יותר או פחות תעבורה.

איתור הקהילות לא העלה תובנות מפתיעות, החלוקה לקלאסטרים תואמת במידה רבה את החלוקה לאזורים המקובלים בעיר.

תופעה מעניינת שעלתה מניתוח מדד ה-Betweenness centrality היא כי נקודות בעלות מדד centrality גבוה ממוקמות בסמוך לתחנות רכבת מרכזיות. תופעה זו מדגישה את השימוש בקורקינטים שיתופיים בתחבורה משלימה לרכבת. נוסעים מאזורים שונים מבצעים נסיעות אל

ומתחנות הרכבת כך שהתחנות מחברות את הרשת. עובדה זו מעלה את הצורך בהקמת עמדות חניה והטענה בסמוך לרכבת.

ניתוח רשתות חלקיות בסינון לפי שעות וימי נסיעה מראה התנהגות שונה בין סוף השבוע ואמצע השבוע ובין הערב לבוקר במהלך השבוע, כצפוי. זיהינו מסלולים בעלי משקל גבוה בין מרכזי תעסוקה מרכזיים לתחנות רכבת בכיוונים הפוכים בין הבוקר לערב. כך שלמרות שהרשת מאוזנת באופן כללי קיימת תופעה של הרעה והצטברות בשעות מסוימות במהלך היום שצריך גם לה לתת מענה בצורת חניות ועמדות הטענה. בנוסף, ראינו כי בסופי השבוע אזורי הפארקים מרכזים יותר תעבורה וקודקודים הממוקמים בפארקים מקבלים דרגה גבוהה יותר באופן יחסי. כמו כן ראינו כי מרבית הנסיעות מתבצעות בשעות הערב. שתי המגמות האלו מראות כי אופי השימוש בקורקינטים הוא בעיקר בהקשר של פנאי.

מניתוח משקלי הנסיעות עלה כי מרבית הנסיעות מתבצעות בין אזורי המרכז, ועל כן באזורים אלו יש צורך במסלולי נסיעה ייחודיים לרכבים דו-גלגליים, על מנת לאפשר זרימה של התנועה ונסיעה בטוחה, מבלי להפריע למרקם התחבורה הקיים בעיר.

בניתוח הנתונים לפי תקופות, ניתן לראות שהחל ממספר שבועות לתוך הפיילוט הרשתות קטנות משבוע לשבוע מבחינת מספר הקשרים (כמות הנסיעות) וגם מאבדות קודקודים מסוימים, פרט בפריפריה, ו"מתמרכזות". להערכתנו, הדבר מרמז על תופעה של טרנד חולף וירידה במספר הנסיעות הקצרות יחסית אשר מקושרות לנסיעות מבחן או הנאה מחזקת את המסקנה כי מרבית השימוש בקורקינטים הוא לצורכי פנאי ומשחק.

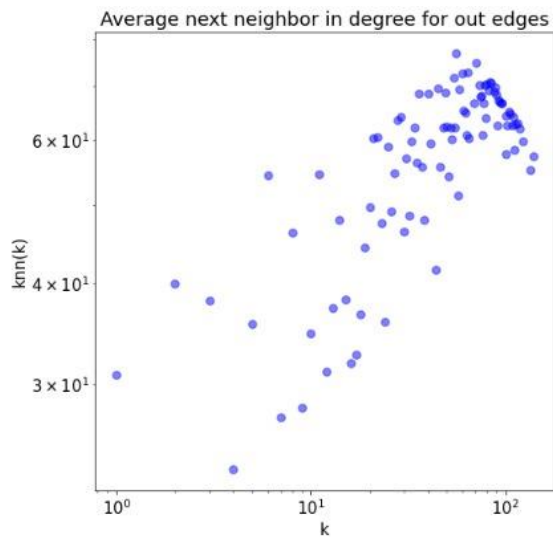
על אף המסקנות שעלו מהפרויקט, נציין כי לדעתנו קיימים מספר פערים בפרויקט, ובדאטה עליו הוא התבסס. על כן, על מנת לקבל תוצאות יותר מובהקות, יש לבצע ניתוחים נוספים, על דאטה גדול יותר (בפרויקט שלנו היו רק 255 צמתים), המכיל יותר פרטי מידע, בפרט פרטים כמו גיל המשתמשים וכן כמות הקורקינטים הזמינים בכל שלב. בנוסף, על מנת לקבל תמונת מצב יותר מהימנה, יש לתת לרשת להיבנות ללא התערבות המפעילים, כך שהקורקינטים יעברו ממקום למקום בצורה "טבעית".

