

קורס רשתות מורכבות פרויקט מסכם

Academic Networks



מרצה: ד"ר אסף אלמוג

עוזר הוראה: עמית שרייבר

תאריך הגשה: 18.06.2021

עומר לוי 305215071

עומרי ג'ובאני 037098985

תקציר מנהלים:

בעבודה זו חקרנו רשתות שיתופי פעולה באקדמיה. בעוד שהרבה עבודות החוקרות את המחקר באקדמיה מתמקדות בציטוטי מאמרים וניתוח רשתות ציטוטים, אנו בחרנו להתמקד בשיתופי פעולה בין חוקרים ובין אוניברסיטאות. מטרתנו במחקר רשתות אלה הייתה להבין את טיב שיתופי הפעולה בין חוקרים ובין מוסדות שונים ולהבין האם ניתן להסיק מסקנות שונות משיתופי פעולה אלה כגון; האם שיתופי פעולה רבים מצביעים על טיב החוקר/המוסד האקדמי, אילו מוסדות וחוקרים נוטים לשתף יותר פעולה ולמה, והאם ניתן לזהות קהילות אקדמיות מהנתונים.

לאחר בחינת מדדי הרשתות ותכונותיהם, הסקנו שאכן כפי ששיערו וציפינו, רשתות שיתופי הפעולה מתנהגות כרשתות חברתיות המתארות קשרים בין בני אדם. בנוסף, הגענו למסקנה כי כמות שיתופי הפעולה אינה בהכרח מדד לציון טיב החוקר והצטיינותו בתחום המחקר בו הוא עוסק. זיהינו תופעה בה בדיספלינות ממדעי החיים כגון ביולוגיה, וביואינפורמטיקה מתבצעים שיתופי פעולה רבים יחסית בעבודת מחקר אחת. אולם, נדרשת עבודת מחקר נוספת ומקיפה יותר לבחינת השערה זו. גילינו שכמות שיתופי הפעולה של מוסדות אקדמיים הוא אכן מדד טוב לבחינת טיב המוסד והשפעתו. המוסדות היוקרתיים והמובילים בעולם אכן משתפים פעולה עם אוניברסיטאות רבות ומבצעים פעילות מחקרית ענפה בתחומים רבים. לאוניברסיטאות אלה השפעה ניכרת על המחקר באקדמיה והתפתחות תחומי מחקר רבים בעולם. לסוף, בחנו לעומק את אוניברסיטת תל אביב, גילינו שבשונה מרשת החוקרים הגדולה, רשת זו הינה דיסאסורטיבית והשערנו לגבי ממצא זה היא שאוניברסיטת תל אביב הינה אוניברסיטה לימודית בה ישנם הרבה שיתופי פעולה בין חוקרים מובילים לבין חוקרים עם רקורד אקדמי דל יותר (מאסטרנטים ודוקטורנטים). כמו כן, מצאנו כי הפקולטה למדעי המחשב היא הפקולטה המובילה הן בפרסום מאמרים והן בשיתופי הפעולה וכי ארבעה מתוך חמישה החוקרים המובילים בשיתופי פעולה באוניברסיטת תל אביב הם מהפקולטה למדעי המחשב.

תוכן עניינים:

1.	אפיון הנתונים	3
2.	ניתוח מאקרו אודות הנתונים	4
3.	ניתוח רשת שיתופי פעולה בין חוקרים – שנת 2019	7
3.1.	חקר מדדי הרשת	7
3.2.	ניתוח Quantity = Quality?	10
4.	רשת שיתופי פעולה בין אוניברסיטאות	13
4.1.	חקר מדדי הרשת	13
4.2.	ניתוח Quantity = Quality?	18
5.	אוניברסיטת תל אביב case study	19
5.1.	ייצוג האוניברסיטה ברשת החוקרים – עיקרי הממצאים	19
5.2.	ייצוג רשתות באמצעות Gephi	21

1. אפיון הנתונים:

הנתונים אשר שימשו אותנו בפרויקט זה נלקחו מאתר Kaggle וכוללים מידע על 4,894,081 מאמרים שונים מתחומים שונים. מרבית הנתונים הינם מ-70 השנים האחרונות. השנים בהן כמות הנתונים הינה המשמעותית ביותר הינן שנים החל משנת 2010 ועד לשנת 2019 כולל, עם ממוצע של 283,700 מאמרים לשנה.

עבור כל מאמר, קיימים שדות שונים המכילים מידע על המאמר כגון תקציר, שנת הפרסום, אילו מאמרים ציטטו מאמר זה, מס' ציטוטים, ומידע על מחברי המאמר (שם ומוסד אקדמי).

עבודות רבות נעשו על רשתות שבחנו קשרים בין מאמרים כאשר המאפיין המקשר היה הציטוטים המופיעים בכל מאמר (מי ציטט את מי).

במחקר שלנו בחרנו דווקא להתמקד ולבחון את שיתופי הפעולה בין החוקרים ובין האוניברסיטאות. לשם כך התמקדנו בניתוח 2 רשתות מרכזיות:

1. רשת לשיתוף פעולה בין חוקרים
2. רשת לשיתוף פעולה בין אוניברסיטאות

עבור רשת החוקרים, השדות המרכזיים בהם השתמשנו הינם מס' זיהוי חוקר, שם החוקר, והמוסד האקדמי של אותו חוקר. כל צומת ברשת הינה חוקר וכל קשת ברשת הינה שיתוף פעולה עם חוקר אחר מהרשת. משקל כל קשת l_{AB} , הינה כמות שיתופי הפעולה של חוקר A עם חוקר B. על מנת קבל תובנות נוספות, הצלבנו מידע זה עם מידע נוסף מאתר Google scholar המכיל מידע נוסף על מאמרים וחוקרים שונים. מאתר זה השתמשנו בעיקר בשדה h-index (עליו יפורט בהמשך) לקבוע את טיב החוקר ותרומתו האקדמית.

עבור רשת האוניברסיטאות, השדה העיקרי הוא שם האוניברסיטה כך שכל צומת ברשת הינה אוניברסיטה וכל קשת הינה שיתוף פעולה עם אוניברסיטה אחרת. משקל כל קשת $l_{u_{AB}}$ הינה כמות שיתופי הפעולה בין אוניברסיטה A עם אוניברסיטה. עבור ניתוח שמות האוניברסיטאות היינו צריכים לעשות עבודת עיבוד טקסט מכיוון שהאוניברסיטאות הוצגו כטקסט בלבד וללא מספר זיהוי ייחודי לאוניברסיטה. לצורך כך השתמשנו בשיטות שונות של similarity כדי לאחד פקולטות שונות של אותו המוסדות תחת שם מוסד אחד וזיהוי מוסדות זהים שהופיעו בצורות שונות בטקסט. על מנת לקבל תובנות נוספות, הצלבנו את המידע שקיבלנו מניתוח הרשת עם דירוגי אוניברסיטאות מוכרים. הראשון- מדד שנגחאי והשני- מדד InfluenceRank. מטרתנו הייתה לבדוק האם שיתופי פעולה רבים אכן

עומדים בקנה אחד עם דירוג טיב האוניברסיטה ודירוג השפעת האוניברסיטה על המחקר האקדמי העולמי.

2. ניתוח מאקרו אודות הנתונים:

בשל גודלו העצום של הדאטה ומגבלות חומרה, לא ניתן לחקור בצורה מיטבית רשתות המכילות את כלל המידע הקיים עבור כל השנים הקיימות בדאטה. לכן, על מנת לבצע מחקר מיטבי על הרשתות החלטנו להתמקד בתחילה בעשר השנים המלאות האחרונות הקיימות בדאטה קרי 2010-2019. מתוך שנים אלו לקחנו מדגם של 4 שנים מייצגות: 2010, 2013, 2016, 2019. עבור כל שנה, בנינו רשת מלאה של קשרי החוקרים באותה השנה. מטרת השוואה זו, לבחון האם התנהגות הרשתות זהה עבור כל שנה או לא. במידה וכן, עובדה זו תאפשר לנו להתמקד בשנה ספציפית ולבצע מחקר מיטבי יותר עבור כמות דאטה קטנה יותר, שכן התנהגות הרשת בכל שנה הינה זהה ואין ערך מוסף בהסתכלות על כלל השנים יחד.

שנה	כמות צמתים	כמות קשתות	דרגה ממוצעת	% שינוי בדרגה הממוצעת
2010	359,383	832,372	4.63	-
2013	455,658	1,146,028	5.03	8.6%
2016	560,265	1,589,391	5.67	6.98%
2019	613,406	1,764,452	5.75	1.04%

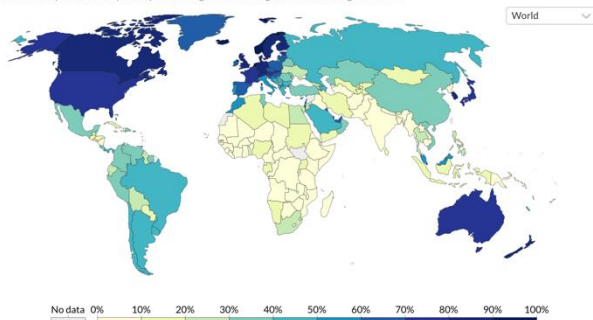
טבלה 1 - נתוני רשתות חוקרים לפי שנה

מנתונים אלו המופיעים בטבלה מס' 1 ניתן לראות כי קיימת גדילה דועכת ככל שמתקדמים בשנים. גדילה זו ניתן להסביר בעיקר בשל התקדמות טכנולוגית וגישה לדאטה, ירידה בעלויות איסוף ואחסון דאטה והתפתחות האינטרנט בשנים אלו.

באיור מספר 1 ניתן לראות את ההתפתחות הטכנולוגית של האינטרנט בעולם עפ"י אחוז האוכלוסייה המשתמשת באינטרנט. באיור זה ניתן להבחין בגידול בין השנים אשר דועך עם הזמן. גם באיור 2 ניתן לראות שוב את התופעה של התקדמות האינטרנט וחיבור האנשים בעולם בשנים אלו. התפתחות האינטרנט גורמת לשיתוף פעולה ושיתוף מידע בצורה קלה יותר ולכן באופן טבעי מגדילה את כמות שיתופי הפעולה.

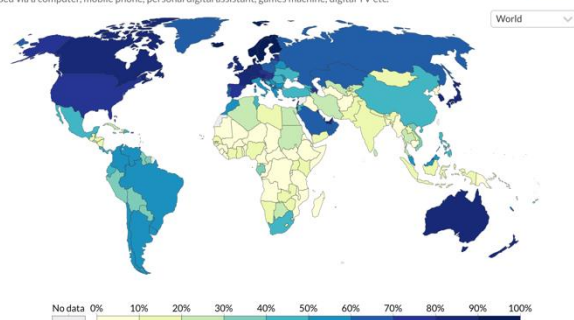
Share of the population using the Internet, 2010

All individuals who have used the Internet in the last 3 months are counted as Internet users. The Internet can be used via a computer, mobile phone, personal digital assistant, games machine, digital TV etc.



