פרויקט גמר · אלגוריתמים ברשתות

# מהי תחנה קריטית?

## מסגרת השוואתית לניתוח מרכזיות ועמידות ברשת התחבורה הציבורית בישראל

מתי "תחנה קריטית" היא באמת קריטית? עבודה זו בוחנת שמונה הגדרות בלתי-תלויות של קריטיות — מבנית, זרימתית, נפחית, מרחבית, תלוית-זמן, מבוססת-ביקוש, מבוססת-עלות-נוסע ומבוססת-הטמעות (embeddings) — ומראה כי הן מזהות תחנות שונות כמעט לחלוטין.


קורס: אלגוריתמים בגרפים / אלגוריתמים ברשתות

מנחה: ד"ר אבנר פריאל


מקור הנתונים: פיד ה-GTFS של משרד התחבורה (אפריל 2026)

יולי 2026

## תקציר

ניתחנו את מערך התחבורה הציבורית בישראל כרשת מורכבת שנבנתה מפיד ה-GTFS הארצי: 30,463 תחנות פעילות ו-51,772 קשתות, שנגזרו מ-15.7 מיליון רשומות לוח-זמנים ב-420,133 נסיעות. שאלת המחקר — *אילו תחנות הן הקריטיות ביותר* — התבררה כשאלה שאין לה תשובה יחידה. הממצא המרכזי: **קריטיות אינה תכונה אינהרנטית של הרשת אלא תוצר של המדד הנבחר.** שמונה הגדרות שונות של קריטיות מניבות דירוגים המסכימים ביניהם במתאם Spearman ממוצע של 0.21 בלבד; קריטיות תפעולית (נפח שירות) וקריטיות מבנית (betweenness) כמעט בלתי-תלויות ($\rho = 0.19$) — ולא בשל רעש דגימה, שכן רצפת-הרעש של המדד עצמו היא 0.96. השתמשנו במגוון שיטות — מדדי מרכזיות, נקודות-חיתוך וגשרים, גילוי קהילות Louvain, מודלי-אפס משמרי-דרגות, סימולציות תקיפה, משקול לפי זמן-נסיעה, פירוק לפי אופני-תחבורה, מודל גרביטציה לביקוש, מודל ניתוב-מחדש והטמעות Node2Vec — כל אחת עם בקרות שמפרידות אות אמיתי מארטיפקט. המסקנה בעלת ההשלכה המדיניותית: כל מדיניות הגנת-תשתית המבוססת על מדד קריטיות יחיד היא שרירותית.


## 1. מבוא ומוטיבציה

תחבורה ציבורית היא תשתית קריטית. השבתת תחנה מרכזית — בשל תאונה, עבודות, תקלה תפעולית או אירוע ביטחוני — עלולה להשפיע על אלפי נוסעים ולנתק אזורים שלמים. ברשת ארצית לא מספיק לדעת אילו תחנות עמוסות; יש להבין אילו תחנות הרשת *תלויה בהן מבנית*. השאלה נשמעת פשוטה, אך כפי שנראה — אין לה הגדרה אחת.

שאלת המחקר המרכזית:

> **אילו תחנות הן הקריטיות ביותר ברשת התחבורה הציבורית בישראל, וכיצד ניתן למדוד את תרומתן לקישוריות ולעמידות הרשת?**

הגישה הרווחת מזהה "תחנה קריטית" עם מדד בודד — לרוב betweenness centrality או נפח השירות. התרומה המרכזית של העבודה אינה אלגוריתם חדש אלא **מסגרת השוואתית ביקורתית**: אנו מפעילים שמונה הגדרות בלתי-תלויות של קריטיות על אותה רשת, ובוחנים עד כמה הן מסכימות. התוצאה — שהן כמעט אינן מסכימות — הופכת שאלה תמימה ("מהי התחנה הקריטית?") לטענה מחקרית לא-טריוויאלית: *בחירת המדד היא בחירה מדיניותית, לא טכנית*. לצד זאת אנו מקפידים על בקרות (רצפות-רעש, מודלי-אפס, תיקון FDR, negatives קשים) שמבדילות בין ממצא אמיתי לבין ארטיפקט של שיטה — ההבדל שבין עבודה מרשימה לבין עבודה נכונה.

## 2. עבודה קשורה

הבסיס התאורטי נשען על ניתוח רשתות מורכבות. Newman מתאר מדדים מבניים — דרגה, קשירות, קהילות ומרכזיות — ככלים להבנת מערכות תשתית ותחבורה, ומגדיר את החתימות המבחינות בין רשתות אקראיות, scale-free ועולם-קטן. Easley ו-Kleinberg מציגים רשתות כשפה כללית לניתוח זרימה ומבנה. מדד ה-betweenness centrality — צומת חשוב אם הוא נמצא על מסלולים קצרים רבים — מחושב ביעילות באלגוריתם של Brandes

(2001), שבו אנו משתמשים בגרסת דגימת-מקורות בשל גודל הרשת. לגילוי קהילות אנו משתמשים בשיטת Louvain (Blondel et al., 2008) הממקסמת modularity. את יומרת ה"scale-free" של רשתות מבוססות-מרכזים אנו בוחנים אל מול מודלי-אפס משמרי-מבנה, ברוח ההבחנה של Newman בין תכונה מבנית אמיתית לבין תוצר-לוואי של רצף הדרגות. לחיזוי קשרים אנו משווים הטמעות Node2Vec (Grover & Leskovec, 2016) למדדי דמיון קלאסיים, תחת פרוטוקול הערכה מתוקן-דליפה. תיקון ריבוי-ההשוואות נעשה בשיטת Benjamini–Hochberg (1995).

## 3. נתונים ובניית הגרף

מקור הנתונים הוא פיד ה-GTFS של משרד התחבורה, בתוקף מ-19.4.2026 עד 19.5.2026. הגדרנו את הגרף המכוון $G = (V, E, W)$: צומת הוא תחנה פעילה (מופיעה בלוח-זמנים לפחות פעם אחת); קשת מכוונת $u \to v$ קיימת אם קיימת נסיעה שבה $v$ מגיעה מיד אחרי $u$; והמשקל הוא מספר הנסיעות החוצות את המקטע (תדירות השירות). הגרף הלא-מכוון נבנה בסימטריזציה ומשמש לקשירות, נקודות-חיתוך, גשרים, קהילות ועמידות. את הקובץ הגדול (816MB, 15.7M שורות) קראנו בזרימה, בלי לטעון אותו לזיכרון.

**אימות ולא הנחה.** קוד נאיבי מניח שהפיד ממוין לפי (trip_id, stop_sequence). אנו בודקים זאת: מעבר-אימות על הקובץ המלא מצא 562 נסיגות ב-stop_sequence מתוך 15.7 מיליון שורות (0.004%), וללא שום שזירה בין נסיעות — הפרה זניחה, אך מדווחת ולא מוסתרת.

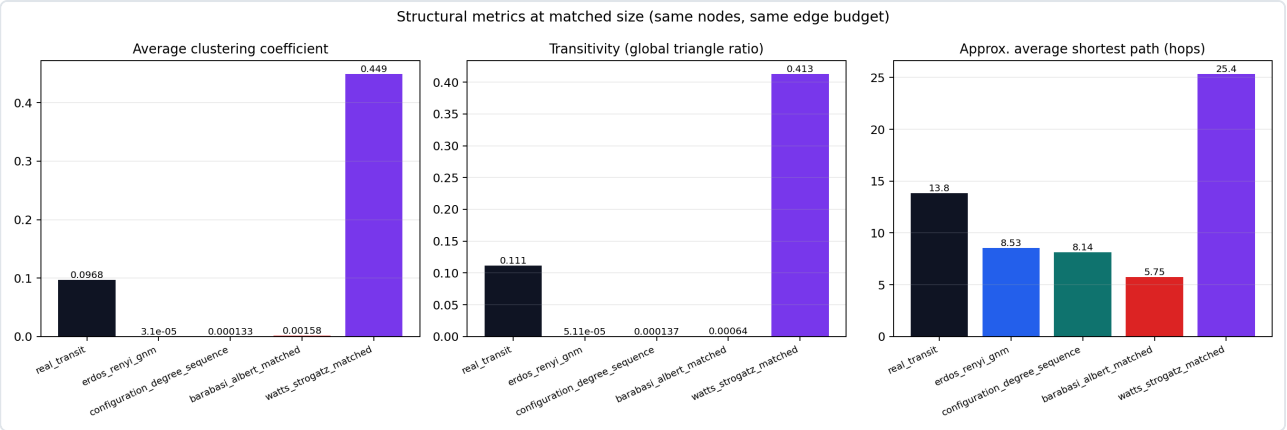
**טבלה 1: מאפייני הרשת הארצית**

| מדד | ערך | מדד | ערך |
|---|---|---|---|
| צמתים (תחנות) | 30,463 | רכיב קשיר גדול | 30,237 (99.26%) |
| קשתות לא-מכוונות | 51,772 | דרגה ממוצעת | 3.40 |
| צפיפות | 0.000112 | דרגה מרבית | 62 |
| רכיבי קשירות | 12 | נקודות-חיתוך | 900 (2.95%) |
| מסלול-קצר ממוצע | 13.8 קפיצות | גשרים | 971 (1.9%) |

