

נספחים

נספח א' - טבלאות מלאות של חלוקות אלגוריתמיות

להלן הטבלאות המלאות של תוצאות החלוקות האלגוריתמיות, המוזכרות בפרק התוצאות.

חלוקות עם ציון המודולריות הגבוה ביותר

במידה ולא מגבילים את עצמנו לחלוקות קיימות, ניתן לראות כי החלוקות המובילות הן של הרצות ספציפיות של ליידן בעיקר, ברזולוציה 1. כמות הקבוצות בכלל הרצות אלו הייתה 9, וגודל הקבוצה הקטן ביותר היה 11. עם זאת ניתן לראות שסוג חלוקה זו אינו יציב דיו, שכן הושגה תוצאה זהה רק ב-4 הרצות מתוך כלל המאה. כל שינוי גבולות, גם הקטן ביותר, מקשה על מתיחת הגבול בין אזורי הזיקה. ראו טבלה 4:

טבלה 4: הרצות טובות ביותר מבחינת מודולריות

אלגוריתם	רזולוציה	מודולריות	ציון צללית ממוצע לקבוצות (זמן כמרחק)	ציון צללית ממוצע לקבוצות (מרחק כמרחק)	גרעינים	כמות הרצות
ליידן	1	0.62571	0.29025	0.39439	49,83,98	3
ליידן	1	0.62571	0.29020	0.39422	36,43,68,91	4
ליידן	1	0.62570	0.29042	0.39413	28	1
ליידן	1	0.62569	0.29032	0.39422	9	1
ליידן	1	0.62567	0.29134	0.39574	8	1
ליידן	1	0.62567	0.29083	0.39516	21,71	2
ליידן	1	0.62564	0.28994	0.39428	54	1
ליידן	1	0.62559	0.28990	0.39363	87,89	2
ליידן	1	0.62553	0.28857	0.39235	97	1
ליידן	1	0.62551	0.28727	0.39150	34,60	2

חלוקות עם ציון הצללית הגבוה ביותר (עבור מודולריות גבוהה מ-0.5)

ניתן לראות כי אלגוריתמים לניתוח אשכולות מניבים את הציון הצללית הגבוה ביותר, כאשר שיטת ה-ward.d2 על מטריצת מרחקי מרחק על רשת הכבישים, בחלוקה ל-5 קבוצות מניבה את הציון הגבוה ביותר. עם זאת, מודולריות החלוקה הנובעת משיטה זו נמוכה יחסית, ולא מתקרבת להרצות באמצעות ליידן/לוביין ברזולוציה נמוכה. נעשה חיתוך של תוצאות אלו עבור מודולריות גבוהה מ-0.5 בלבד, על מנת למקד את הבדיקה בחלוקות בעלות משמעות תחבורתית גבוהה יותר. ראו טבלה 5:

טבלה 5: הרצות טובות יותר מבחינת צללית זמן

אלגוריתם	רזולוציה	מטריצה	כמות קבוצות בהרצה	כמות אזורי תנועה בקבוצה הקטנה	מודולריות	ציון צללית ממוצע לקבוצות (זמן כמרחק)	ציון צללית מינימלי לקבוצות (מרחק כמרחק)	גרעינים	כמות הרצות
וורד_2	-	מרחק	5	11	0.551	0.438	0.424	-	1
וורד_2	-	זמן	5	11	0.517	0.434	0.316	-	1
וורד_2	-	זמן	4	115	0.517	0.423	0.322	-	1
וורד	-	זמן	4	118	0.533	0.421	0.374	-	1
וורד_2	-	מרחק	4	146	0.551	0.421	0.321	-	1
לוביין	0.5	-	5	11	0.586	0.419	0.275	3,47	2
לוביין	0.5	-	5	11	0.586	0.419	0.276	72	1
לוביין	0.5	-	5	11	0.586	0.419	0.273	97	1
וורד	-	מרחק	4	146	0.549	0.418	0.325	-	1
ליידן	0.5	-	5	11	0.586	0.415	0.275	100,60	2