לדעתנו ניתן להשתמש בתוצאות ומסקנות הפרויקט ולהשליכן על ערים נוספות בהן קיים שימוש בקורקינטים שיתופיים, על מנת לאפשר שימוש בטוח ויעיל ולשלבם במרקם חיי התחבורה.

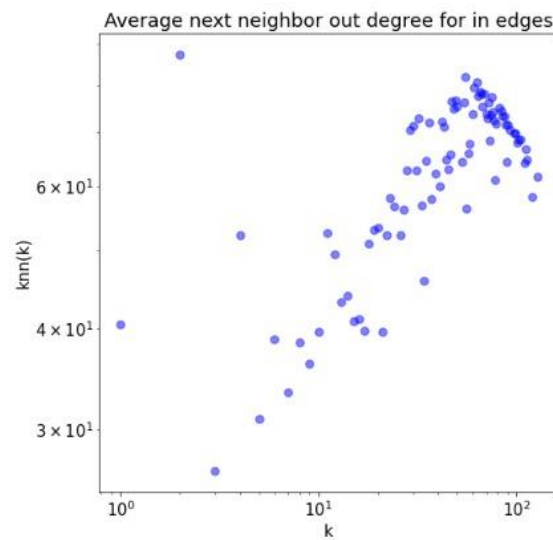


## נספחים:

### נספח מספר 1 – KNN

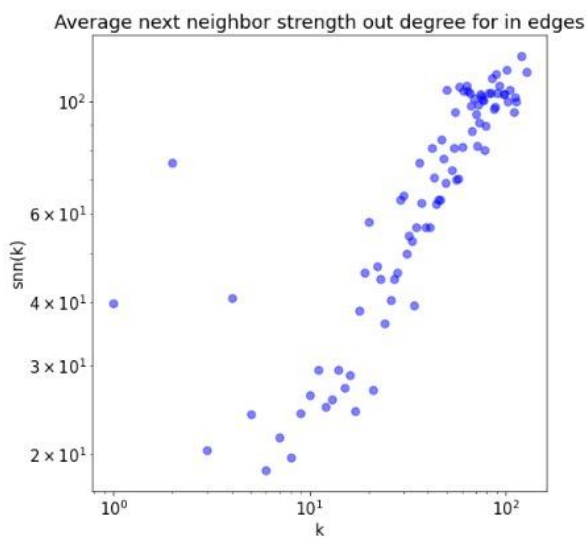


גרף 33 – ממוצע הדרגות הנכנסות עבור שכנים המקושרים ע"י קשתות יוצאות  
 $\langle r \rangle = 0.144$

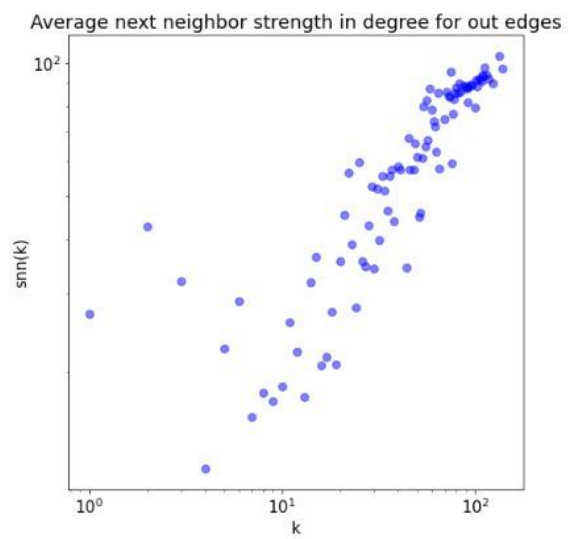


גרף 32 – ממוצע הדרגות היוצאות עבור שכנים המקושרים ע"י קשתות נכנסות  
 $\langle r \rangle = 0.223$