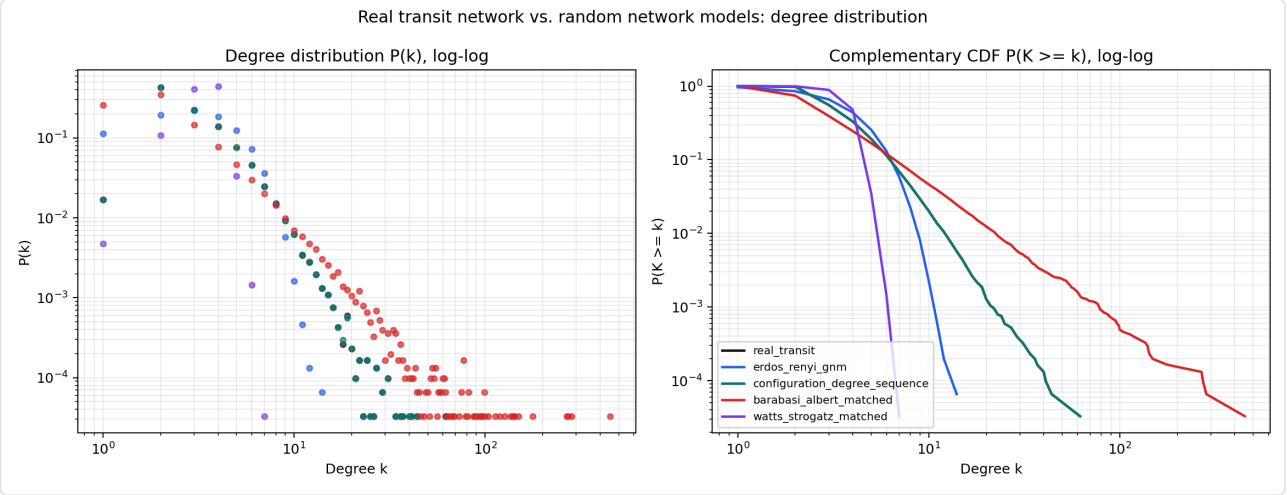
## 4. מבנה הרשת: עולם-קטן שאינו scale-free

הציפייה התמימה מרשת מבוססת "מרכזי תחבורה" היא שתהיה scale-free — עם מפרקי-על בעלי דרגה עצומה. הבדיקה מפריכה זאת. כדי לבדוק אילו תכונות מבניות הן *אמיתיות* ולא תוצר-לוואי של רצף הדרגות, השווינו את הרשת לארבעה מודלי-אפס בגודל תואם: Erdős–Rényi, מודל configuration משמר-דרגות, Barabási–Albert (scale-free) ו-Watts–Strogatz (סריג עולם-קטן).

הבקרה החדה ביותר היא מודל ה-configuration: הוא משחזר את רצף הדרגות במדויק ($KS = 9.8\times10^{-5}$) אך מפרק כמעט את כל מבנה המשולשים. מקדם ה-clustering של הרשת האמיתית, 0.097, גבוה פי 726 מזה של המודל משמר-הדרגות ($1.3\times10^{-4}$) ופי 3,123 מ-Erdős–Rényi. המשמעות: ההתקבצות ברשת אינה נובעת מהצמתים עתירי-הדרגה — היא **חתימה מרחבית-גאוגרפית** אמיתית. במקביל, זנב הדרגות של הרשת דק בהרבה מזה של מודל Barabási–Albert (דרגה מרבית 62 מול 452, פי 7.3; אומדן Hill $\alpha \approx 3.46$, תלול מתחום ה-2–3 של scale-free) — הרשת *אינה* scale-free. עם זאת היא כן עולם-קטן מובהק ($\sigma = 1924$).

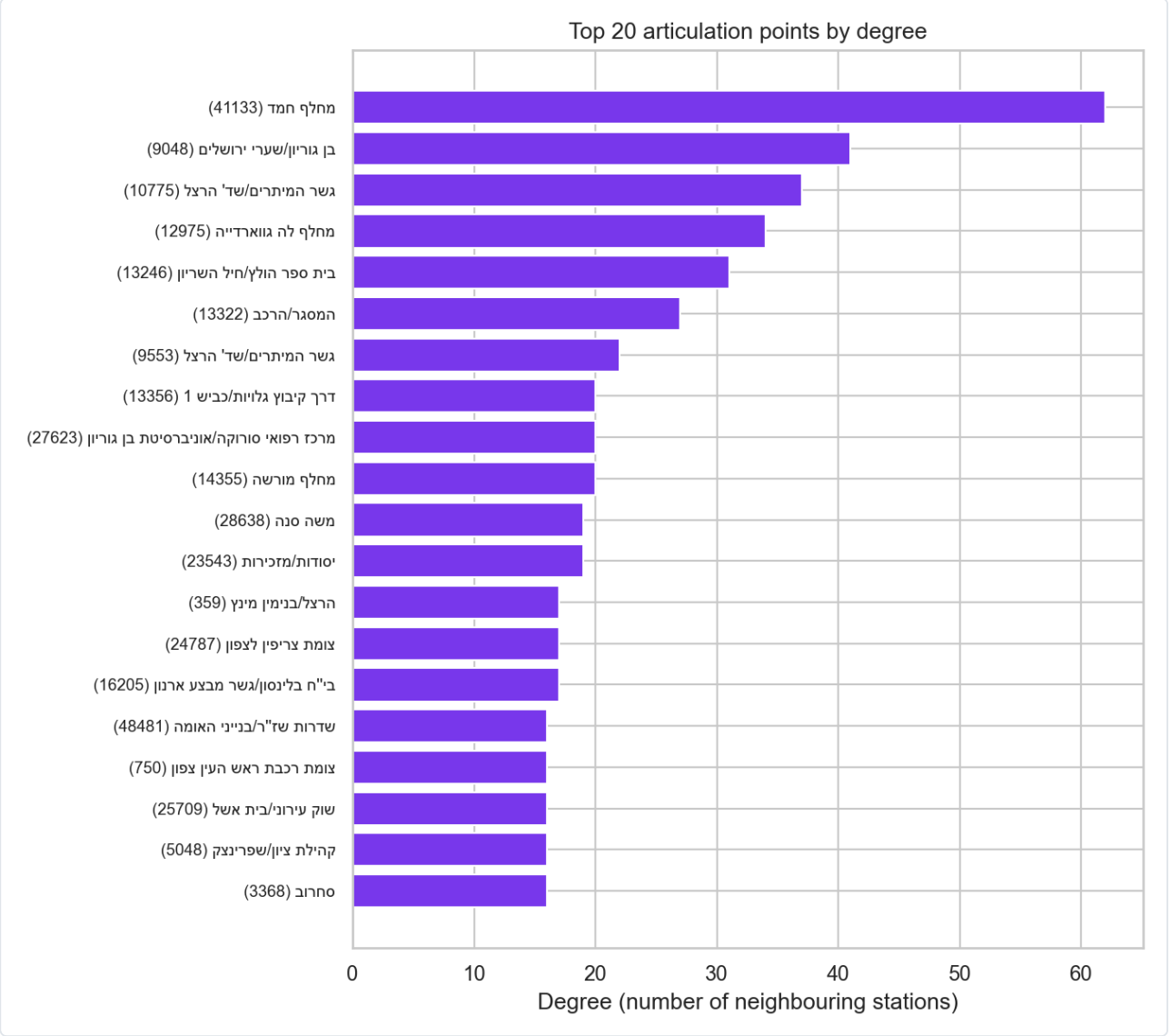



**איור 1.** השוואת מדדים מבניים בין הרשת האמיתית לארבעה מודלי-אפס בגודל תואם. מקדם ה-clustering של הרשת (0.097) גבוה פי כ-726 מזה של מודל ה-configuration בעל רצף-הדרגות הזהה ($1.3\times10^{-4}$), והמסלול הקצר-ביותר הממוצע (13.8 קפיצות) ארוך פי 1.62 מ-Erdős–Rényi; רק סריג Watts–Strogatz מייצר clustering ומסלולים ארוכים מן הרשת.




**איור 2.** התפלגות הדרגות וההתפלגות המצטברת המשלימה P(K≥k) בסקאלה log-log. זנב הרשת חופף למודל ה-configuration בלבד; Erdős–Rényi בעל זנב דק מדי (חתך חד סביב k=14), ואילו Barabási–Albert מייצר זנב שמן שמגיע ל-~450 — ומכאן שהרשת אינה scale-free.

ממצא נוסף מהפך אינטואיציה: ההצמדות לגאוגרפיה *מאריכה* את המסלולים. המסלול-הקצר הממוצע ברשת האמיתית (13.8 קפיצות) ארוך מכל מודל אקראי — פי 1.62 מ-Erdős–Rényi ופי 2.4 מ-Barabási–Albert. עֵרְבּוּב הקשתות היה מקצר את המסלולים; המחיר המרחבי מדיד וממשי. לבסוף, הרשת **חסינה-בקשתות אך שבירה-בצמתים**: רק 1.9% מהקשתות הן גשרים (98.1% מגובות במעגלים), אך 900 צמתים הם נקודות-חיתוך — בדיוק ה"תחנות הקריטיות" שבמוקד העבודה (איור 3).