חלוקות ליידין/לוביין עם ציון הצללית הגבוה ביותר

הרצות ליידין / לוביין ברזולוציה 0.5 הניבו את ציון הצללית הגבוה ביותר מבין כלל הרצות של אלגוריתמים אלו, כאשר לוביין יותר בולטת מליידן. עם זאת, מודולריות החלוקות הנובעות מהרצות אלו נמוכה משמעותית מזו של החלוקות המובילות במודולריות. ראו טבלה 6:

טבלה 6: עבור הרצות עם מודולריות טובה, ההרצות הטובות ביותר של צללית זמן

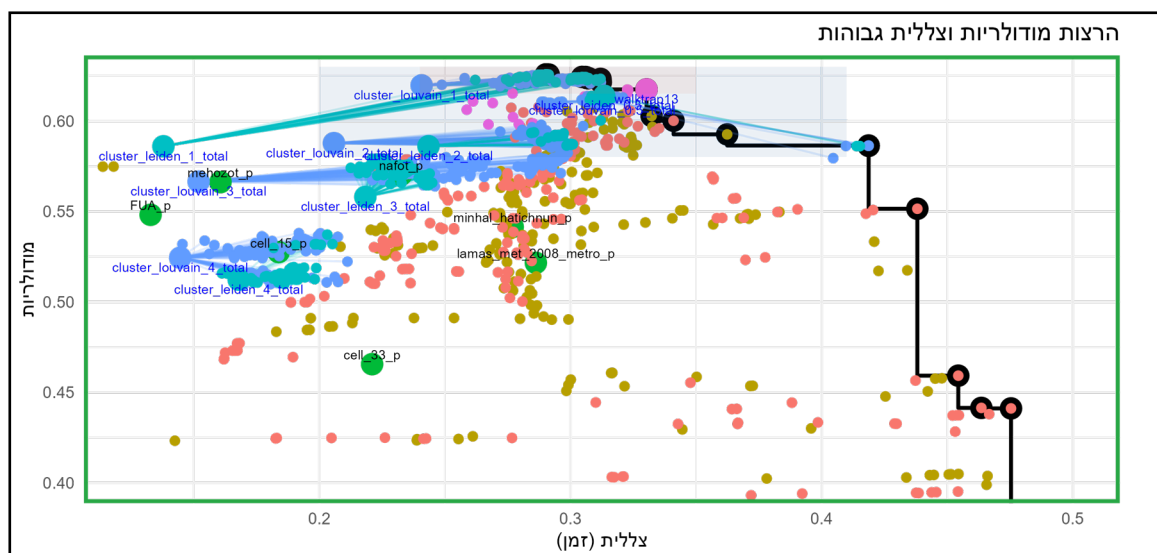
אלגוריתם	רזולוציה	כמות קבוצות בהרצה	כמות אזורי תנועה בקבוצה הקטנה	מודולריות	ציון צללית ממוצע לקבוצות (זמן כמרחק)	ציון צללית מינימלי לקבוצות (זמן כמרחק)	גרעינים	כמות הרצות
לוביין	0.5	5	11	0.586	0.419	0.234	47,3	2
לוביין	0.5	5	11	0.586	0.419	0.232	72	1
לוביין	0.5	5	11	0.586	0.419	0.233	97	1
ליידן	0.5	5	11	0.586	0.415	0.235	100,60	2
ליידן	0.5	5	11	0.586	0.414	0.235	23	1
לוביין	0.5	5	11	0.586	0.410	0.203	60	1
לוביין	0.5	4	170	0.579	0.405	0.137	73	1
לוביין	0.5	6	11	0.613	0.321	0.190	83	1
לוביין	0.5	6	11	0.613	0.319	0.212	52,84	2
לוביין	0.5	6	11	0.613	0.318	0.212	80	1