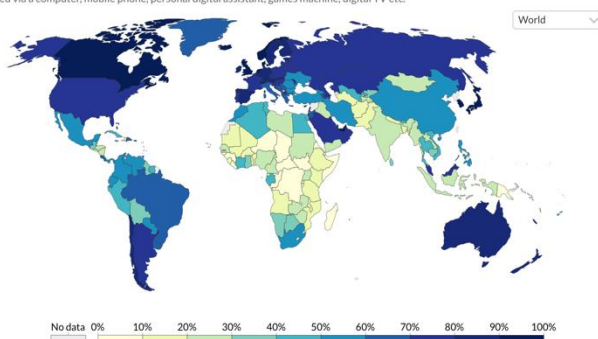
Share of the population using the Internet, 2013

All individuals who have used the Internet in the last 3 months are counted as Internet users. The Internet can be used via a computer, mobile phone, personal digital assistant, games machine, digital TV etc.



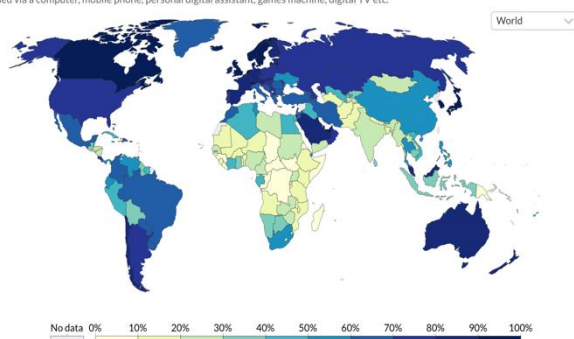
Share of the population using the Internet, 2016

All individuals who have used the Internet in the last 3 months are counted as Internet users. The Internet can be used via a computer, mobile phone, personal digital assistant, games machine, digital TV etc.



Share of the population using the Internet, 2017

All individuals who have used the Internet in the last 3 months are counted as Internet users. The Internet can be used via a computer, mobile phone, personal digital assistant, games machine, digital TV etc.

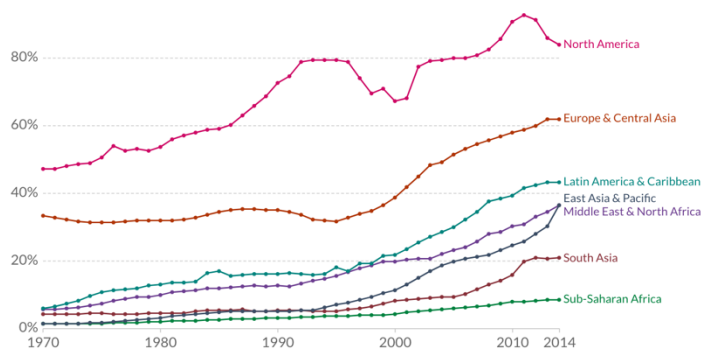


איור 1 - אחוז האוכלוסייה המשתמשת באינטרנט לפי שנים 2010, 2013, 2016, 2017

Gross enrollment ratio in tertiary education, 1970 to 2014

Total enrollment in tertiary education, regardless of age, expressed as a percentage of the total population of the five-year age group following on from secondary school leaving.

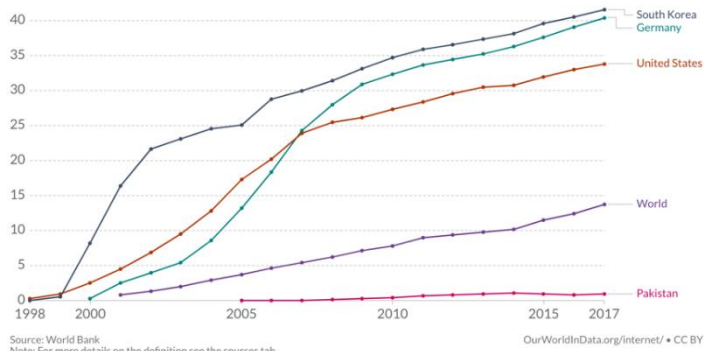
+ Add country



Broadband subscriptions per 100 people, 1998 to 2017

Broadband subscriptions refer to fixed subscriptions to high-speed access to the public Internet (a TCP/IP connection), at downstream speeds equal to, or greater than, 256 kbit/s.

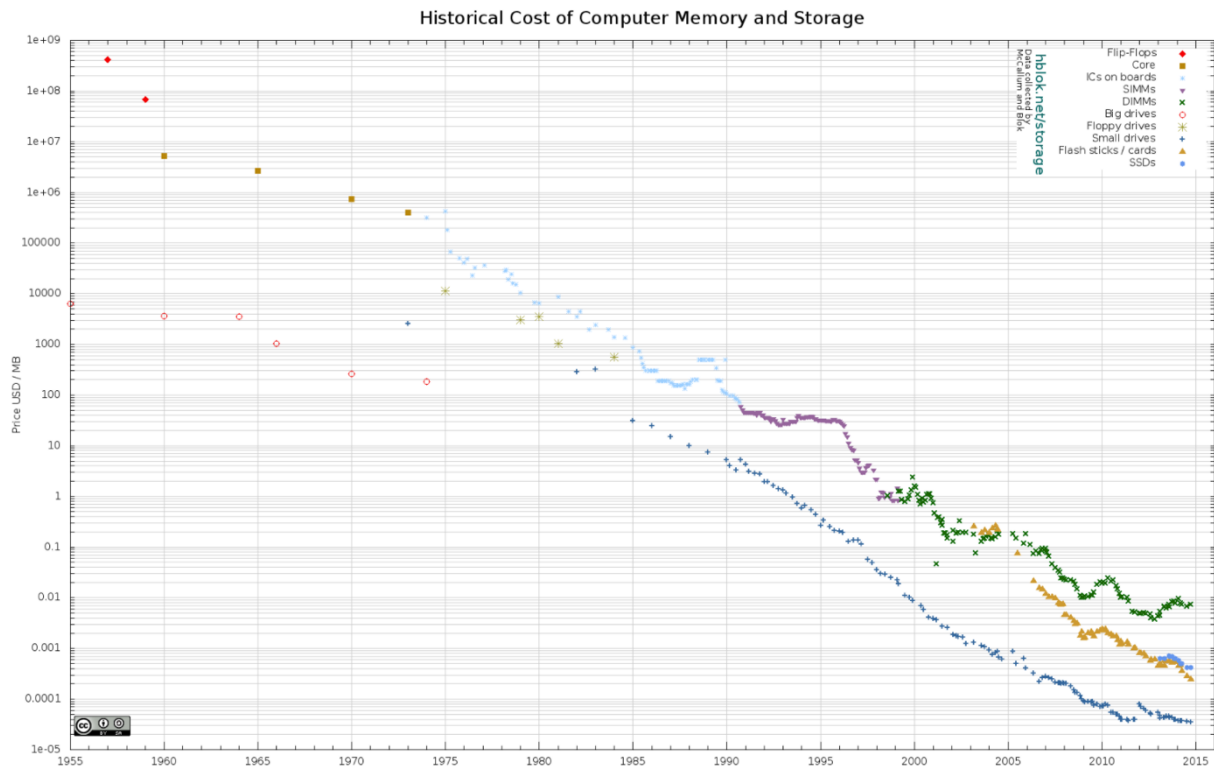
+ Add country



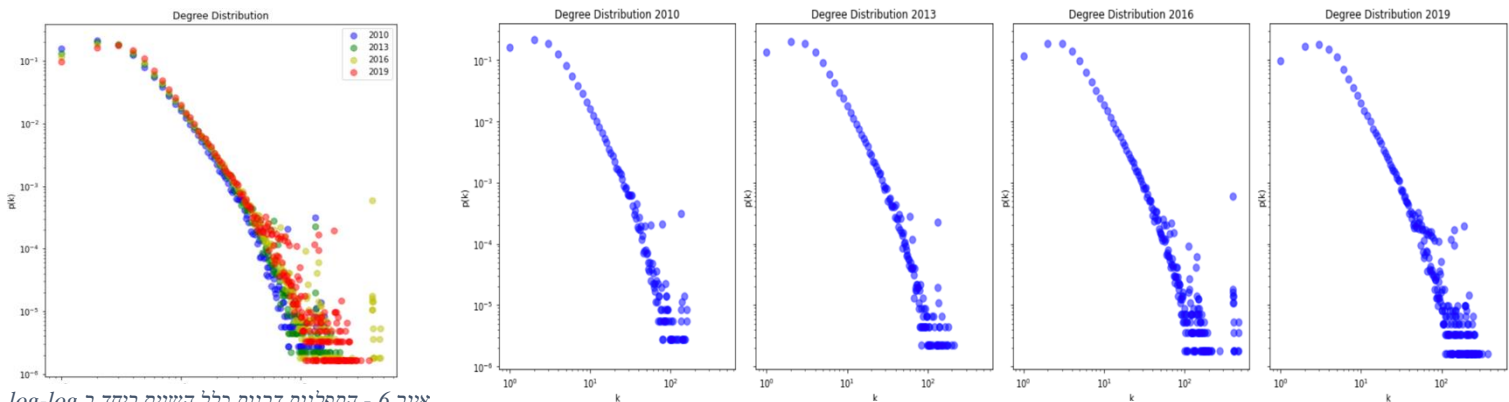
איור 3 - יחס הרשמה ברוטו ללימודי השכלה גבוהה

איור 2 - כמות מניינים לפס הרחב באינטרנט פר 100 אנשים

השערה נוספת לגידול זה יכולה לנבוע אולי גם מהנגשת ההשכלה הגבוהה בשנים אלו לאוכלוסייה רבה יותר והצורך בתארים אקדמיים בעולם שהופך טכנולוגי יותר כפי שניתן לראות באיור 3. עלייה בכמות הנרשמים להשכלה הגבוהה גורמת לחוקרים רבים יותר וכפועל יוצא פרסומים רבים יותר. באיור 4 ניתן לראות את הטרנד בירידת עלויות אחסון וזיכרון לאורך השנים ובפרט בשנים הני"ל. אחסון דאטה בצורה זולה יותר מאפשר יצירת ארכיונים ואיסוף מידע (במקרה שלנו מאמרים) באופן זול וזמין יותר.



איור 4 – עלות היסטורית של אחסון וזיכרון מחשבים



איור 6 - התפלגות דרגות כלל השנים בידוד ב $\log\text{-}\log$

איור 5 - התפלגות דרגות עבור כל שנה בנפרד ב $\log\text{-}\log$

באיורים 5 ו-6 ניתן לראות כי למעט הגידול הקיים בדרגה בממוצעת בין השנים, התנהגות הרשת בכל שנה די דומה. לכן, עבור המשך ניתוח הרשת בחרנו להתמקד בשנה האחרונה בדאטה עבורה קיים דאטה מלא- שנת 2019.