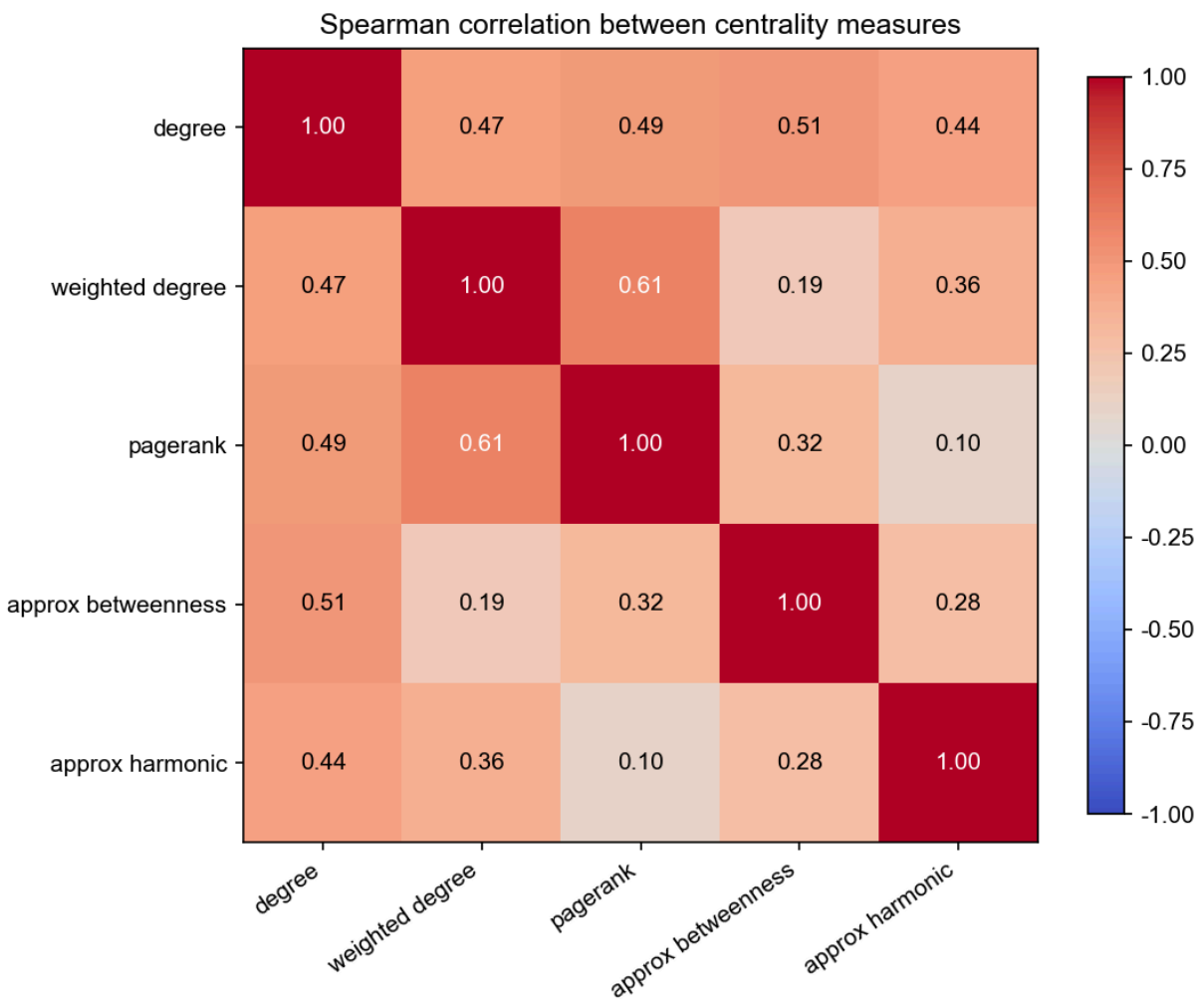



**איור 3.** עשרים נקודות-החיתוך בעלות הדרגה הגבוהה ביותר — צמתים שהסרתם מפצלת את הרשת. הצומת המחובר ביותר, מחלף חמד (ירושלים, דרגה 62), הוא גם נקודת-כשל יחידה.

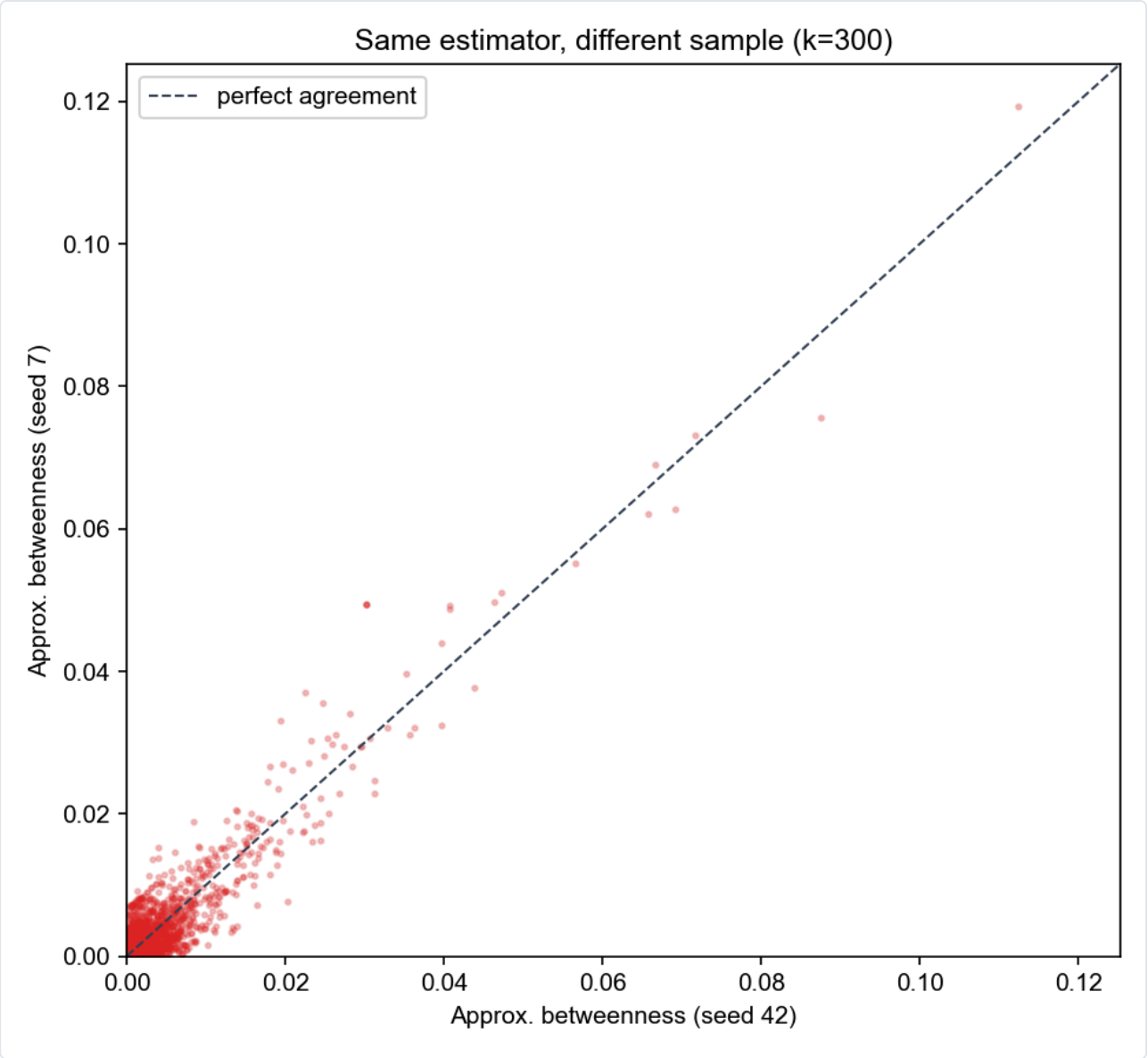
## 5. שתי פנים של קריטיות: תפעולית מול מבנית

כאן מתחילה ההפרכה של הנחת-היסוד. חישבנו לכל תחנה חמישה מדדי מרכזיות — דרגה, דרגה משוקללת, PageRank, approx betweenness (דגימת 300 מקורות) ו-harmonic — ובדקנו את המתאם ביניהם. קריטיות תפעולית (נפח שירות, נמדד ב-weighted degree) וקריטיות מבנית (צווארי בקבוק, נמדדים ב-betweenness) כמעט **בלתי-תלויות**: $\rho = 0.19$ בלבד (איור 4).

ההתנגדות המתבקשת — "אולי זה רעש הדגימה של betweenness" — נדחית באמצעות בקרה. הרצנו את האומדן פעמיים בזרעים שונים: ההסכמה בין שתי ההרצות היא $\rho = 0.96$ (חפיפת Top-50 של 0.84). כלומר רצפת-הרעש היא 0.96, והמתאם הצולב הנמוך (0.19) הוא **דיסוציאציה אמיתית** ולא ארטיפקט (איור 5). התוצאה מוחשית ברמת התחנה הבודדת: התחנה העמוסה ביותר, קניון עזריאלי (דרגה משוקללת 21,396), בעלת betweenness אפסי (0.0008); ואילו התחנה הראשונה ב-betweenness — ת.התרעננות/דור אלון מגל 6 — נושאת פי 97 פחות תנועה אך betweenness גבוה פי 136.