נספח ב' - גרפים ומפות משלימות

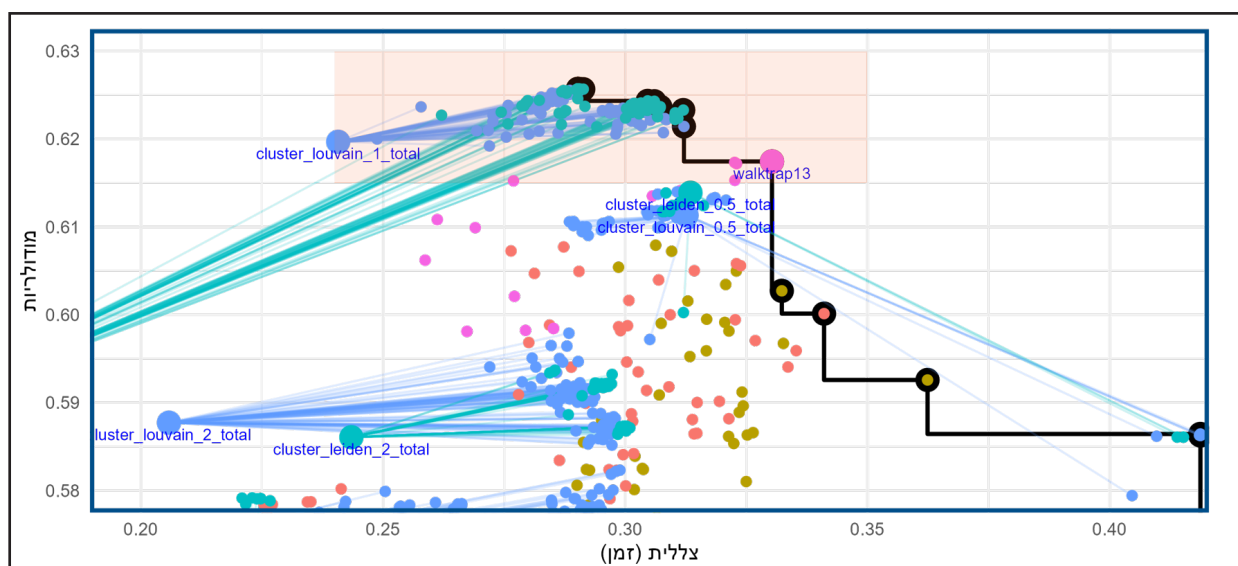
גרפי הקשר בין מודולריות לצלילית – זומים

הגרפים הבאים מציגים התמקדות ברמות שונות על הקשר בין מודולריות לצלילית, מזום רחב ועד זום צר על האזור בעל המודולריות הגבוהה ביותר.

איור 15: זום ממוקד לציונים סבירים



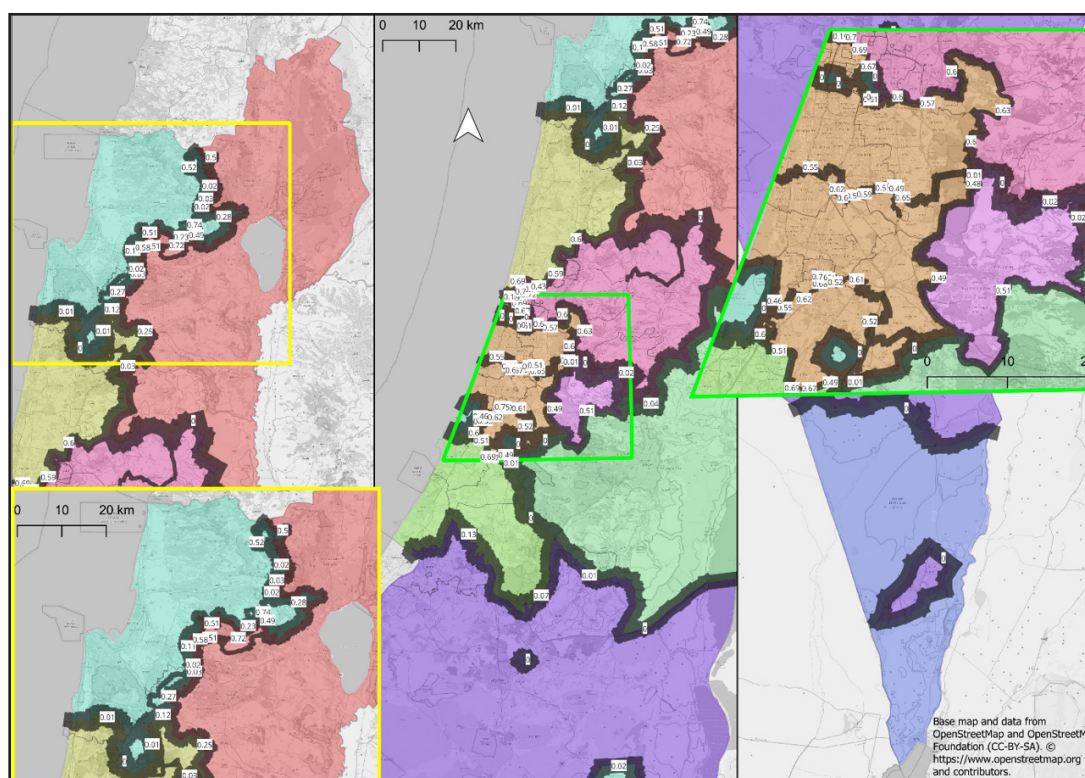
איור 16: זום ממוקד לחלוקות חדשות גבוהות יותר