### נספח מספר 2 – SNN

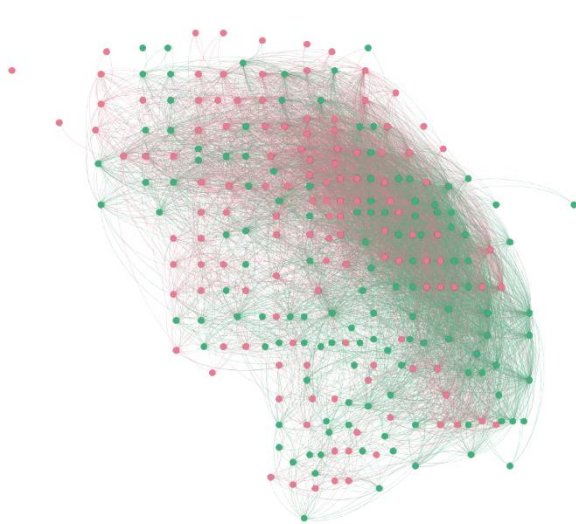


גרף 35 – חוזק השכנים היוצאים הממוצע עבור שכנים המקושרים ע"י קשתות נכנסות

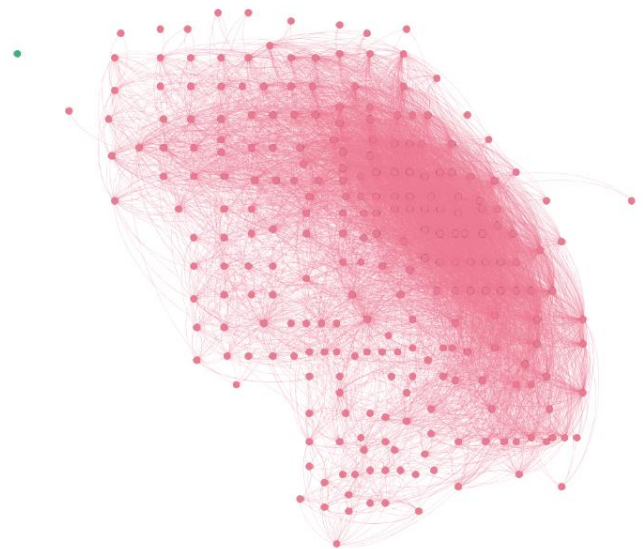


גרף 34 – חוזק השכנים הנכנסים הממוצע עבור שכנים המקושרים ע"י קשתות יוצאות

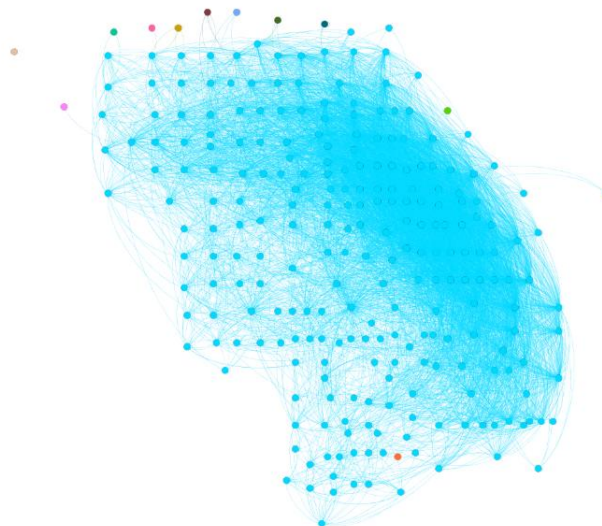
נספח מספר 3 – תוצאות אלגוריתמי Community detection :