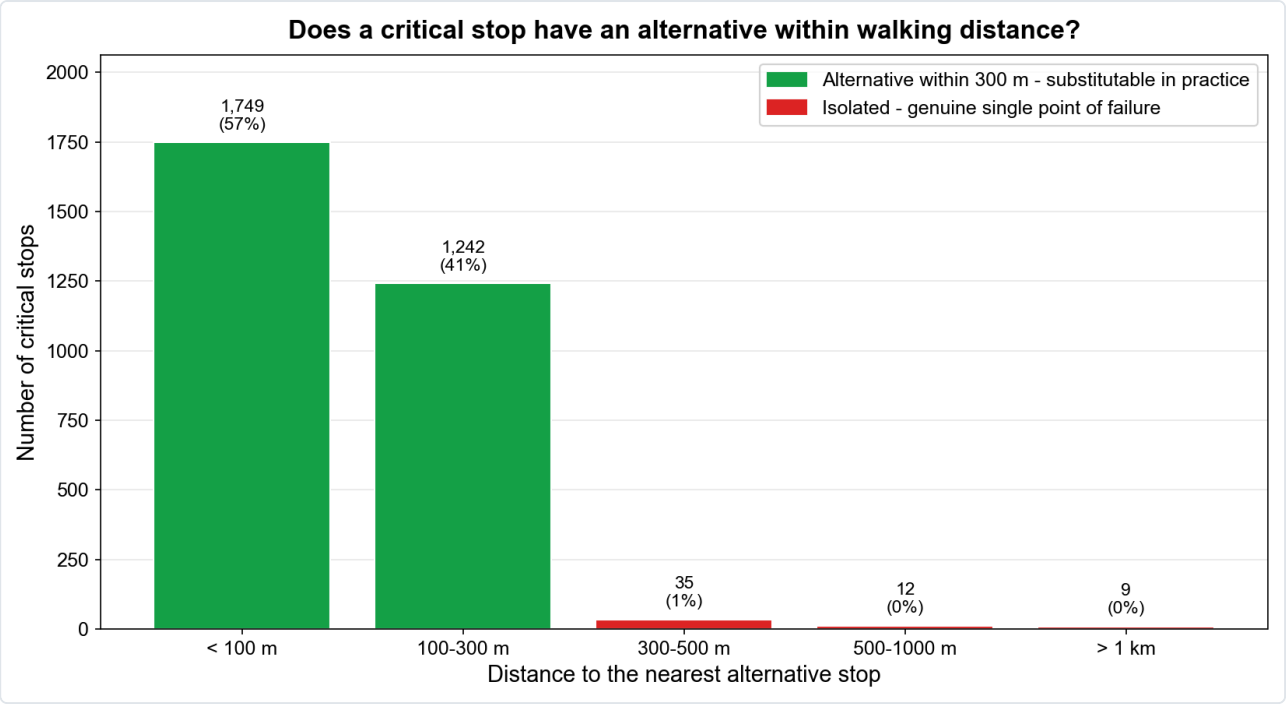



**איור 4.** מטריצת קורלציית Spearman בין חמשת מדדי המרכזיות. הקורלציה בין weighted degree (נפח) לבין approx betweenness (צווארי בקבוק) עומדת על 0.19 בלבד, ובין PageRank ל-harmonic על 0.10 — מרכזיות תפעולית ומבנית כמעט בלתי-תלויות.




**איור 5.** בקרת-רעש. אומדן approx betweenness עבור שני זרעים בלתי-תלויים (42 ו-7). הנקודות צמודות לקו ההסכמה המושלם (ρ = 0.96), ולכן הקורלציה הצולבת הנמוכה (0.19) אינה תוצר של רעש דגימה אלא דיסוציאציה אמיתית.

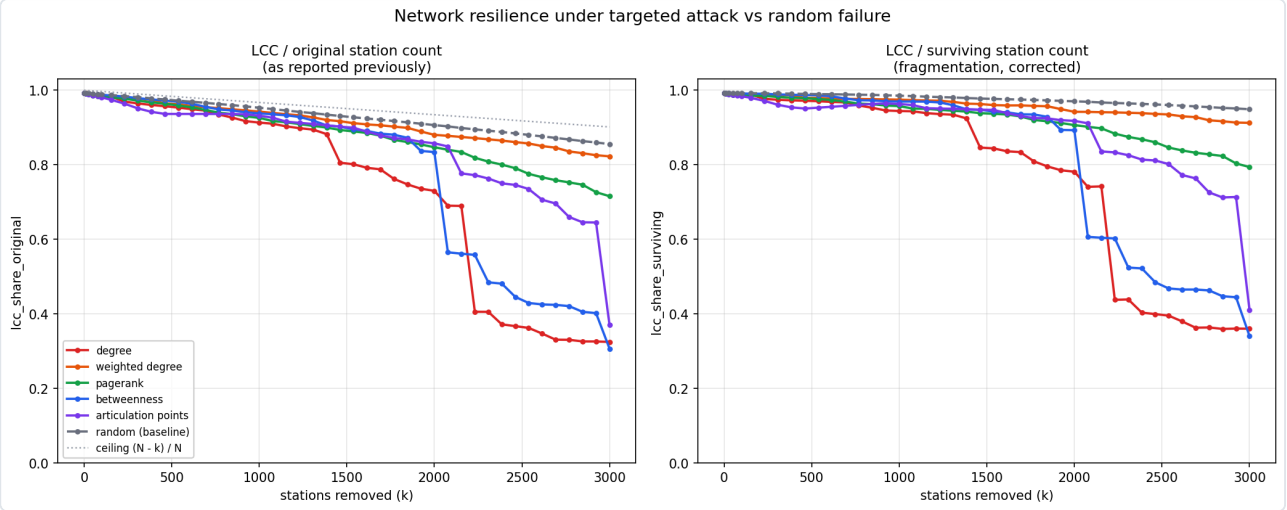
גם ההגדרה השנייה — "תחנה קריטית = תחנה שאין לה חלופה" — מתמוטטת כשמכניסים את המציאות המרחבית. מתוך 3,047 התחנות שסומנו קריטיות (10% העליונים ב-betweenness), **98.2% ברות-החלפה**: קיימת תחנה חלופית עד 300 מ' (57.4% אף עד 100 מ'). רק 56 תחנות (1.8%) הן נקודות-כשל בודדות אמיתיות — וכמעט כולן מחלפי-פריפריה, קיבוצים ומושבים, לא ליבות מטרופולין (איור 6). קריטיות טופולוגית אינה מתורגמת לפגיעות תפעולית כשקיימת חלופה במרחק הליכה.




**איור 6.** האם לתחנה קריטית יש חלופה במרחק הליכה? שתי העמודות הירוקות (98.2% ברות-החלפה) מגמדות את שלוש האדומות (1.8% מבודדות). ממצא כן ש״מרסן״ את יומרת הקריטיות.

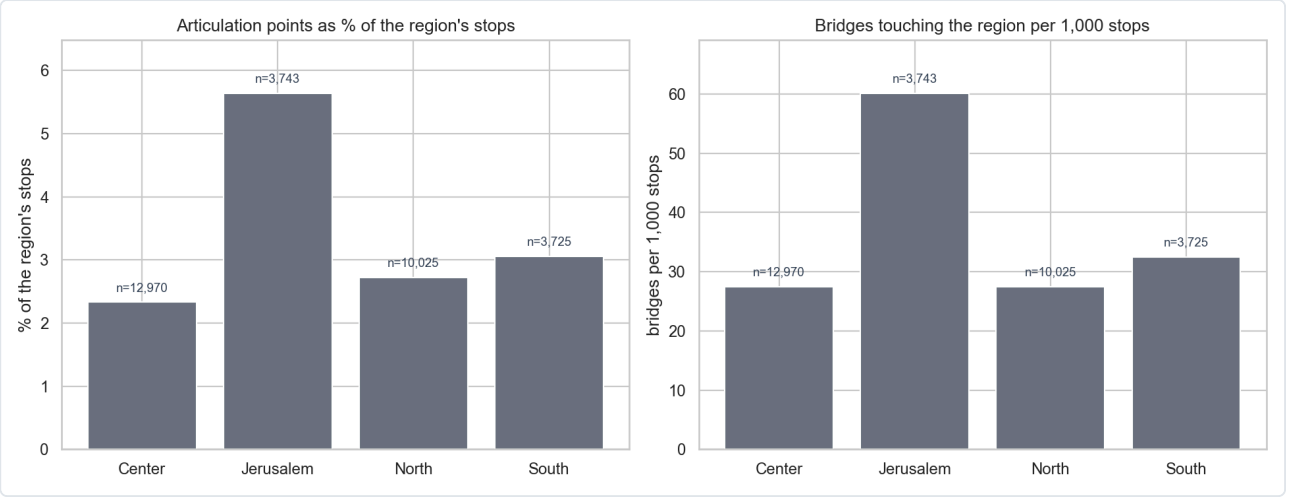
## 6. עמידות: חסינה-אך-שבירה, ושתי גאוגרפיות של פגיעות

סימולציית התקיפה הסירה תחנות לפי שש אסטרטגיות מול בסיס אקראי, תוך מעקב אחר גודל הרכיב הקשיר הגדול (איור 7). הרשת מפגינה חתימת ״חסינה-אך-שבירה״: כשל אקראי כמעט אינו פוגע (AUC = 0.928, צמוד לתקרת N-(k)/N), בעוד תקיפה מכוונת הרסנית — פער של פי 2.7 ב-k=3000. מפתיע: התקיפה היעילה ביותר לחישוב, לפי דרגה, היא גם ההרסנית ביותר במצטבר (AUC = 0.734), ואילו betweenness יוצרת קריסה חדה ומאוחרת — שתי חתימות-תקיפה שונות.