מפות יציבות לוביין וליידן

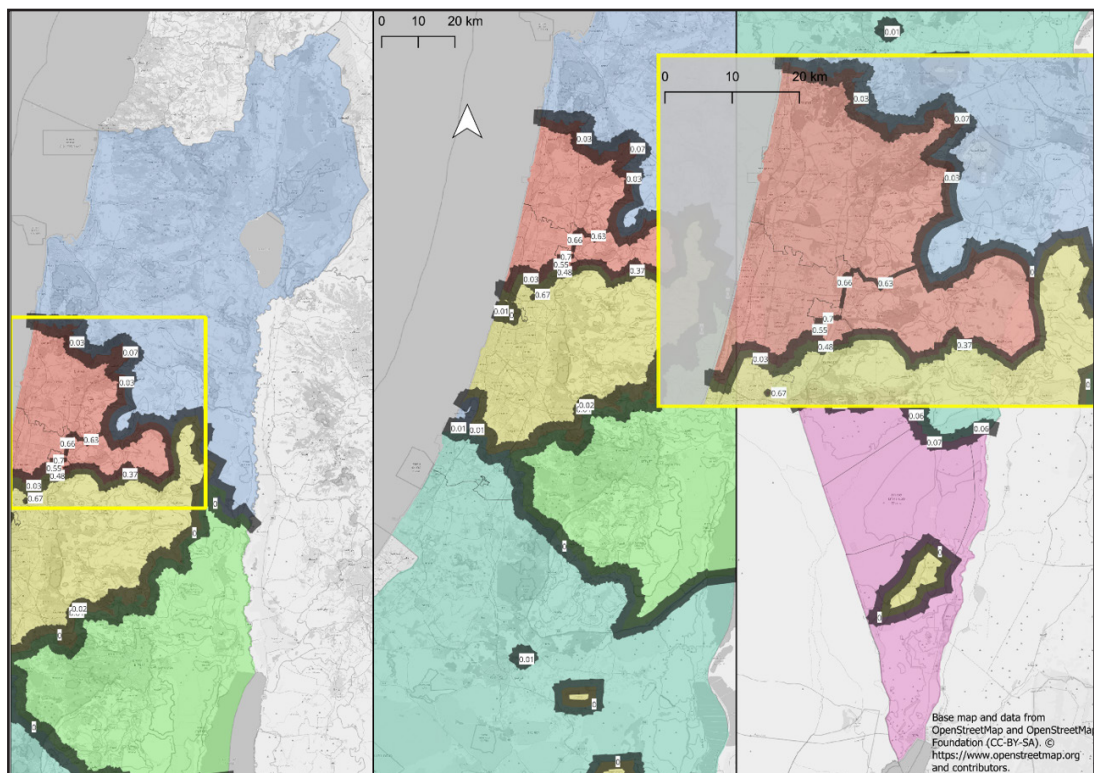
במפה הבאה (איור 18) ניתן לראות את החלוקה האגרטיבית של לוביין ברזולוציה 1, כאשר האזורים באותו צבע הם אזורים שהופיעו יחדיו ב-75% מההרצות. ניתן לראות כי קיימת יחסית באזור הצפון והדרום, בגוש דן קיימות חלוקות ביניים רבות. עובי קו הגבול מתאר כמות פעמים גדולה יותר בה האזורים הסמוכים לא הופיעו באותה הקבוצה.