גרף 37 - KL

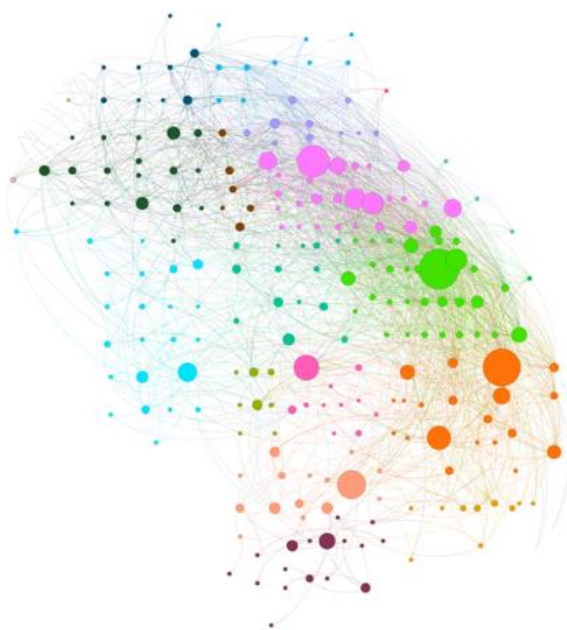


גרף 36 - Greedy Modularity



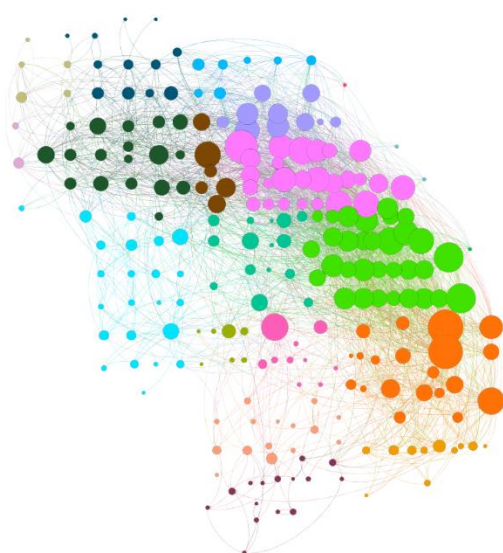
גרף 38 - Girvan Newman

**נספח מספר 4 – Betweenness של גרף שעות הבוקר של אמצע השבוע**

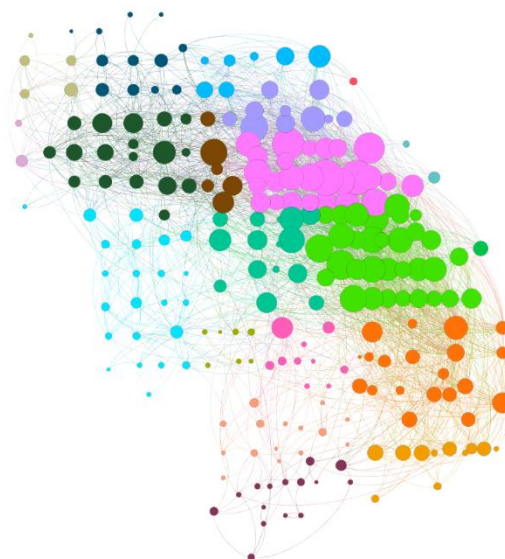


גרף 39- Betweenness שעות הבוקר, אמצע השבוע

**נספח מספר 5 – אמצ"ש שעות הערב, In מול Out :**

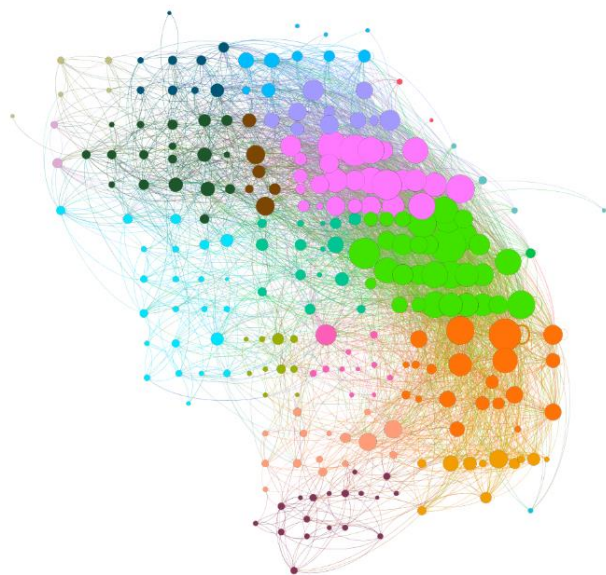


גרף 41 – אמצ"ש, שעות הערב, Out degree

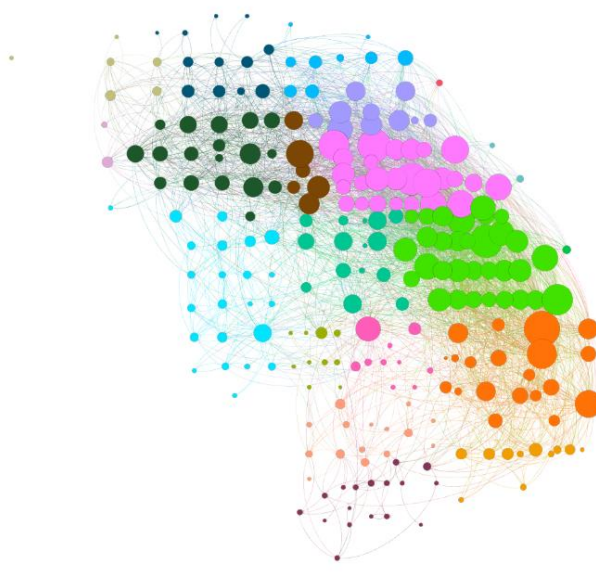


גרף 40 – אמצ"ש, שעות הערב In degree

**נספח מספר 6 – אמצ"ש מול סופ"ש בכל השעות**



גרף 43 – סופ"ש



גרף 42 – אמצ"ש

**נספח מספר 7 – אמצ"ש בשעות הערב מול סופ"ש בשעות הערב**



גרף 45 – סופ"ש בשעות הערב



גרף 44 – אמצ"ש בשעות הערב