3. ניתוח רשת שיתופי פעולה בין חוקרים – שנת 2019:

3.1. חקר מדדי הרשת:

כפי שתואר לעיל, את שיתופי הפעולה בין החוקרים בחרנו לתאר כרשת לא-כיוונית כאשר צומת הינה חוקר וקשת הינה שיתוף הפעולה בין חוקרים שונים. משקל הקשת הינו כמות הפעמים בה שיתפו החוקרים פעולה.

1. כמות צמתים (N): 613,406

2. כמות קשתות (L): 1,764,452

3. דרגה ממוצעת: 5.75

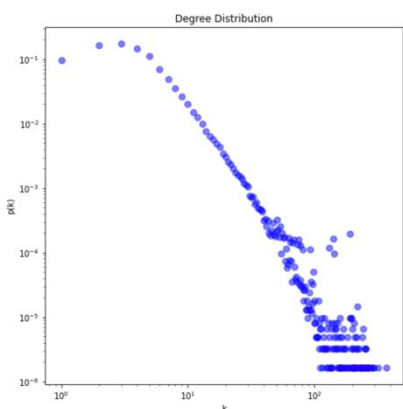
4. התפלגות דרגות ברשת $(P(k))$:

ניתן לראות באיור 7 כי הרשת שלנו מתנהגת כרשת scale free. מכיוון שהרשת שלנו מתארת קשר בין אנשים, נתון זה עולה בקנה אחד עם ההשערה כי לרוב, קשרים בין אנשים כדוגמת רשתות חברתיות ושיתופי פעולה מתנהגות באופן זה ולכן נתון זה אינו מפתיע.

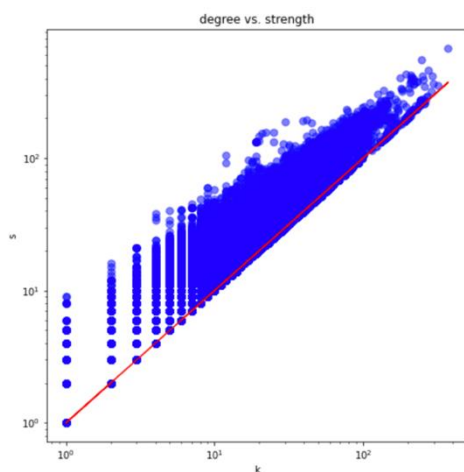
5. התפלגות משקל w וחוזק s ברשת $(P(w), P(s))$:

באופן זהה להתפלגות הדרגות, גם התפלגות המשקולות והחוזק ברשת מראה התנהגות

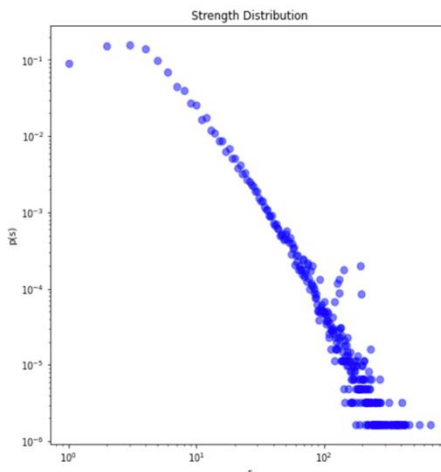
scale-free כפי שניתן לראות באיורים 8 ו-9



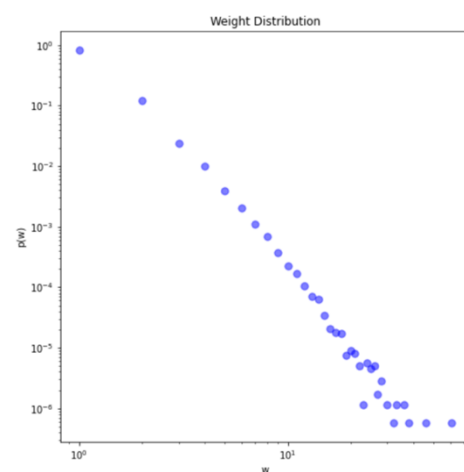
איור 7 - התפלגות דרגות ב log-log scale



איור 10 - התפלגות דרגה vs חוזק ב log-log scale



איור 9 - התפלגות חוזק ב log-log scale



איור 8 - התפלגות משקלים ב log-log scale

על מנת להבין את הקשר בין התפלגות הדרגות להתפלגות החוזק ברשת, יצרנו גרף המכיל את ההתפלגויות של s ושל k יחד, כפי שניתן לראות באיור 10. מגרף זה ניתן להבין כי במרבית המקרים החוקרים משתפים פעולה אחד עם השני יותר מפעם אחת כאשר בדרגות הנמוכות שיתופי הפעולה החוזרים גבוהים וככל שעולים בדרגה כמות שיתופי הפעולה החוזרים הולכת ויורדת (על האלכסון האדום בגרף ניתן לדמות מה היה קורה אילו כל חוקר היה משתף פעולה פעם אחת בלבד עם חוקר אחר). ניתן אולי להסביר זאת כי החוקרים בעלי הדרגה הגבוהה, משתפים יותר פעולה עם חוקרים **שונים** ולכן אולי מטעמי זמן, אין יכולת לבצע שיתוף פעולה חוזר באותה השנה.

6. **Average path length <math>\langle l \rangle:**

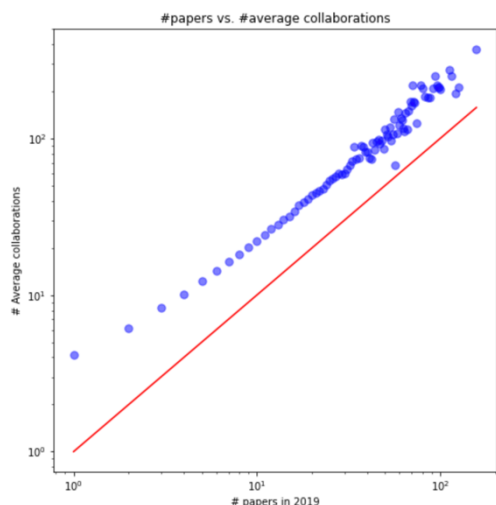
מכיוון שהרשת שלנו אינה קשירה לא ניתן למדוד פרמטר זה באופן אנליטי, לכן נאמר כי האורך הינו אינסופי.

7. **Average clustering coefficient $C(k)$:**

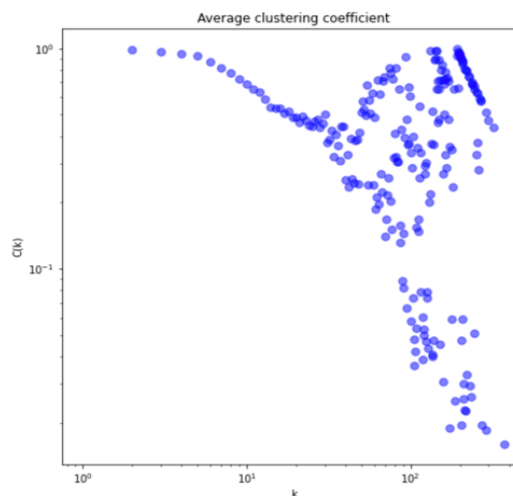
ה Average clustering הממוצע ברשת הינו 0.79 שזהו ערך גבוה יחסית (לשם ההשוואה, ערך c של 0.5 אומר שחצי מהקשרים האפשריים בין החוקרים ששיתפו פעולה עם חוקר ספציפי הינם מחוברים). ראשית, ניתן להסביר ערך זה מעצם האופן בו נבנתה הרשת. כאשר חוקר A , משתף פעולה במאמר מסוים עם חוקרים B ו- C , אוטומטית גם חוקר B משתף פעולה עם C וזאת ללא תלות בכמות הכותבים עבור מאמר ספציפי.

8. **Hierarchy ($C(k)$ vs k):**

בנוסף, ניתן לראות באיור 11 כי ככל שהדרגה הולכת ועולה כך השכנים נוטים להתחבר פחות אחד לשני. תכונה צפויה זו, מאפיינת רשתות scale free המתארות קשרים בין בני אדם ובפרט ברשת שלנו. כאשר הדרגה נמוכה, כמות המאמרים שאותו חוקר כתב (וכפועל יוצא כמות שיתופי הפעולה) בממוצע אינה גבוהה. מאופן בניית הרשת כפי שתואר לעיל, יוצא כי כלל החוקרים ששיתפו פעולה חייבים להיות מחוברים אחד לשני, ולכן נקבל c גבוה. לעומת זאת כאשר כמות שיתופי הפעולה הולכת ועולה, כמות המאמרים השונים שאותו חוקר כתב (וכפועל יוצא עליה בשיתוף הפעולה עם חוקרים **שונים**) הולכת ועולה ולכן c הולך ויורד. את הקשר בין כמות המאמרים לכמות שיתופי הפעולה ניתן לראות באיור 12. באיור זה ניתן לראות באופן ברור כי ככל שחוקר כותב יותר מאמרים, כך כמות שיתופי הפעולה **השונים** בממוצע הולכת ועולה כפי ששיערנו.



איור 12 – מספר שיתופי פעולה ממוצע כתלות במספר מאמרים



איור 11 – מקדם $c(k)$ כתלות בדרגה

9. Degree correlation (K_{nn} vs k):

ניתן לראות באיור 13 כי הרשת הינה אסורטיבית, כלומר חוקרים בעלי דרגה גבוהה, נוטים להתחבר גם כן לחוקרים אחרים בעלי דרגה גבוהה, וכן חוקרים בינוניים לבינוניים וחוקרים בתחילת דרכם עם דומים להם. כלומר האבים מתחברים אחד לשני. נתון זה אינו מפתיע שכן אם יוצאים מנקודת הנחה שחוקרים בעלי שיתופי פעולה רבים הינם מוכרים/נחשבים יותר, ככל הנראה הם יעדיפו להתחבר לחוקרים אחרים הדומים להם מבחינה זו. תכונה זו מאפיינת קשרים בין בני אדם (the rich club) ובפרט גם ברשת שלנו. גם הפרמטר r שנמדד ברשת מצביע על תכונה זו. ה- r שנתקבל גדול מאפס ושווה ל-0.7. מסקנות אלו עולות בקנה אחד עם מה שראינו בסעיף הקודם בניתוח ההיררכיה.

10. ניתוח centrality:

Author name	Author academic institution	Degree centrality
Mohsen Guizani	Qatar university, Doha, Qatar	0.000610
Ruth C Lovering	University College London	0.000523
Helen E. Parkinson	European Bioinformatics Institute	0.000483
Muhammad Imran	University of Glasgow	0.000473
Robert D Finn	European Bioinformatics Institute (EMBL-EBI)	0.000468

טבלה 2- ניתוח degree centrality חוקרים

בשל גודלה של הרשת, לא ניתן לחשב centrality מסוג betweenness, closeness. ניתוח אודות 5 החוקרים בעלי הדרגה הגבוהה ביותר יפורט בהמשך בסעיף הבא המודד את טיב החוקר אל מול כמות שיתופי הפעולה

11. Community detection – ניתוח זה בוצע ברשת האוניברסיטאות בלבד.

3.2. ניתוח Quantity = Quality ?