איור 18: לוביין 1 אגרטיבי – אזורים המודגשים הופיעו יחדיו ב-75% מההרצות



כאשר מקטינים רזולוציה, כפי שניתן לראות באלגוריתם ליידין (איור 19), החלוקה נהיית יציבה הרבה יותר – אך עם מודולריות נמוכה יותר. לעיתים הגבולות בין האזורים נמצאים ב"הנדסת יתר", כמו למשל בלוביין באזור המרכז.

איור 19: ליידן 0.5 אגרגטיבי – חלוקה יציבה יותר עם מודולריות נמוכה יותר



נספח ג' - תהליך הניתוח המפורט

להלן הפירוט המלא של שלבי הניתוח, כולל הנוסחאות המתמטיות.

1. הכנה ועיבוד נתונים

- שימוש במטריצת נסיעות מוצא-יעד של כלל אזורי התנועה של הסקר הסלולרי, עבור כלל שעות היום, כמשקלות עבור מדידת מודולריות.
- שימוש במטריצת מרחקי זמן מוצא-יעד (בהתבסס על רשת הדרכים באוגוסט 2025; ראו נספח ד) לחישוב ניתוח אשכולות מרחבי ומדד צלילית.

2. חישוב מודולריות (Q)

חישוב מודולריות כללית של החלוקה על גרף המיוצג על ידי מטריצת ה-OD. עבור גרפים עם קשתות משוקללות (משקלים המייצגים כמות נסיעות) יש להשתמש בנוסחה המתאימה למודולריות משוקללת; סכימה על סך המשקל הכולל W . חישוב מודולריות פר אזור תנועה ופר קהילה: תרומת האזור למודולריות מחושבת כ- (משקל הקשתות הפנימיות – משקל צפוי תחת המודל האקראי) $W /$; נמדד גם יחס תרומה/דרגה.

נוסחת המודולריות:

$$Q = \frac{1}{2W} \sum_{i,j} \left(A_{ij} - \gamma \frac{k_i k_j}{2W} \right) \delta(c_i, c_j)$$

כאשר:

- A_{ij} = משקל הקשת בין צומת i לצומת j (כמות הנסיעות)
- k_i, k_j = הדרגה המשוקללת של צומתים i ו- j (סך הנסיעות היוצאות מכל צומת)
- W = סך כל המשקלות בגרף
- γ = פרמטר הרזולוציה (resolution parameter) המשמש לשליטה בגודל הקהילות המזוהות
- $\delta(c_i, c_j)$ = פונקציה המחזירה 1 אם i ו- j באותה קהילה, ו-0 אחרת

נוסחת המודולריות לכל קבוצה (קהילה):

$$Q_c = \frac{1}{2W} \sum_{i,j \in c} \left(A_{ij} - \gamma \frac{k_i k_j}{2W} \right)$$

כאשר c = קבוצה/קהילה מסוימת.

נוסחת תרומת הצומת (אזור תנועה) למודולריות:

$$Q_i = \frac{1}{2W} \left(\sum_{j \in c_i} A_{ij} - \gamma \frac{k_i \sum_{j \in c_i} k_j}{2W} \right)$$

כאשר c_i = הקבוצה שאליה משתייך צומת i . ערך חיובי גבוה יותר מעיד על תרומה משמעותית של הצומת לרציפות הקבוצה.

נוסחת היחס תרומה/דרגה:

$$\text{Contribution Ratio}_i = \frac{Q_i}{k_i}$$

מדד זה מנרמל את התרומה למודולריות על ידי דרגת הצומת, ומספק מדד של "יעילות" התרומה של כל צומת למודולריות קבוצתו.

3. חישוב צללית (s)

חישוב צללית פר תצפית על בסיס מטריצת מרחקי זמן: עבור כל אזור i חושב $a(i)$ = ממוצע המרחקים לאזורים באותה קבוצה ו- $b(i)$ = המינימום של ממוצעי המרחקים לקבוצות אחרות. חישוב צללית ממוצעת לכל אזור, קהילה/אשכול, ובצורה גלובלית.