**איור 7.** עקומות חוסן תחת תקיפה ממוקדת מול כשל אקראי. משמאל — נרמול מ-N המקורי; מימין — נרמול מהתחנות ששרדו (מדד פרגמנטציה מתוקן). הסרה לפי דרגה (אדום) גורמת לנזק המצטבר הגדול ביותר; betweenness (כחול) יוצרת קריסה כמעט-אנכית סביב k≈2000; כשל אקראי (אפור) צמוד לתקרה.

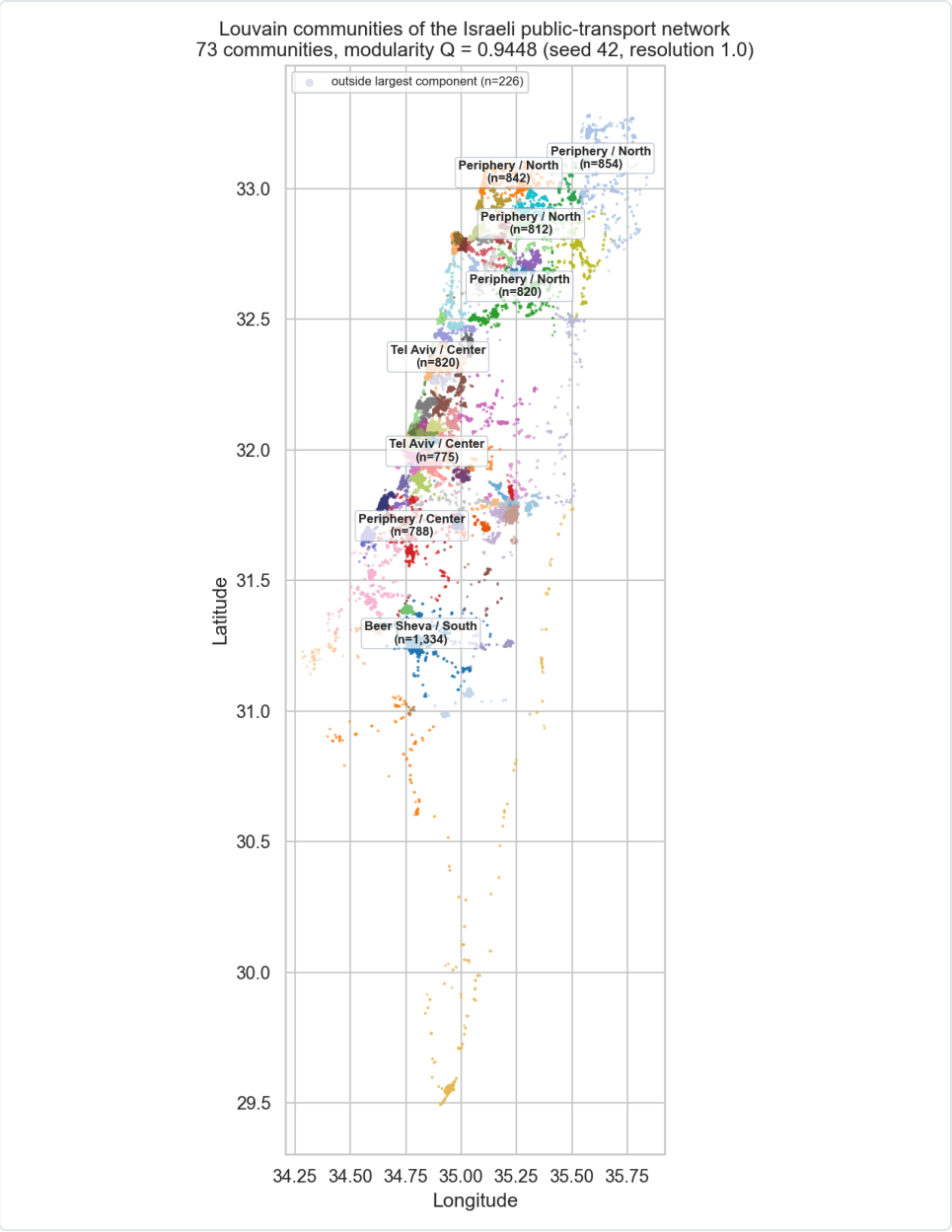
הפגיעות האזורית חושפת ש״פגיעות״ עצמה אינה מושג יחיד. הצפון מוביל בשיעור צווארי-הבקבוק (11.55% תחנות קריטיות מול 8.91% במרכז), אך **ירושלים** חריגה במדד אחר לגמרי: ריכוז נקודות-חיתוך (5.64%, פי 2.4 מהמרכז) וצפיפות גשרים (60.1 ל-1,000 תחנות, פי 2.2 מהצפון) — מבנה עצי דל-יתירות שמקורו בטופוגרפיה ההררית (איור 8). חשוב לא פחות: *האזור הגאוגרפי כמעט אינו מנבא קריטיות*. מבחן chi-square מובהק ($p = 7\times10^{-10}$) רק משום ש-n = 30,463, אך גודל האפקט זניח (Cramér's V = 0.039); מלוא הפער בין האזורים הוא 2.64 נקודות אחוז. זו תוצאה שלילית כנה שמרסנת פרשנות-יתר.




**איור 8.** שבריריות מבנית לפי אזור בשני מדדים בלתי-תלויים: שיעור נקודות-חיתוך (משמאל) וצפיפות גשרים ל-1,000 תחנות (מימין). ירושלים חריגה בשני המדדים בו-זמנית — עדות למבנה עצי דל-יתירות, ולא לצווארי-בקבוק של betweenness (שם דווקא הצפון מוביל).

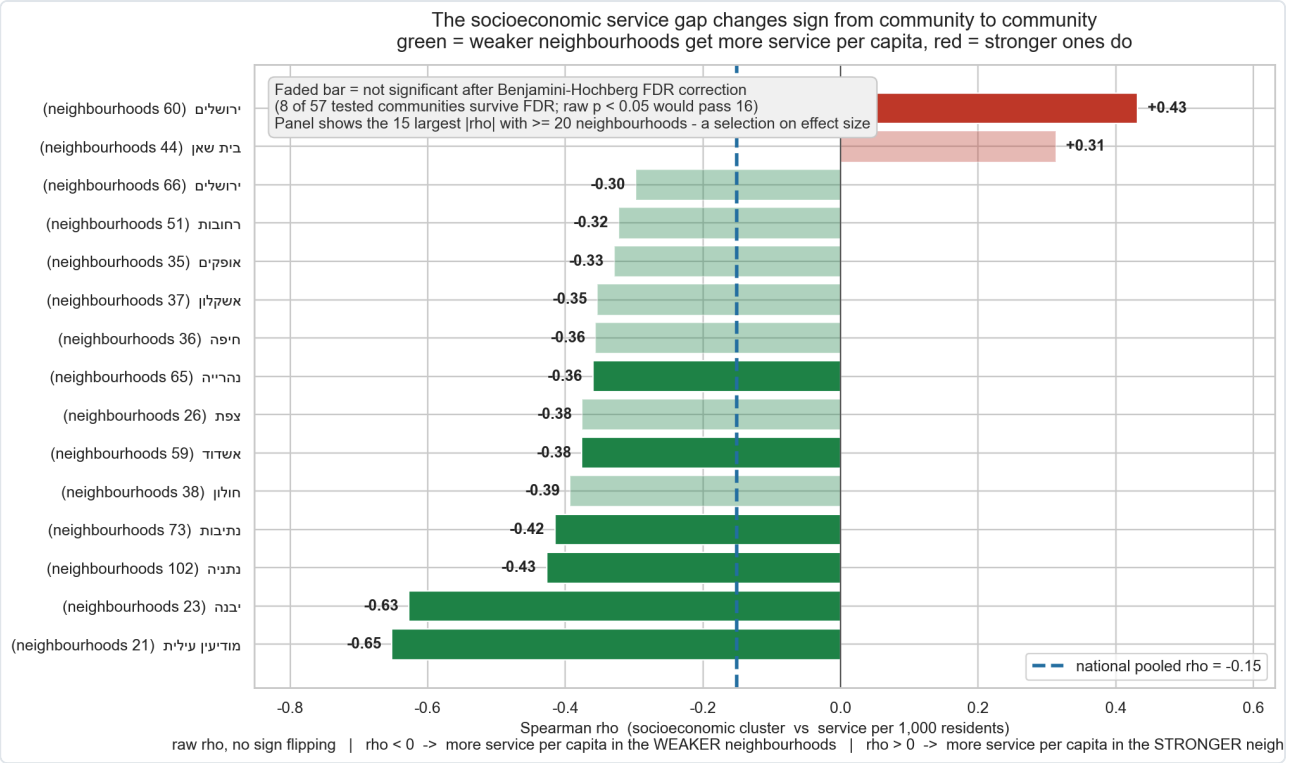
## 7. קהילות ושוויוניות: פרדוקס Simpson

אלגוריתם Louvain מצא **73 קהילות** עם מודולריות קיצונית (Q = 0.9448): רק 3.2% ממשקל הקשתות חוצה גבולות קהילה. הרשת מתפקדת כאוסף ״איים״ אזוריים מקושרים רופפות, לא כרשת ארצית משולבת. הקהילות נובעות מהטופולוגיה בלבד אך משחזרות את הגאוגרפיה הממשית (איור 9), והמבנה יציב על פני זרעים (ARI = 0.77–0.86) ורזולוציות — כלומר מאפיין אמיתי ולא תוצר של בחירת האלגוריתם (בסיס label-propagation נותן Q = 0.55 בלבד).