על מנת לחקור בצורה טובה את הרשת אותה בנינו עבור שיתופי פעולה בין חוקרים, שאלה מעניינת עליה ניסינו לענות היא האם חוקרים עם שיתופי פעולה רבים נחשבים גם למובילים בתחומם. כלומר האם כמות שיתופי הפעולה של החוקר מעידה על טיבו ותרומתו לתחום אותו הוא חוקר. לשם כך ביצענו אינטגרציה עם מידע מ Google scholar, המכיל מידע רב על חוקרים רבים וטיבם, מאמרים, ואף כמות ציטוטים עבור כל חוקר וכל מאמר. מכיוון ש Google לא מאפשרים שליפה מרוכזת של נתונים מ- Google scholar, ביצענו משיכה ידנית עבור חמשת החוקרים בעלי הדרגה הגבוהה ביותר וכן על חוקרים נוספים מובילים בתחומם לשם השוואת הנתונים.

בחרנו ב- 3 מדדים מרכזיים לבחינת טיב החוקר:

1. מוסד אקדמי – כלומר לאיזה מוסד אקדמי משתייך החוקר (ההנחה היא כי חוקר מוביל לרוב שייך לאוניברסיטה מובילה)
 2. כמות הציטוטים שיש לחוקר בשנת 2019
 3. מדד H-index מדד זה עוזר להבין עד כמה תרומתו של חוקר בתחום מסוים משמעותית. לדוגמא, חוקר עם H-index 100, מציין כי לחוקר לפחות 100 מאמרים שכל אחד מהם צוטט לפחות 100 פעמים.
- באמצעות מידע זה והמידע הקיים ברשת שלנו, לקחנו את 5 החוקרים בעלי הדרגה (שיתופי הפעולה) הגבוהה ביותר, וחיברנו יחד. את התוצאות ניתן לראות בטבלה מספר 3

Author name	Institution	Field of research	Degree	#citations in 2019	H-index	#papers Google	#papers network	#papers ratio
Mohsen Guizani	Qatar university, Doha, Qatar	IOT, deep learning, communication sys.	374	6,243	83	189	158	0.84
Ruth C Lovering	University College London	Human genome	321	1,287	42	16	5	0.31
Helen E. Parkinson	European Bioinformatics Institute	Bioinformatics, biology	296	3,226	57	16	7	0.43
Muhammad Imran	University of Glasgow	Communication sys.	290	3,107	59	133	94	0.71
Robert D Finn	European Bioinformatics Institute	Bioinformatics	287	6,087	59	20	7	0.35

טבלה 3 – נתוני חמשת החוקרים המובילים בשיתופי פעולה ברשת

אל מול החוקרים בעלי שיתופי הפעולה הרבים ביותר ברשת שלנו, בחרנו לשם ההשוואה, חוקרים ידועים בעלי שם עולמי וזוכי פרסים יוקרתיים, כפי שניתן לראות בנתונים המופיעים בטבלה מספר 4.

Author name	Institution	Bio	Degree	#citations in 2019	H-index	#papers Google	#papers Net.	#papers ratio
Chris Sander	Harvard Medical School	DeLano Award winner for Computational Biosciences in 2018	15	23,100	190	16	2	0.125
Andrew Ng	Stanford University	Founder of Google brain, Coursera	5	22,749	133	20	1	0.05
Yoshua Bengio	University of Montreal	2018 Turing Award winner, "father" of deep learning	6	87,258	184	103	2	0.02

טבלה 4 – נתוני חוקרים בעלי שם עולמי ומובילים בתחומם

בנוסף עבור כל אחד מהחוקרים המופיעים בניתוח זה, בדקנו את כמות המאמרים בשנת 2019 שיש לכל חוקר בדאטה הקיים ברשותנו והשוונו לכמות המאמרים המופיעה ב Google scholar. עמודת #papers ratio בטבלאות 3 ו-4 מציינת את היחס בין כמות המאמרים בדאטה שלנו וכמות המאמרים בגוגל.

מסקנות מרכזיות מהניתוח:

1. ניתן לראות כי אף אחד מהחוקרים המובילים בשיתופי פעולה אצלנו ברשת אינו שייך לאוניברסיטה מ-10 המובילות בעולם (עפ"י מדד Shanghai), וזאת בניגוד לחוקרים שהבאנו כהשוואה, כפי שניתן לראות בטבלה מס' 5

Institution	Shanghai rank
Qatar university, Doha, Qatar	601-700
University College London	16
European Bioinformatics Institute	Not ranked
University of Glasgow	151-200

טבלה 5 – ציון אוניברסיטאות החוקרים המובילים ברשת עפ"י מדד שנגחאי

2. כמות הציטוטים בשנת 2019 של החוקרים המובילים ברשת שלנו מרשימה יחסית אך אין היא מתקרבת לכמות הציטוטים של החוקרים המובילים שהבאנו כהשוואה
3. גם מדד ה- H-index של החוקרים ברשת שלנו אינו תואם את הציפיות. נתון זה הינו גבוה יחסית, אך אינו גבוה מספיק, בטח בהשוואה לחוקרים המצטיינים בתחומם שהבאנו כהשוואה

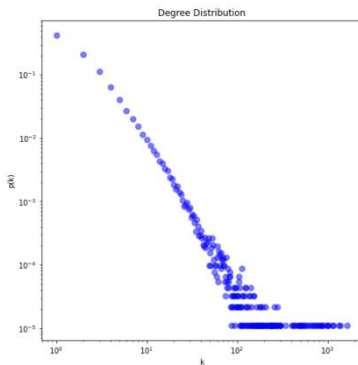
השערות להסבר התופעה:

1. המאמרים בדאטה אצלו לא משקפים באופן מיטבי עבודות של חוקרים מובילים. ניתן לראות השוואה בין כמות המאמרים בדאטה שלנו לכמות בגוגל סקולר והיחס ביניהם בטבלאות 3 ו-4. יחס ייצוג לא אחיד בדאטה יכול לגרום לכך שקיימים חוקרים מובילים שכמות שיתופי הפעולה שלהם אינם באים לידי ביטוי בדאטה שלנו
2. קיים הבדל בין חוקרים בעלי קשרים, המון שיתופי פעולה והרצון לפרסם כמה שיותר מאמרים אל מול חוקרים שמבצעים מחקרי עומק. 2 מהחוקרים ברשת שלנו (Mohsen Muhammad Imran, Guizani) היו עורכים ב-IEEE, ז'ורנל מאוד נחשב בו מפרסמים מחקרים. ניתן לראות שחוקרים אלו נהנים מייצוג גבוה בדאטה שלנו ואף רשומים במעל 100 מחקרים בשנת 2019. ניתן אולי להסיק מכך שחוקרים רבים ירצו לשתף פעולה עימם ולו רק ששמם יופיע על המאמר על מנת לתת תנופה לפרסום המאמר.
3. קיימים תחומים שבהם יש צורך בהמון שיתופי פעולה. כלומר מאמרים בנושאים אלו, לרוב מכילים שיתופי פעולה של המון חוקרים כגון ביולוגיה, ביואינפורמטיקה. בדוגמאות שאנחנו בחנו ומופיעות בטבלה 3 ניתן לראות כי עבור כמות מאמרים מעטה יחסית (פחות מ-10) נוצרים מאות שיתופי פעולה. לבחינת השערה זו דרושה עבודה נוספת שכן בדאטה שלנו לא קיים מפורט נושא המחקר ותחום החוקר.
4. השערה נוספת לתופעה זו יכולה להיות שעדיין לא עבר מספיק זמן, כדי ששיתופי פעולה של החוקרים המובילים ברשת שלנו (והתוצר המתקבל מהם במחקר) יקבלו הכרה לתרומתם האקדמית ומתוך כך עליה בכמות הציטוטים ומדד ה- H-index.

4. רשת שיתופי פעולה בין אוניברסיטאות:

4.1. חקר מדדי הרשת:

גם את הרשת שיתופי הפעולה בין בחרנו לתאר כרשת לא-כיוונית כאשר צומת הינו אוניברסיטה וקשת הינה שיתוף הפעולה בין אוניברסיטאות, כלומר שני חוקרים שפרסמו מאמר יחדיו והם מאוניברסיטאות שונות. חשוב לציין שהתמקדנו בשיתוף פעולה בין אוניברסיטאות והתעלמנו משיתופי פעולה תוך אוניברסיטאיים, כלומר השתדלנו לנקות את הנתונים משיתופי פעולה בין פקולטות או חוגים מאוניברסיטאות זהות. עשינו זאת על ידי ניקוי מילים וסימונים חוזרים מהשמות של האוניברסיטאות והשוואה בין השמות השונים. ניקינו שמות שהיו דומים מעל סף מסוים שמצאנו לנכון עבור משימה זו. עם זאת, חשוב לציין שישנן זליגות שמהן לא יכולנו להימנע. למשל, ישנם מקרים בהם מופיעה אוניברסיטה מסוימת וגם בית הספר לרפואה מאותה האוניברסיטה. במקרים אלה אנו מתייחסים לפקולטה זו כמוסד מחקרי בפני עצמה.



איור 14 - התפלגות דרגות ב $\log\text{-}\log$ scale

1. כמות צמתים (N): 98,229

2. כמות קשתות (L): 202,447

3. דרגה ממוצעת: 4.16

4. התפלגות דרגות ברשת $(P(k))$:

ניתן לראות באיור 14 כי הרשת שלנו מתנהגת כרשת scale free.

הרשת שלנו מתארת קשרים בין אוניברסיטאות ולכן התפלגות זו

עולה בקנה אחד עם הציפיות שלנו. מטבע הדברים, מספר

מצומצם של אוניברסיטאות הן מובילות, משפיעות ומשתפות

פעולה עם מוסדות אחרים ולצד זה, שאר האוניברסיטאות הן פחות משפיעות ומשתפות

פעולה עם פחות מוסדות מחקריים. זהו הסבר שיכול להסביר את התפלגות הדרגות.