נוסחת הצללית:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}$$

כאשר:

- $a(i)$ = ממוצע המרחקים מתצפית i לכל שאר התצפיות באותו אשכול
- $b(i)$ = המינימום של ממוצעי המרחקים מתצפית i לתצפיות בכל אשכול אחר
- הערך נע בין -1 ל-1, כאשר ערכים חיוביים מעידים על הקצאה טובה לאשכול

4. סריקות פרמטרים וריצות סטוכסטיות

- עבור אלגוריתמים סטוכסטיים (Louvain/Leiden) בוצעו 100 ריצות לפי seed בר שחזור, ונשמרו תוצאותיהן, עבור 5 משקלי רזולוציה שונים (0.5, 1, 2, 3, 4).
- עבור ניתוח היררכי/linkage בוצעו חישובים על מטריצת מרחקים + מטריצת זמנים, על כלל שיטות הקישור ההיררכי (linkage), עבור מספר קבוצות מ-2 עד 30. השיטות שנבחנו: single (קישור יחיד), complete (קישור מלא), average (ממוצע), mcquitty, ward.D (וורד), centroid (מרכזי), median (חציוני), ward.D2 (וורד מרובע), DIANA (דיביזיבי).

5. יציבות וקונסנזוס

עבור האלגוריתמים הסטוכסטיים, בוצע חישוב של כמות הפעמים שמספר אזורי מופיעים יחד. אוחדו אזורי שהיו יחדיו ב-75% מהפעמים. חישוב 75% מהפעמים בוצע על סמך הנתונים ובדיקה האם שיעור זה מייצר שבירה בנתונים. נתונים "מייצגים" אלו מוצגים בצורה בולטת יותר.

6. דיווח תוצאות

- ברמה גלובלית
- ברמה של קהילה/אשכול
- ברמה של אזור תנועה

7. בחירת חלוקות אופטימליות

בחירה רב-תבחינית: נבנה גרף המשקף את חזית הפארטו, המהווה את המצב בו אין דרך לשפר מדד אחד בלי לפגוע במדד אחר, בין מודולריות של חלוקה וצללית ממוצעת.

8. מיפוי

ימופו חלוקות מעניינות לפי מודולריות, צללית ומטריצת מוצא-יעד.

נספח ד' - השגת מטריצת מרחקי זמן מוצא יעד

התבצע חישוב המטריצה בצורה הבאה: בתוך כל אזור תנועה, נדגמו מספר נקודות בצורה עד כמה שיותר מפוזרות במרחב, בדגימת דיסק פואסנית, של כמות נקודות שחושב כשורש 0.3 של גודל שטח האזור. כלל הנקודות הוצמדו לרשת osm באמצעות אלגוריתם locate של ולהלה. הנקודות נוקו מכפילויות, ונשמרו רק נקודות ייחודיות. הופקה מטריצה זמני נסיעה/מרחקי נסיעה באמצעות אלגוריתם sources_to_targets של ולהלה המחשבת את המרחק בין כלל הנקודות בכל אזור תנועה. נבחרה הנקודה הקרובה ביותר למרכז הגיאומטרי של האזור, ונשמרה כמייצגת את אזור התנועה. ביצוע חישוב המסלול בין 1270 אזורי התנועה בוצע על הנקודות הללו באמצעות אלגוריתם sources_to_targets של ולהלה, המחשב את זמני הנסיעה בין כלל הנקודות על גבי רשת כבישים. זמן הנסיעה חושב בכל מקום ללא זרימת תנועה, עבור רכב פרטי ולפי מהירויות המונחות על ידי ולהלה. מקור רשת: OpenStreetMap (snapshot 2025-08-01) + תיקוני רשת מקומיים.

ד"ר עדו קליין הוא חוקר ויועץ בתחום התחבורה, המתמחה בניתוח נתונים, מודלים תחבורתיים ומערכות מידע גאוגרפיות (GIS). עבודתו משלבת שיטות חישוביות, ניתוח מרחבי ותהליכי קבלת החלטות מבוססי נתונים עבור גופים ציבוריים ופרטיים. תחומי העניין שלו כוללים תכנון תחבורה, מודלים של ביקוש לנסיעות, ניתוח רשתות, ניהול ידע ויישום טכנולוגיות קוד פתוח. הוא עוסק גם בפיתוח כלים ושיטות חדשניות לחקר מערכות מורכבות ולהנגשת מידע מקצועי למקבלי החלטות.

idshklein@gmail.com