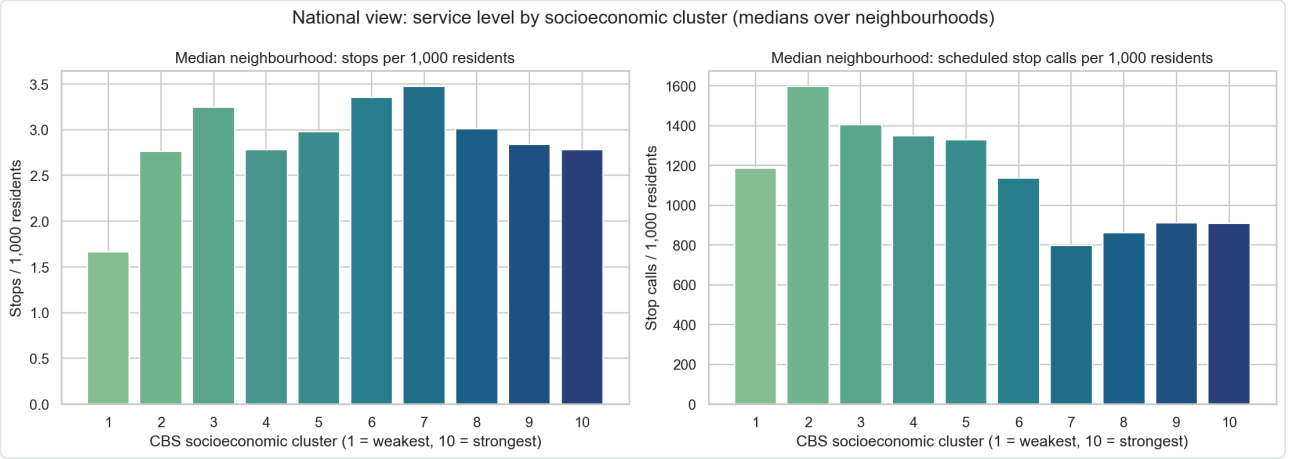

**איור 9.** מפת 73 קהילות Louvain (Q = 0.9448). קהילות שזוהו מהמבנה הטופולוגי בלבד משחזרות את הגאוגרפיה: אשכולות מרחביים קוהרנטיים סביב באר-שבע/דרום (n=1,334), מטרופולין תל-אביב, וקהילות פריפריה/צפון.

הקהילות אינן רק ממצא מבני — הן המפתח לתוצאה החברתית החזקה ביותר בעבודה. ההשערה התמימה "שכונות חלשות מקבלות פחות שירות" מתגלה כבלתי-ניתנת-להגדרה ברמה הארצית. המתאם הארצי המצרפי בין דירוג סוציו-אקונומי לעצימות השירות הוא $\rho = -0.15$ — אך זהו **ארטיפקט אקולוגי**. בתוך קהילות בודדות הסימן *מתהפך*: מ-−0.65 במודיעין עילית (יותר שירות לנפש בשכונות החלשות) עד +0.43 בקהילת ירושלים (איור 10). זהו פרדוקס Simpson קלאסי: המתאם המצרפי אינו מייצג אף קהילה בודדת. יתרה מכך, רק 8 מתוך 57 הקהילות שנבדקו שורדות תיקון Benjamini–Hochberg ל-FDR (מול 16 שהיו "מובהקות" ללא תיקון) — היושרה הסטטיסטית היא מה שהופך את הממצא לאמין.




**איור 10.** מתאם Spearman בין דירוג סוציו-אקונומי לעצימות שירות, בנפרד בתוך כל קהילת Louvain. הסימן מתהפך בין קהילות (ירוק = יותר שירות לחלשים, אדום = ההפך). עמודות דהויות אינן מובהקות לאחר תיקון FDR (רק 8 מ-57 שורדות). הקו המקווקו מסמן את המתאם הארצי המצרפי 0.15−.

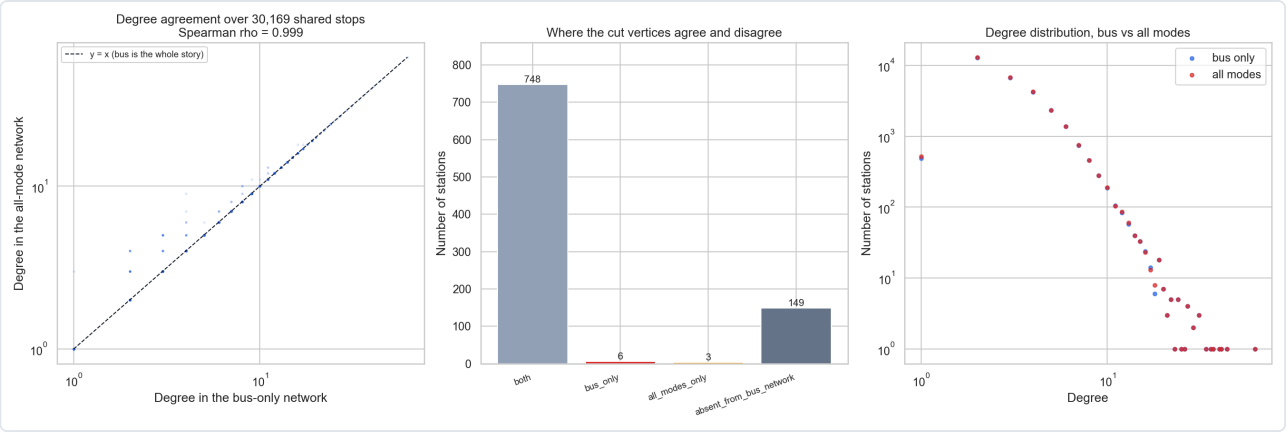
ולבסוף, אפילו כיוון המסקנה תלוי במדד: לפי *מספר עצירות* לנפש (תשתית) המתאם הוא 0.03+ (החלשים מקבלים מעט *פחות*), ואילו לפי *עצירות-נסיעה מתוזמנות* לנפש (עצימות) הוא 0.15− (החלשים מקבלים יותר) — איור 11. שוויוניות התחבורה ברשת הזו היא תכונה מקומית ותלוית-מדד, לא תכונה ארצית.




**איור 11.** חציון רמת השירות לפי אשכל סוציו-אקונומי. משמאל — עצירות ל-1,000 תושבים (תשתית): הנמוך ביותר דווקא באשכל החלש. מימין — עצירות-נסיעה מתוזמנות ל-1,000 (עצימות): הגבוה ביותר דווקא באשכלים החלשים. שני המדדים נותנים כיוון-אקוויטי הפוך.

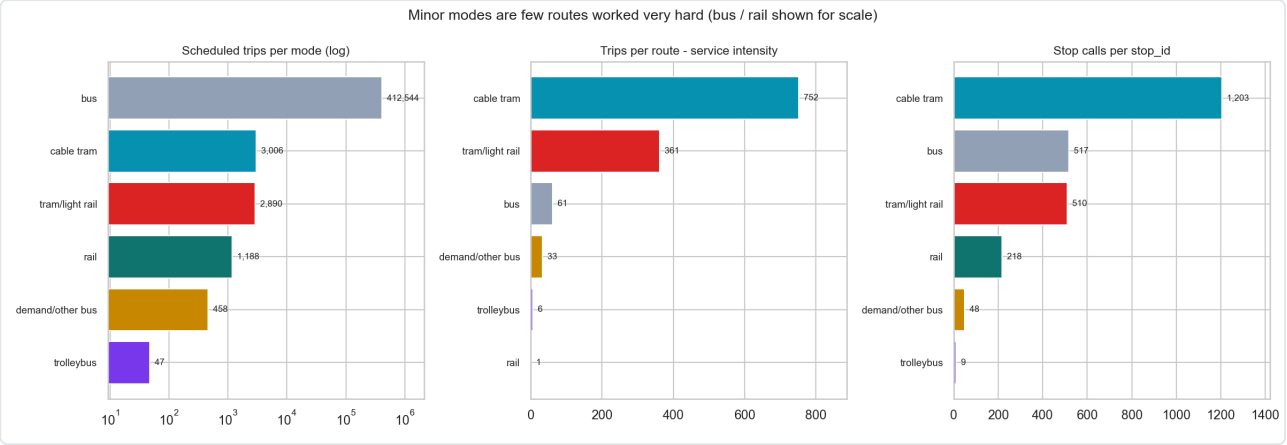
## 8. פירוק לפי אופני-תחבורה

הרשת הארצית מסתירה שונות מבנית עצומה בין אופני-התחבורה. פירקנו אותה לשש רשתות נפרדות. ראשית, בקרה: האוטובוס מהווה 98.2% מהנסיעות, ולכן הרשת הארצית זהה כמעט לחלוטין לרשת האוטובוס בלבד (מתאם דרגות ρ = 0.999, זהות ב-99.78% מהתחנות; איור 12) — יש להיזהר מלייחס ל"מרכזיות כל-המודים" משמעות עצמאית.