5. התפלגות משקל w וחוזק s ברשת $(P(w), P(s))$:

באופן זהה להתפלגות הדרגות, גם התפלגות המשקולות והחוזק ברשת מראה התנהגות

scale-free כפי שניתן לראות באיורים 15 ו-16. מכיוון שרשת זו נגזרת למעשה מרשת

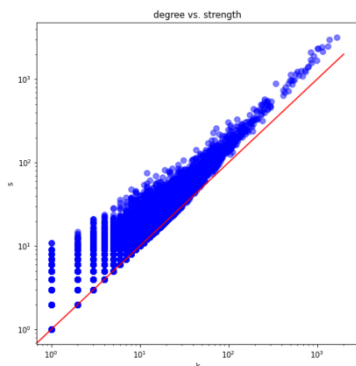
החוקרים, אנו ציפינו לראות תוצאות אלה. באיור 17 שמציג גרף חוזק אל מול דרגה, ניתן

לראות תופעה זהה לזו שראינו בגרף הזהה של החוקרים, ככל שעולים בדרגה, העלייה

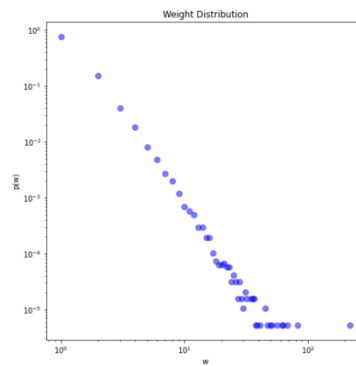
בחוזק יורדת. הקו האדום מסמל מה היה קורה אם לכל אוניברסיטה בדרגה נתונה היה

שיתוף פעולה אחד עם כל מוסד וכל הנקודות מעל קו זה מראות שיתופי פעולה חוזרים. מכך

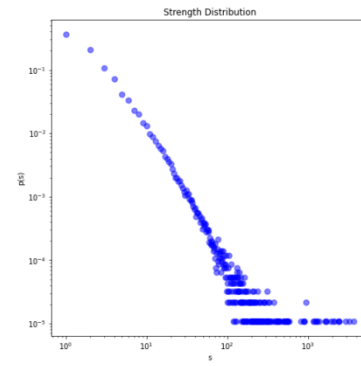
אנחנו מסיקים שככל שעולים בדרגה, עולה הגיוון בשיתופי הפעולה ויחד עם זאת יורדת כמות שיתופי הפעולה החוזרים.



איור 17 - התפלגות דרגות vs חוזק ב- \log scale



איור 16 - התפלגות משקלים ב- \log - \log scale



איור 15 - התפלגות חוזק ב- \log - \log scale

6. $\langle l \rangle$: Average path length

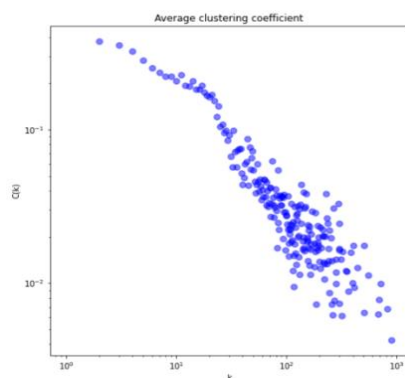
מכיוון שהרשת שלנו אינה קשירה לא ניתן למדוד פרמטר זה באופן אנליטי, לכן נאמר כי האורך הינו אינסופי.

7. $C(k)$: Average clustering coefficient

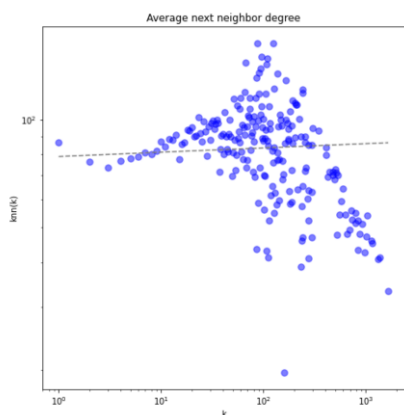
ה Average clustering הממוצע ברשת הינו 0.19. מספר זה הוא יחסית נמוך ואומר שבממוצע אין הרבה 'קליקות'. עצם שיתוף פעולה של מוסד מחקרי עם מוסדות אחרים לא יוביל אותנו בהכרח לחשוב שהמוסדות שמשתפים עמו פעולה משתפים גם כן פעולה אחד עם השני.

8. $C(k)$ vs k : Hierarchy

ניתן לראות באיור 18 כי ככל שעולים בדרגה, מקדם ההתקבצות קטן. דבר זה הוא הגיוני ועולה בקנה אחד עם הציפיות. מוסדות המשתפים פעולה עם מספר מאוד מצומצם של מוסדות אחרים נוטים לשתף פעולה עם מוסדות תוך אוניברסיטאיים או מוסדות מאותה המדינה. הגיוני שמוסדות אלה יהווה 'קליקה'. עליה בדרגה משמעותה שיתוף פעולה עם יותר מוסדות והגיוני לצפות לירידה במקדם ההתקבצות מכיוון שמגוון שיתופי פעולה עם מוסדות שונים ומדינות שונות כנראה מושפע מצרכים אקדמיים ולא מקליקות טבעיות.



איור 18 - מקדם התקבצות ממוצע כתלות בדרגה



איור 19 – דרגת שכן ממוצעת כתלות בדרגה

9. Degree correlation (Knn vs k):

ניתן לראות באיור 19 כי ישנה מגמת עלייה חלשה במדד ה k_{nn} ככל שעולים בדרגה. עם זאת, אנו מקבלים מדד אסורטיביות r קרוב ל-0 ($r=-0.03$). מזאת ניתן להסיק שמבחינת אסורטיביות, הרשת היא נייטרלית. כלומר, מוסדות מחקר לא דווקא נוטים לשתף פעולה עם בהכרח עם מוסד שהם שונים או דומים מהם.

10. Centrality:

גם רשת האוניברסיטאות גדולה בכדי לחשב betweenness ו closeness ולכן חישבנו degree centrality בלבד.

degree centrality:

Name of academic institution	Degree centrality	Degree
University of Cambridge	0.018	1654
Tsinghua University	0.015	1377
Zhejiang Univ	0.0144	1325
IBM	0.0125	1153
Carnegie - Mellon	0.0124	1143
Microsoft	0.0112	1033
MIT	0.11	1012

טבלה 6 – ניתוח Degree centrality מוסדות מחקר

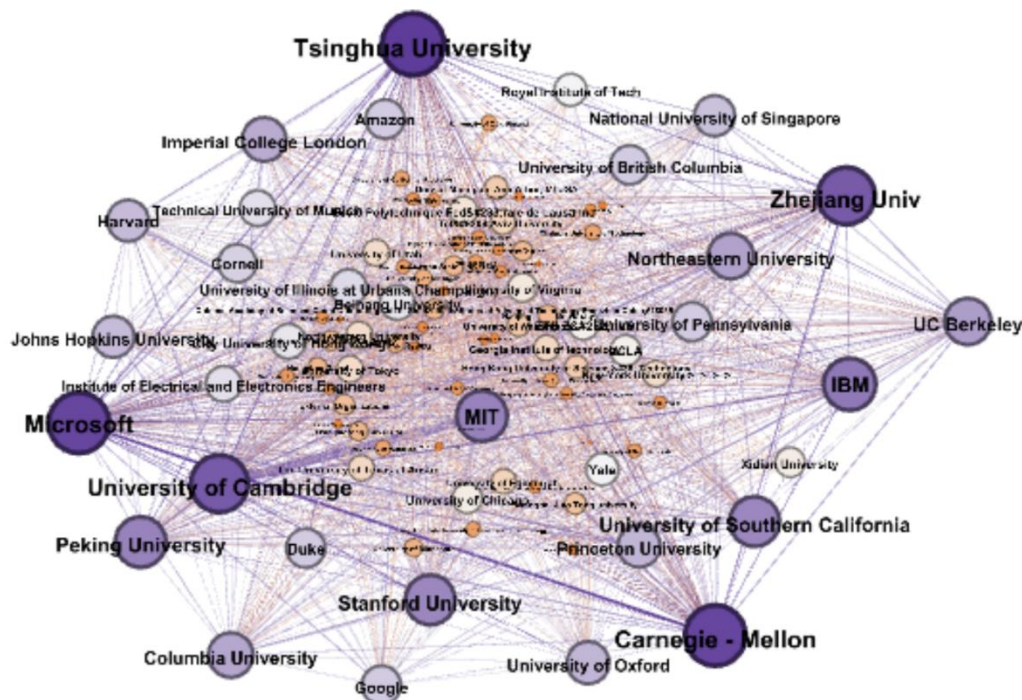
בטבלה 6 ניתן לשים לב שהמוסדות בעלות הדרגה הגבוהה ביותר ברשת הינן אוניברסיטאות מוכרות ומובילות ומכוני מחקר של שתי חברות טכנולוגיה מהמובילות בעולם. נציין שעבור רוב המוסדות המובילים איחדנו את השמות הדומים לכדי שם אחיד של מוסד כך שעבור ואריאציה של IBM, Microsoft או כל מוסד מוביל אחר, איחדנו את המוסד תחת שם אחד. כלומר, גם אם חלק מהמחקרים נעשו למשל ב- IBM ישראל, ארה"ב או סין, אלה נכנסו תחת שם מוסד אחד- IBM. מכוני מחקר של חברות אלה ושל חברות דומות (גוגל, אמאזון וכד') נמצאות ברשימות המוסדות המובילים בעיקר בגלל פעילות המחקר הענפה שלהן בשיתוף פעולה עם

אוניברסיטאות ומכוני מחקר רבים אך גם בזכות זאת שהם פרוסים בכל העולם ומחזיקים מכוני מחקר נפרדים בהרבה מדינות.

11. Community detection:

11.1. צמצום הדאטה לתת רשת של 100 המוסדות המובילים בשיתופי פעולה:

בחלק זה נתקלנו בבעיה בגלל גודל הרשת. כאמור, רשת האוניברסיטאות מכילה יותר מ 90 אלף צמתים ולכן חלוקת המוסדות לקלאסטרים שונים היא משימה קשה והתוצאה אינה אינפורמטיבית כלל. בשביל לקבל מושג טוב לגבי חלוקה לקלאסטרים וחישוב מודולאריטי צמצמנו את הנתונים ל- 100 המוסדות בעלי הדרגות הגבוהות ביותר ובדקנו את שיתופי הפעולה ביניהם. כלומר, ציצמנו את רשת מוסדות המחקר ממעל 90 אלף צמתים ל 100 צמתים של מוסדות מחקר מובילים ובדקנו את שיתופי הפעולה בין מוסדות אלה. באיור 20 ניתן לראות את הרשת המצומצמת של 100 האוניברסיטאות המובילות. ברשת זו מוצגים שיתופי הפעולה בין המוסדות על ידי הקשתות והקודקודים מופיעים בגדלים וצבעים שונים. ככל שקודקוד הוא יותר גדול ככה שיתופי הפעולה של המוסד הם יותר מרובים (דרגה גבוהה יותר). הוספנו גם סקאלת צבע- ככל שהדרגה גבוהה יותר, הקודקוד נבצע בצבע סגול כהה יותר. וככל שהדרגה נמוכה יותר, הסקאלה כתומה יותר.



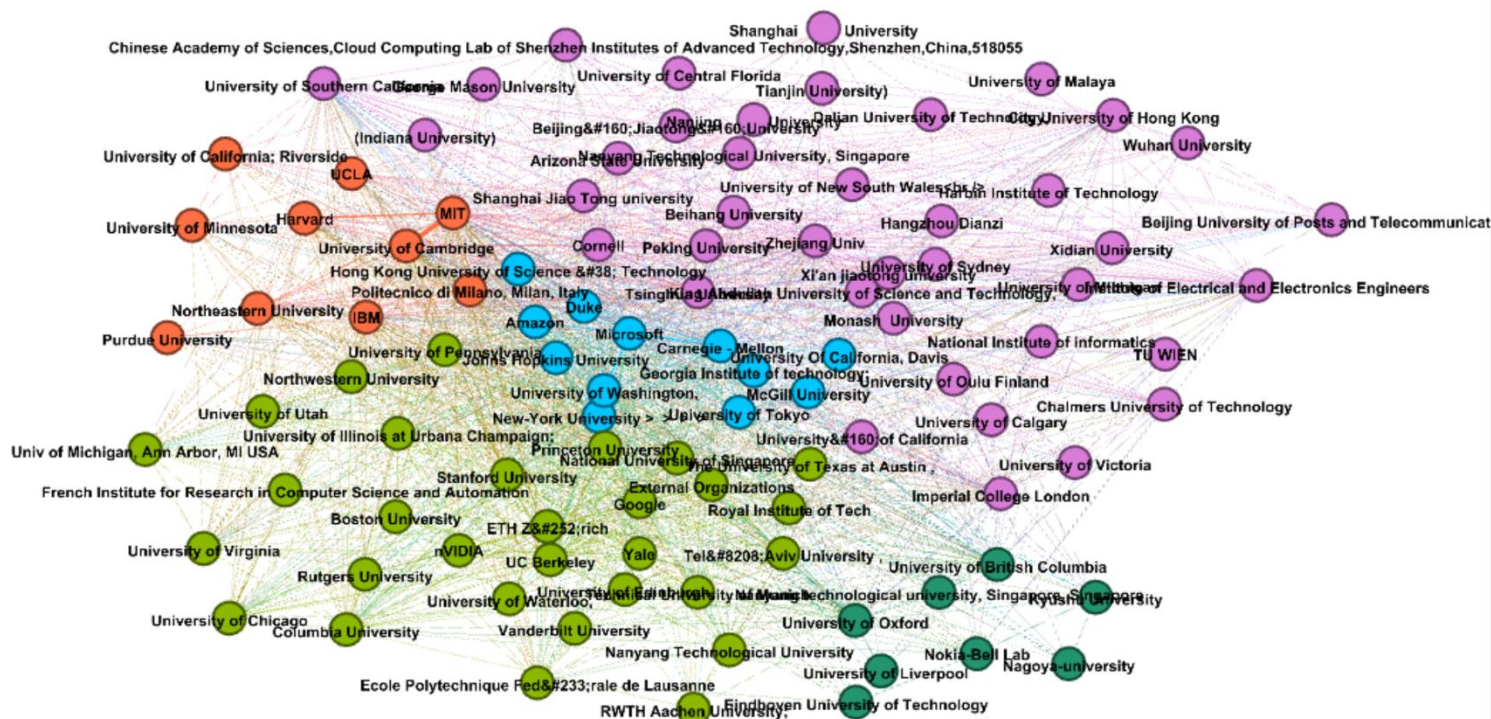
איור 20 – רשת 100 מוסדות המחקר המובילים בשיתופי פעולה

11.2. שיטות מודולאריות:

השתמשנו בשתי שיטות לחלוקת המוסדות לקלאסטרים ולאחר מכן חישבנו את המודולאריות עבור כל שיטה.

השיטה הראשונה - greedy modularity - קיבלנו חלוקה ל-40 קבוצות ומודולאריות של 0.005. השיטה השנייה - Louvain modularity - קיבלנו חלוקה לחמש קבוצות ומודולאריות של 0.18.

ניתן לראות כי במקרה שלנו Louvain היא שיטה יותר יעילה שכן היא נותנת ציון מודולאריות גבוה יותר וגם חלוקה יותר הגיונית לקבוצות. באיור 21 ניתן לראות את החלוקה לקבוצות השונות. ניתן לראות שרוב המוסדות המובילים מהמזרח מסין (בעיקר), סינגפור והונג קונג משויכים לאותה הקבוצה (קבוצה ורודה). הגיוני להסיק מכך שהרבה חוקרים אלה מרגישים יותר בנוח לשתף פעולה עם חוקרים ממדינתם שדוברים את שפתם וגם ישנה קירבה תרבותית וגיאוגרפית. קבוצה אחרת (בירוק כהה) מאופיינת בעיקר במוסדות לא אמריקאים ובעיקר במוסדות בריטיים וקנדיים מובילים. הקבוצה הכתומה מורכבת ממוסדות אמריקאים מובילים (Tier 1 universities) ומכילה את המוסדות המובילים – Cambridge, MIT, Harvard, IBM. קיים קשר אדוק בין מוסדות אקדמיים ומכוני מחקר של חברות טכנולוגיות. הקבוצה הכחולה דומה באופיה לקבוצה הכתומה ומכילה הן אוניברסיטאות מובילות מארה"ב ומהמזרח והן מכוני מחקר של חברות טכנולוגיות שמשתפים פעולה עם אוניברסיטאות אלה.



איור 21- חלוקת רשת 100 המוסדות המובילים בשיתופי פעולה על פי אלגוריתם לובין

4.2. האם Quantity = Quality?

נרצה לבדוק האם כמות שיתופי הפעולה עם אוניברסיטאות שונות מעידים על איכות האוניברסיטה ועל מידת ההשפעה שלה. נשתמש בשני מדדים. המדד הראשון הוא **דירוג שנגחאי** - זהו מדד מפורסם המדרג אוניברסיטאות בעולם על פי מדדים שונים ונחשב לאחד המדדים האיכותיים והאובייקטיביים. המדד השני הוא מדד **Influence Ranking** - מדד מפורסם גם כן שמשתמש בדאטה רב ומודד את מידת ההשפעה של האוניברסיטאות על האקדמיה, מספר פרסומים בז'ורנלים שונים ובוחן עוד מספר מדדים שונים. בסעיף זה נסתכל רק על אוניברסיטאות ונתעלם ממכוני מחקר של חברות טכנולוגיות כגון גוגל ומייקרוסופט.

כפי שניתן לראות בטבלה 7, ובשונה מתמונת המצב אותה קיבלנו מניתוח החוקרים, אנו רואים תמונה די מובהקת. מוסדות המובילים בשיתופי פעולה אכן נמצאים גבוה בשני הדירוגים שבחנו. מכך אנו מסיקים ששיתופי פעולה הם אכן מדד טוב לבחינת טיב המוסד האקדמי ולבחינת השפעתו על האקוסיסטם האקדמי-מחקרי. אולם, ניתן להבחין גם בשני מקרים בהם הדירוגים אינם עולים בקנה אחד עם ציפיותינו. אוניברסיטת קרנגי-מלון מובילה בשיתופי פעולה (מקום 4) אך מדורגת במקום 95 בדירוג שנגחאי ובמקום 26 בדירוג ההשפעה. הסבר הגיוני לכך הוא שאוניברסיטה זו היא קטנה משמעותית משאר האוניברסיטאות המובילות והיא ממוקדת בתחומי מחקר מסויימים בה היא מצטיינת (תחומי המדעים המדויקים וההנדסה). לעומתה, האוניברסיטאות המובילות האחרות בטבלה מצטיינות במנעד רחב של תחומים ובעוד אספקטים רבים נוספים ולכן הן מדורגות משמעותית יותר גבוה בשני הדירוגים. אוניברסיטה נוספת בה רואים תמונת מצב שונה היא אוניברסיטת טסינגואה הסינית שמדורגת במקום השני במדד שיתופי הפעולה אך נמצאת רק במקום ה-43 בדירוג שנגחאי ולא נמצאת כלל בין 50 האוניברסיטאות המובילות בדירוג ההשפעה. השערתנו היא שישנה הטיה או דגש על אוניברסיטאות מערביות/ אמריקאיות בדירוגים השונים.

University	Rank by our network degree	Shanghai Rank	Influence Rank
University of Cambridge	1	3	10
Tsinghua University	2	43	Not in top 50
Stanford	9	2	3
Carnegie - Mellon	4	95	26
MIT	5	4	5
University of Oxford	6	7	13
Harvard	12	1	1

טבלה 7 – דירוג אוניברסיטאות על פי דירוג דרגה, מדד שנגחאי ומדד השפעה

5. אוניברסיטת תל אביב case study:

לאחר ניתוח רשתות החוקרים והאוניברסיטאות, החלטנו לקחת את אוניברסיטת תל אביב בתור case study ולבחון כיצד הרשת הפנימית של האוניברסיטה באה לידי ביטוי בכל אחת מן הרשתות. בניתוח נתנו דגש לייצוג המחלקה להנדסת תעשייה בתוך רשת החוקרים, החוקרים המובילים באוניברסיטה וכן את קשריה של האוניברסיטה עם מכוני מחקר ואוניברסיטאות בארץ ובעולם.

5.1 ייצוג האוניברסיטה ברשת החוקרים – עיקרי הממצאים

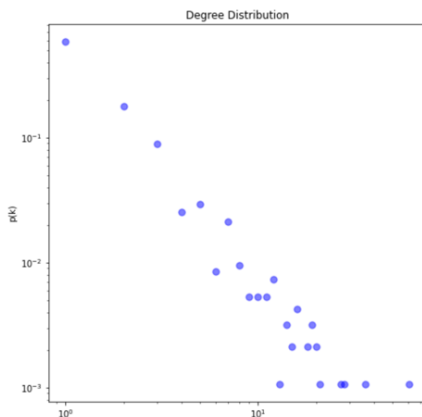
1. כמות צמתים (N): 963

2. כמות קשתות (L): 1234

3. דרגה ממוצעת: 2.56

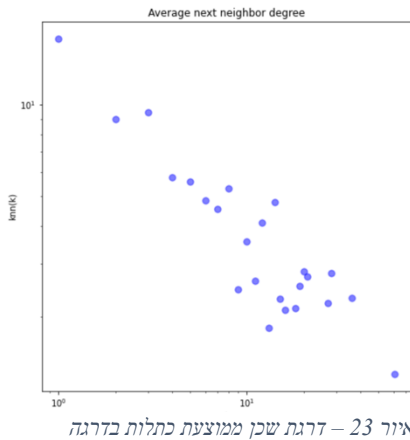
4. התפלגות דרגות ברשת (P(k):

ניתן לראות באיור 22 כי גם ברשת החוקרים של האוניברסיטה, התפלגות הדרגות ברשת שלנו מתנהגת כרשת scale free. עם התפלגות power law וזאת למרות כמות צמתים וקשתות נמוכה משמעותית מהרשת המקורית, ההתנהגות של הרשת נשמרת גם ברשת הפנימית של האוניברסיטה.



5. Degree correlation (Knn vs. k)

באיור 23 ניתן לראות כי באופן מפתיע ובניגוד להתנהגות הרשת המלאה של החוקרים, רשת החוקרים של האוניברסיטה הינה דיסאסורטיבית ובעלת z שלילי של -0.287 . המשמעות של ממצא זה היא כי חוקרים בעלי דרגה גבוהה, נוטים לשתף פעולה בביצוע מחקרים דווקא עם חוקרים בעלי דרגה נמוכה. מניתוח מספר מקרים של הרשת אנו משערים כי ההסבר לתופעה הוא כי מרבית המחקרים שבוצעו ומופיעים בדאטה הם של מאסטרנטים ודוקטורנטים המשתפים פעולה או מקבלים הנחיה מחוקרים מובילים. את השערה זו בדקנו ע"י ניתוח מדגם של קשרים ברשת של חוקרים בעלי דרגה נמוכה עם האבים. באמצעות שימוש בפלטפורמת-LinkedIn ראינו כי מרבית החוקרים בעלי הדרגה הנמוכה ששיתפו פעולה עם חוקרים מובילים היו סטודנטים למאסטר או דוקטורנטים בשנת 2019.