**איור 12.** השוואת רשת האוטובוס לרשת כל-המודים על 30,169 תחנות משותפות: התאמת דרגות כמעט מושלמת (ρ = 0.999), חלוקת נקודות-חיתוך כמעט זהה, והתפלגויות דרגות חופפות. הוספת חמישה מודים כמעט אינה משנה את השלד המבני.

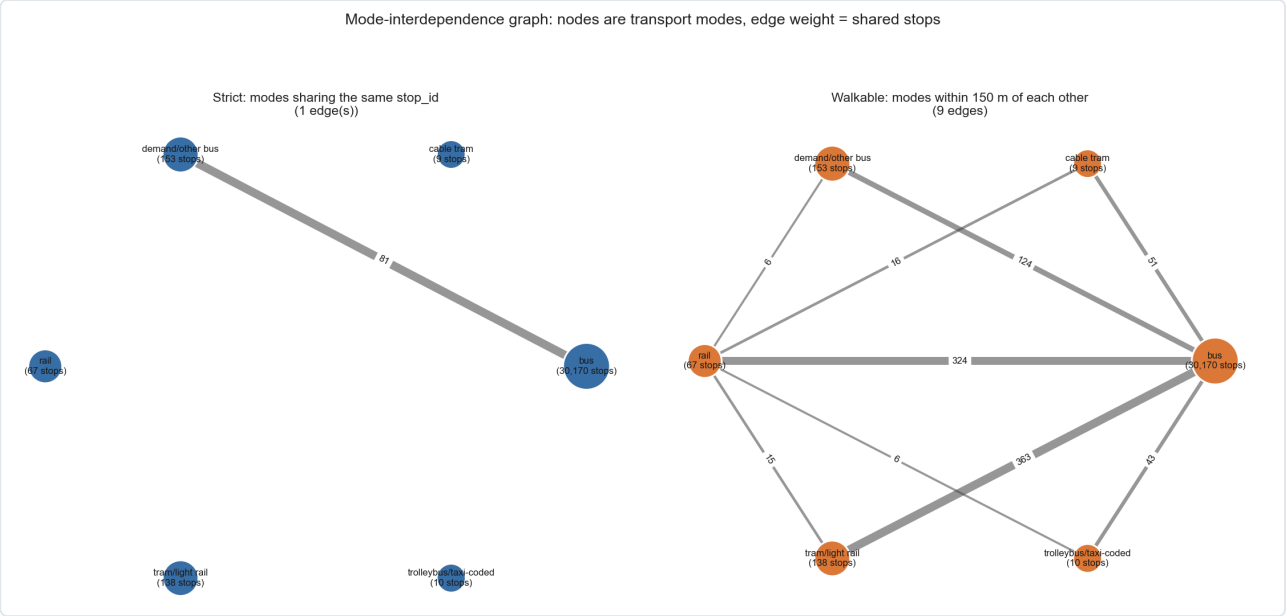
אך המודים הקטנים מספרים סיפור הפוך. הרכבת הקלה היא 138 תחנות שבהן **92.75% נקודות-חיתוך ו-100% מהקשתות גשרים** (מספר ציקלומטי אפס) — טופולוגיית מסלול/עץ חסרת-יתירות לחלוטין, שבה כל מקטע הוא נקודת-כשל יחידה. ועם זאת, דווקא המודים הזעירים נושאים את עצימות-השירות הגבוהה ביותר: הרכבל עם 1,203 עצירות לתחנה והרכבת הקלה עם 361 נסיעות לקו (פי 5.9 מהאוטובוס) — מעט קווים המופעלים באינטנסיביות קיצונית על שלד שברירי (איור 13).




**איור 13.** אף שהאוטובוס מהווה 98.2% מהנסיעות, עצימות השירות לקו ולתחנה גבוהה דווקא במודים הקטנים — cable tram 752 נסיעות/קו ו-1,203 עצירות/תחנה, מול 61 בלבד באוטובוס.

גם ההגדרה של "צומת מעבר רב-אופני" תלוית-מדד. לפי הגדרת ה-GTFS המחמירה (stop_id משותף) יש בישראל 81 תחנות מעבר בלבד, וכולן bus↔demand-bus — הרכבת, הרכבת הקלה והרכבל *לעולם* אינם חולקים stop_id עם דבר, שכן המודל מקצה מזהה נפרד לכל רציף. תחת הגדרה זו אין לתחנות-המעבר שום חשיבות מבנית (OR = 1.26, לא-מובהק). אך תחת הגדרה מרחבית (מודים בטווח הליכה של 150 מ') נמצאות 866 תחנות-מעבר אמיתיות, והן

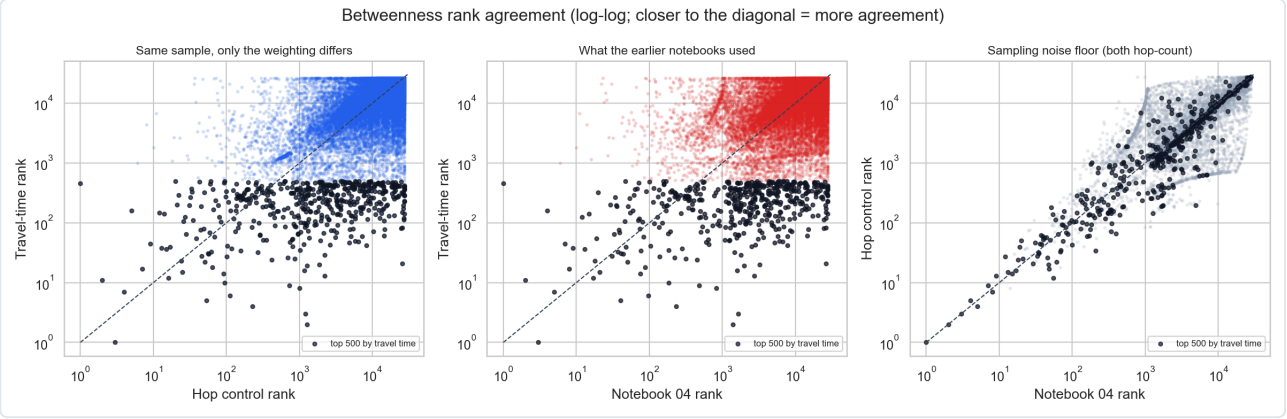
נוטות פי 9.1 להיות נקודות-חיתוך (18.9% מול 2.5%), אפקט ששורד תיקון לדרגה (OR = 7.57). המעבר הרב-אופני בישראל הוא עובדה מרחבית שמודל הנתונים מסתיר (איור 14).




**איור 14.** גרף תלות-הדדית בין מודים (משקל הקשת = מספר התחנות המשותפות). בהגדרה המחמירה (משמאל) קיימת קשת אחת בלבד; בהגדרת ההליכתיות (מימין) האוטובוס משמש מרכז לרשת-כוכב. מעבר רב-אופני הוא תופעה של קרבה מרחבית, לא של רציף משותף.

## 9. זמן-נסיעה, ביקוש וניתוב-מחדש

עד כאן, כמו ברוב הספרות, "מרחק" נמדד במספר קפיצות. זו מגבלה, וכימתנו אותה. בנינו גרף משוקלל בזמן-נסיעה מתוך שדות ה-arrival_time (15.2 מיליון תצפיות תקפות; זמן חציוני למקטע 66 שניות). התוצאה חדה: מסלול לפי קפיצות ומסלול לפי זמן זהים ב-**0.33% מהמקרים בלבד**, וניתוב לפי קפיצות עולה בחציון 2,443 שניות (~41 דקות) לנסיעה. חשוב מכך, betweenness משוקלל-זמן מתואם עם betweenness מבוסס-קפיצות ב-$\rho = 0.61$ בלבד — אל מול רצפת-רעש של 0.96 (איור 15). כלומר *זהות התחנות המרכזיות משתנה ממשית* כשעוברים לזמן אמיתי; זהו סייג על כל תוצאות ה-betweenness הטופולוגיות שקדמו.