6. ניתוח centrality

name	department	Degree centrality
Daniel Cohen-Or	school of computer science	0.0646
Talma hendler	school of psychological sciences	0.0381
Haim kaplan	school of computer science	0.0296
Lior wolf	school of computer science	0.0286
Mooly Sagiv	school of computer science	0.0222

טבלה 8 Degree centrality Top 5 Authors

name	department	Closeness centrality
Daniel Cohen-Or	school of computer science	0.0582
Yishay Mansour	school of computer science	0.0504
Amit Daniely	google research	0.0478
Shay Moran	princeton university	0.0468
Haim kaplan	school of computer science	0.0466

טבלה 9 Closeness centrality Top 5 Authors

name	department	Betweenness centrality
Yishay Mansour	school of computer science	0.0210
Amir Globerson	school of computer science	0.0185
Amit Daniely	google research	0.0168
Noga Alon	school of computer science	0.0141
Haim Kaplan	school of computer science	0.0127

טבלה 10 Betweenness centrality Top 5 Authors

בניתוח ה- centrality של רשת החוקרים של אוניברסיטת תל אביב, ניתן למצוא כמה מס' עובדות מעניינות:

1. פרופסור דניאל כהן אור- החוקר בעל שיתופי הפעולה הרבים ביותר באוניברסיטה בשנת 2019. ניתן לראות כי הוא בעל הדרגה הגבוהה ביותר וכן בעל ה- closeness centrality הגבוה ביותר. בשל העובדה כי המחלקה למדעי המחשב הינה בעלת הייצוג הגבוה והמשמעותי ביותר ברשת, נתון זה אינו מפתיע. בנוסף, ניתן לראות כי פרופסור דניאל כהן-אור הרבה לבצע שיתופי פעולה בשנת 2019 דווקא עם חוקרים בעלי דרגה נמוכה יחסית, ולא עם חוקרים בעלי דרגה דומה. מידע נוסף אודות פרופסור דניאל כהן-אור יפורט בהמשך.
2. פרופסור חיים קפלן – ניתן לראות כי פרופסור חיים קפלן מופיע בחמשת החוקרים המובילים בכלל מדדי ה- centrality שמדדנו. העובדה שגם הוא הינו חוקר מוביל במחלקה למדעי המחשב ושמחלקה זו בעלת הייצוג הגבוה ביותר ברשת מהווה הסבר לנתון זה. בנוסף, כפי שניתן לראות באיור 24, פרופסור חיים קפלן מבצע שיתופי פעולה עם חוקרים שגם הם מרכזיים ובעלי שיתופי פעולה רבים במחלקה למדעי המחשב ולכן מדדי ה- centrality הגבוהים אינם מפתיעים.
3. פרופסור ישי מנסור – גם הוא בדומה לפרופסור חיים קפלן שייך למחלקה למדעי המחשב ומבצע שיתופי פעולה עם חוקרים משפיעים כמותו. גם הוא בעל דרגה גבוהה יחסית ברשת (התשיעי הכי גבוה ברשת)

5.2 ייצוג רשתות עבור אוניברסיטת תל אביב באמצעות Gephi

לשם ניתוח מיטבי של ה- case study אותו בחרנו לנתח, ייצרנו 3 רשתות שונות עבור האוניברסיטה:

1. רשת חוקרים – פירוט חמשת החוקרים המובילים והצגת הרשת בעומק 5 (יפורט בהמשך)
2. רשת חוקרים – ייצוג החוקרים במחלקה להנדסת תעשייה
3. רשת אוניברסיטאות – ייצוג קשריה של האוניברסיטה עם מוסדות מחקר בארץ ובעולם.

על מנת לבצע ניתוח מיטבי, היה עלינו לבצע משימות רבות של עיבוד וניקוי טקסט בדאטה. מכיוון שמוסד המחקר אליו משתייך כל חוקר מגיע כטקסט (ללא id) השתמשנו בטכניקות שונות על מנת לייצר אחידות עד כמה שניתן בדאטה, שכן אותו מוסד מחקר, יכול להופיע בשמות שונים או ברמת פירוט שונה (אוניברסיטה, מחלקה ואוניברסיטה וכו'). לשם כך השתמשנו בשיטות כגון TF-IDF, cosine similarity, n-grams ועוד. כמו כן ביצענו אף עבודת השלמה ידנית באמצעות Google ורשת LinkedIn להשלמת מידע חסר או חלקי.

לשם הצגת הרשת באופן נוח יותר לניתוח, השתמשנו באלגוריתם DFS (Depth First Search), על מנת לנקות רעש מהנתונים. עבור הרשת המתארת את הקשרים של חמשת החוקרים המובילים באוניברסיטה הגדרנו מקסימום עומק של 5. לעומת זאת, עבור הרשת של המחלקה להנדסת תעשייה לא הגדרנו עומק מקסימלי, שכן כמות הצמתים אינה גבוהה באופן יחסי.

1. חמשת החוקרים המובילים באוניברסיטה

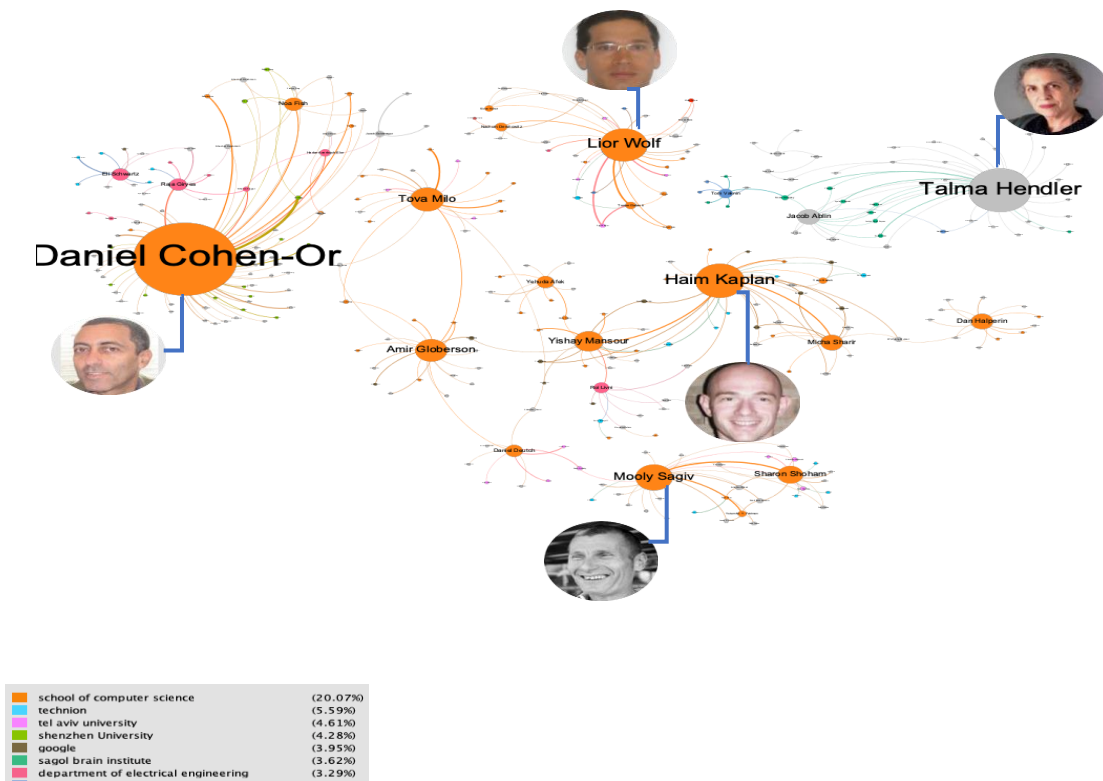
כפי שצוין לעיל, ניתן לראות בטבלה מספר 11 וכן באיור 24 כי המחלקה בעלת הייצוג המשמעותי ביותר ברשת האוניברסיטה הינה המחלקה למדעי המחשב, כאשר 4 מתוך 5 החוקרים בעלי שיתוף הפעולה הרב ביותר מגיעים ממחלקה זו. החוקר המשמעותי ביותר באוניברסיטה וביתרון מובהק הינו דניאל כהן-אור, עם 61 שיתופי פעולה בשנת 2019 עם אוניברסיטאות שונות מהעולם וכן עם מחלקות בתוך האוניברסיטה - המחלקה למדעי המחשב והמחלקה להנדסת חשמל. ניתן לראות כי גם ה-H-index של פרופסור כהן-אור הינו הגבוה ביותר מבין חמשת החוקרים המובילים באוניברסיטה ואף גבוה יותר מחמשת החוקרים בעלי הדרגה הגבוהה בכל רשת החוקרים שניתחנו. חוקרת נוספת אשר בולטת ברשת הינה פרופסור תלמה הנדלר, מהמחלקה למדעי הפסיכולוגיה. למרות ייצוג מועט ביותר ברשת למחלקה זו (2 חוקרים בלבד), פרופסור תלמה הנדלר הינה השנייה ברשת עם שיתופי פעולה הרבים ביותר. מרבית שיתופי הפעולה של פרופסור תלמה היא דווקא עם חוקרים מגרמניה מאוניברסיטת Leipzig, אך קיימים שיתופי פעולה גם עם הפקולטה לרפואה באוניברסיטה, אוניברסיטת אוקספורד ו-Sagol brain institute בתל אביב, עבורו פרופסור תלמה משמשת כדירקטורית (ומכאן שיתופי הפעולה הרבים עם מוסד זה).

Author name	department	Degree	#citations in 2019	H-index
Daniel cohen-or	school of computer science	61	3,115	102
Talma hendler	school of psychological science	36	1,455	66
Haim kaplan	school of computer science	28	565	53
Lior wolf	school of computer science	27	3,501	59
Mooly Sagiv	school of computer science	21	691	61

טבלה 11 – דירוג חוקרים מובילים מאוניברסיטת תל אביב

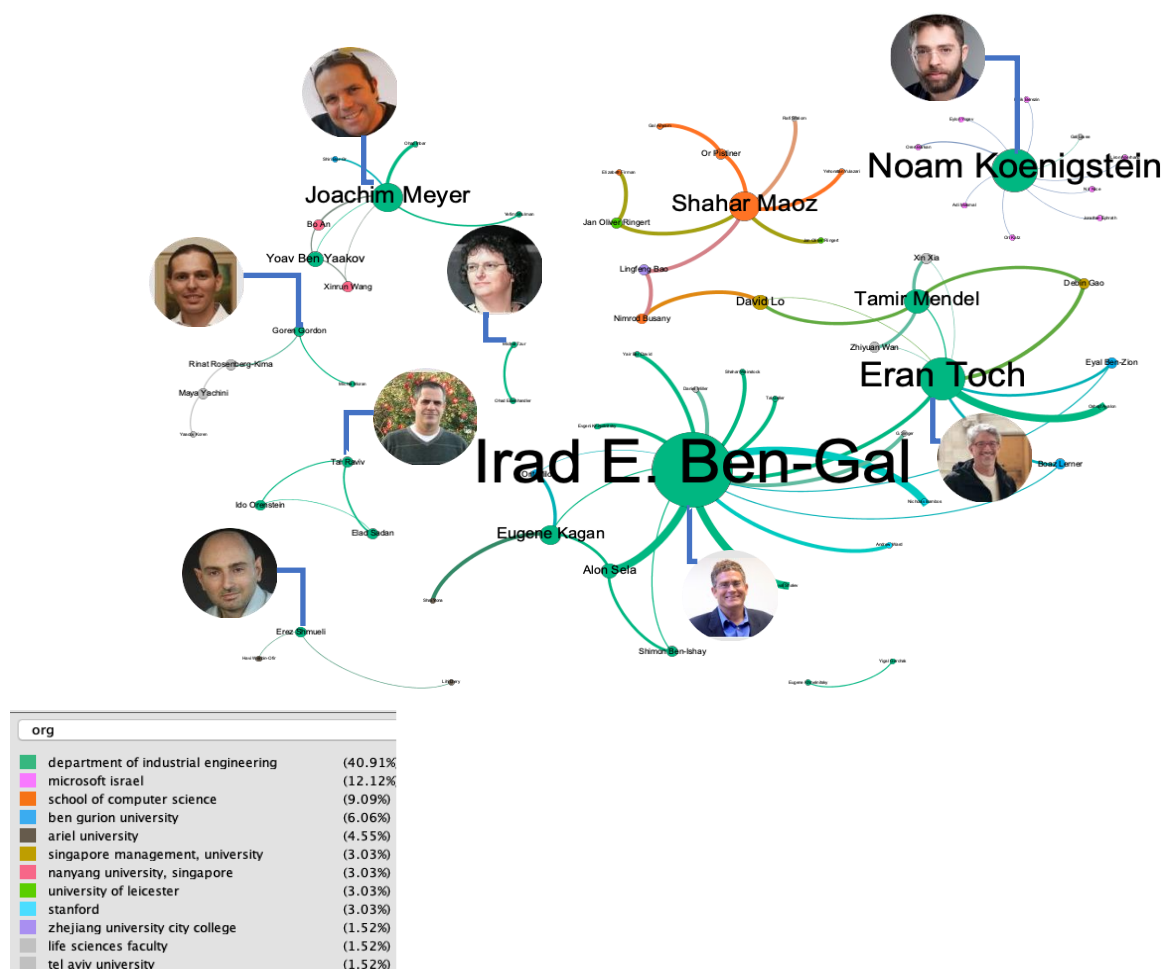
בנוסף, ניתן לראות כי שיתופי הפעולה המרכזיים מתבצעים באופן פנים מחלקתי אך קיימים שיתופי פעולה נוספים גם עם מוסדות נוספים בארץ (טכניון, אוניברסיטת בן גוריון), מוסדות מחקר עסקיים (גוגל, IBM וכו') ואוניברסיטאות בעולם.

שיתוף פעולה פנים אוניברסיטאי בולט ברשת, הינו בין המחלקה למדעי המחשב למחלקה להנדסת חשמל. נתון זה אינו מפתיע שכן קיימים נושאים רבים אשר משיקים לשתי המחלקות ומהווים פוטנציאל לשיתוף פעולה (כגון בינה מלאכותית, AI, ניתוחי תמונה ועוד).



איור 24 – רשת שיתופי הפעולה של חמשת החוקרים המובילים באוניברסיטת תל אביב

2. רשת חוקרים - המחלקה להנדסת תעשייה:



איור 25 – רשת שיתופי פעולה של חוקרים מהפקולטה להנדסת תעשייה וניהול באוניברסיטת תל אביב

מאיור 25 ניתן לראות כי מרבית שיתופי הפעולה של החוקרים במחלקה הינם פנים-מחלקתיים, אם כי ישנם שיתופי פעולה עם אוניברסיטאות שונות בארץ (כגון בן גוריון ואריאל) ואוניברסיטאות בעולם (כגון סטנפורד, אוניברסיטת סינגפור ועוד).

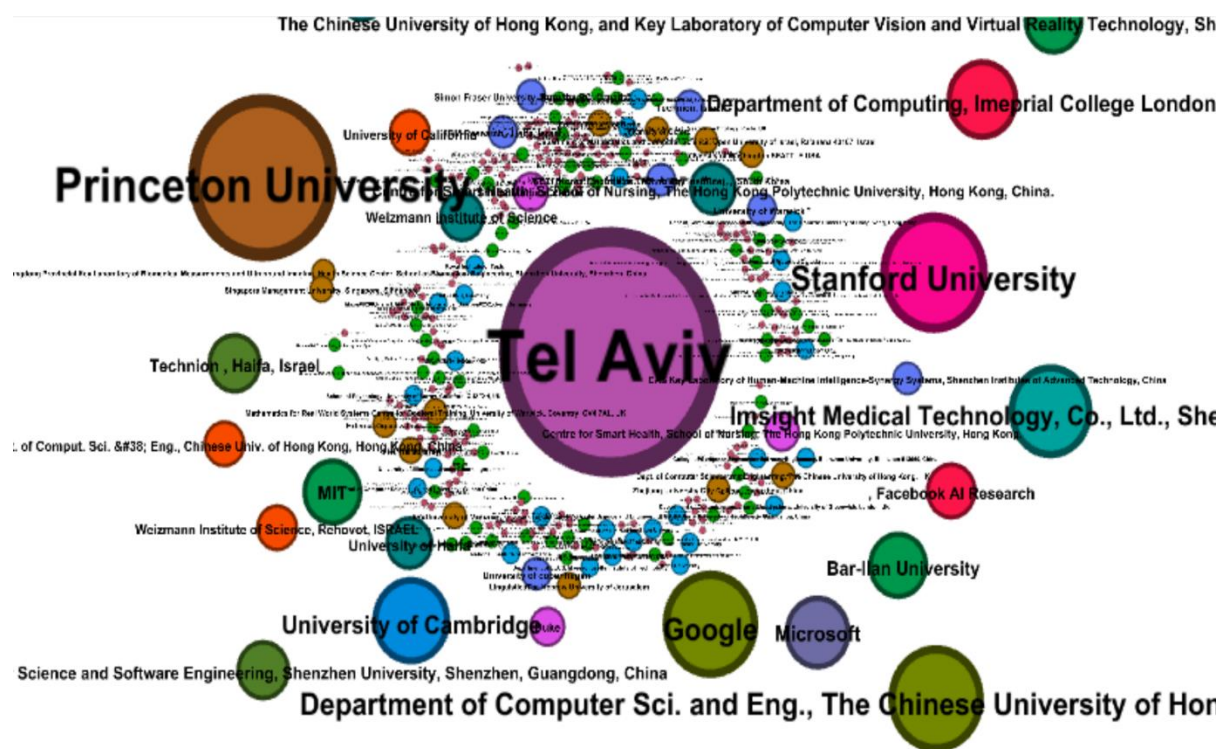
החוקר הבולט ביותר במחלקה הינו פרופסור עירד בן-גל, אשר שימש כראש המחלקה בעבר ובעל שיתופי פעולה הרבים ביותר במחלקה בשנת 2019. נתון זה אינו מפתיע שכן פרופסור בן-גל הינו חוקר ותיק ותפקידו כראש הפקולטה מהווה כר פורה ליצירת שיתופי פעולה רבים בעולם האקדמי והן במגזר העסקי.

נתון בולט נוסף ברשת הינו כמות שיתופי הפעולה הרבים יחסית שיש במחלקה עם מיקרוסופט ישראל. האחראי במחלקה על מרבית שיתופי פעולה אלו הינו ד"ר נעם קונינגשטיין. ממחקר שביצענו ראינו כי

ד"ר קונינגשטיין עבד במיקרוסופט לפני הצטרפותו לאוניברסיטה במהלך שנת 2018. אנו משערים כי שיתופי פעולה אלו הינם מחקרים ושיתופי פעולה שהתחילו עוד בעבודתו במיקרוסופט לפני הצטרפותו לסגל האוניברסיטה והבשילו לכדי מאמרים בשנת 2019.

3. קשרי האוניברסיטה עם אוניברסיטאות אחרות:

נסתכל על הוויזואליזציה של תת הרשת. ניתן לראות באיור 26 שהגדלנו את גודל הצמתים לפי כמות שיתופי הפעולה של המוסד עם אוניברסיטת תל אביב. ניתן לראות מוויזואליזציה זו למשל שחוקרים מאוניברסיטת תל אביב שיתפו פעולה הכי הרבה פעמים עם חוקרים מאוניברסיטת פרינסטון שבארה"ב. כפי שניתן לראות בטבלה מספר 12, מרבית שיתופי הפעולה הם עם אוניברסיטאות ומוסדות מחקר מהמובילים ביותר בעולם ועם מוסדות ישראלים כגון – אוניברסיטת חיפה, בר-אילן, הטכניון, וממכון ויצמן. חוקרים מאוניברסיטת תל אביב שיתפו פעולה עם חוקרים מ 323 אוניברסיטאות שונות בשנת 2019.



איור 26 – רשת המוסדות ששיתפו פעולה עם אוניברסיטת תל אביב בשנת 2019

org1	org2	weight
Tel Aviv	Princeton University	28
Tel Aviv	Stanford University	18
Tel Aviv	Department of Computer Sci. and Eng., The Chin...	16
Tel Aviv	Im sight Medical Technology, Co., Ltd., Shenzhe...	14
Tel Aviv	University of Cambridge	13
Tel Aviv	Google	13
Tel Aviv	, Facebook AI Research	12
Tel Aviv	MIT	12
Tel Aviv	Department of Computing, Imperial College Lond...	12
Tel Aviv	Microsoft	11
Tel Aviv	Bar-Ilan University	10
Tel Aviv	The Chinese University of Hong Kong, and Key L...	10
Tel Aviv	Technion , Haifa, Israel	9
Tel Aviv	IBM	9
Tel Aviv	College of Computer Science and Software Engin...	9
Tel Aviv	University of Haifa	8
Tel Aviv	Guangdong Provincial Key Laboratory of Compute...	8
Tel Aviv	Centre for Smart Health, School of Nursing, Th...	8
Tel Aviv	Weizmann institute of Science	8
Tel Aviv	Dept. of Comput. Sci. & Eng., Chinese Univ. of...	7

טבלה 12 – מספר שיתופי פעולה של מוסדות שונים עם
אוניברסיטת תל אביב בשנת 2019