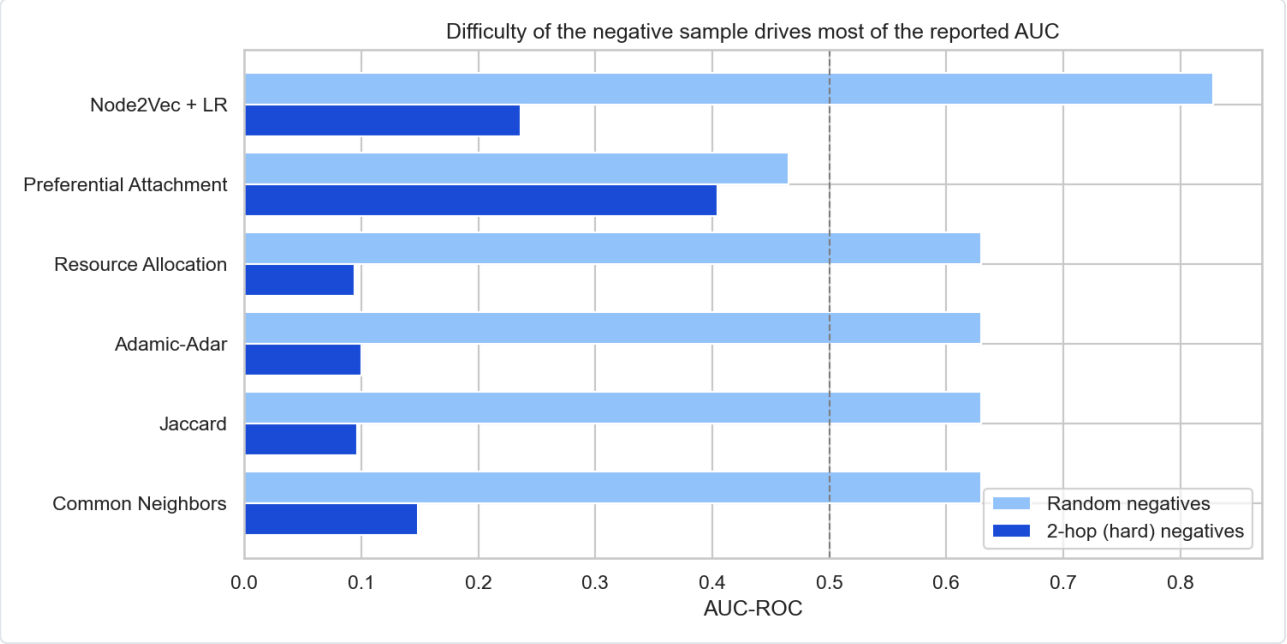



**איור 15.** דירוגי betweenness בשלושה פאנלים. משקול לפי זמן מול משקול לפי קפיצות (שמאל, אמצע) מניב $\rho = 0.61$ בלבד, לעומת רצפת-רעש של $\rho = 0.96$ (ימין, שתי ריצות קפיצות בזרעים שונים). הפער מוכיח שמשקול הזמן משנה ממשית את זהות התחנות המרכזיות.

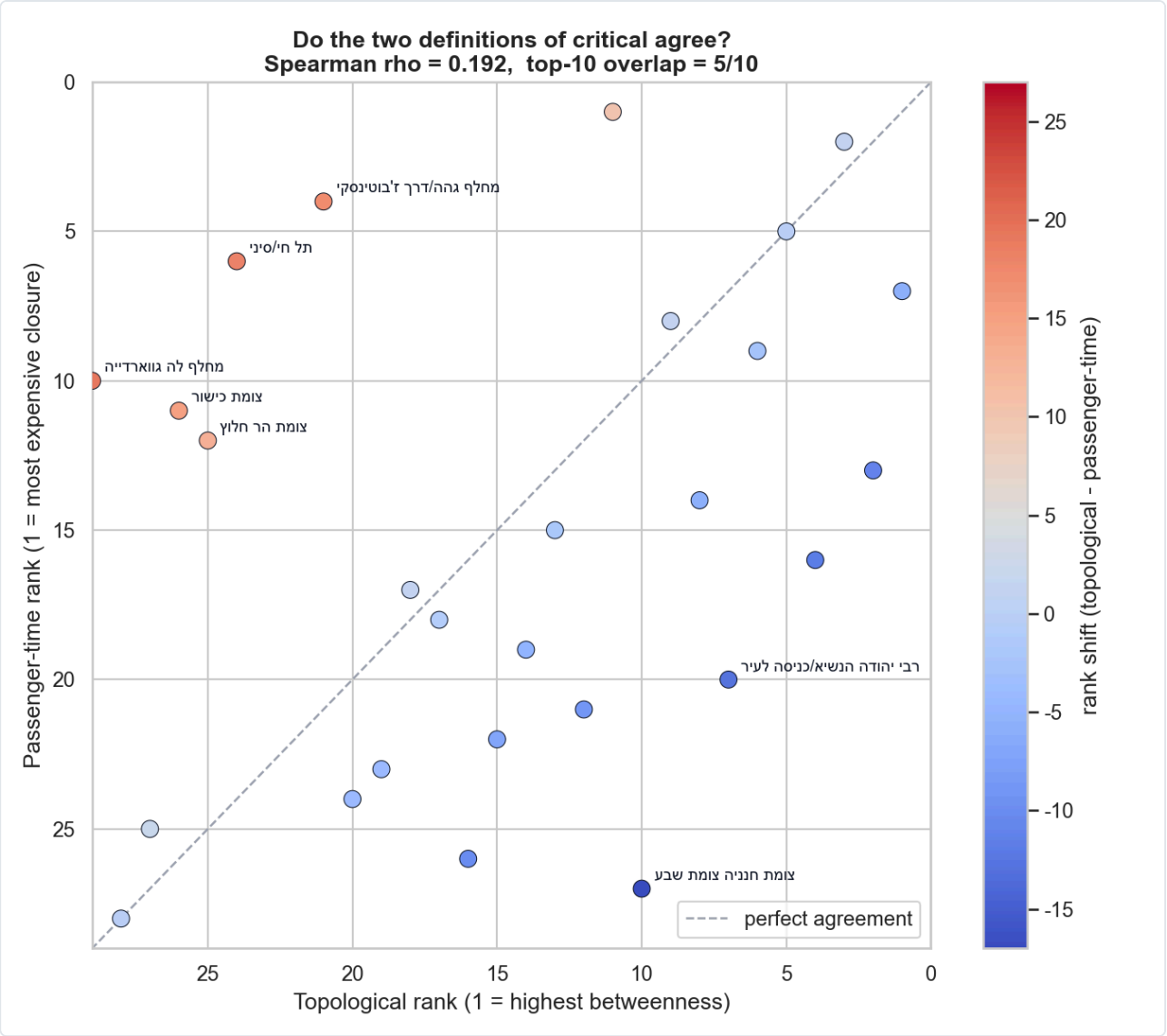
הרחבנו לשני כיוונים נוספים. **ניתוח דינמי:** חמישה גרפים לפי חלונות-שעה חשפו שהרשת מאבדת 14% מתחנותיה בלילה, ודירוגי הקריטיות זזים בין חלונות בחציון של 3,409 מקומות — תחנה קריטית בשעת שיא אינה בהכרח קריטית בלילה. **שקלול לפי ביקוש:** בהיעדר נתוני נסיעות בפועל ב-GTFS, בנינו פרוקסי מבוסס גרביטציה מאוכלוסיית הלמ"ס (8.7 מיליון נפש) — מסומן במפורש כפרוקסי — והוא משנה מהותית את הדירוג (מול בקרת משקל-אחיד ב-$\rho = 0.96$). **ניתוב-מחדש:** עבור 30 התחנות הקריטיות ביותר סימולנו סגירה וחישבנו מסלולים חלופיים על גרף הזמן. 51.5% מהנסיעות עוברות דרכן, אך הרשת סופגת את הסגירות היטב: עיכוב חציוני של 76 שניות בלבד, ורק 4 מתוך 6,595 זוגות מתנתקים. הממצא הלא-טריוויאלי: הדירוג לפי עלות-זמן-נוסע בפועל כמעט *אינו מתואם* עם הדירוג הטופולוגי ($\rho = 0.19$, לא-מובהק) — התחנה היקרה ביותר לסגירה מדורגת רק 11 ב-betweenness (איור 17).

## 10. למידת גרפים: תוצאה שלילית מבוקרת

יישמנו הטמעות Node2Vec לחיזוי קשרים ולסיווג תחנות קריטיות. בגרסה הראשונית קיבלנו AUC = 0.993 — אך האימון בוצע על הגרף המלא כולל קשתות-הבדיקה, כלומר **דליפת מידע**. לאחר תיקון הפרוטוקול צנח ה-AUC ל-0.83. המבחן החד ביותר הוא negatives "קשים" במרחק שתי קפיצות: תחתם *כל* השיטות, כולל Node2Vec, קורסות אל מתחת לקו האקראי 0.5 — Node2Vec מ-0.83 ל-0.24 (איור 16). המשמעות: על גרף גאוגרפי דליל ההטמעות מודדות את *קלות* ה-negatives, לא מבנה נלמד אמיתי. גם מסווג התחנות הקריטיות נכשל (F1 = 0.29, מתחת ל-0.35 של מאפיינים ידניים ובקושי מעל 0.22 של ניחוש-בסיס). אנו מדווחים זאת כתוצאה שלילית — וזו, בעבודה ביקורתית, מעלה ולא חיסרון.



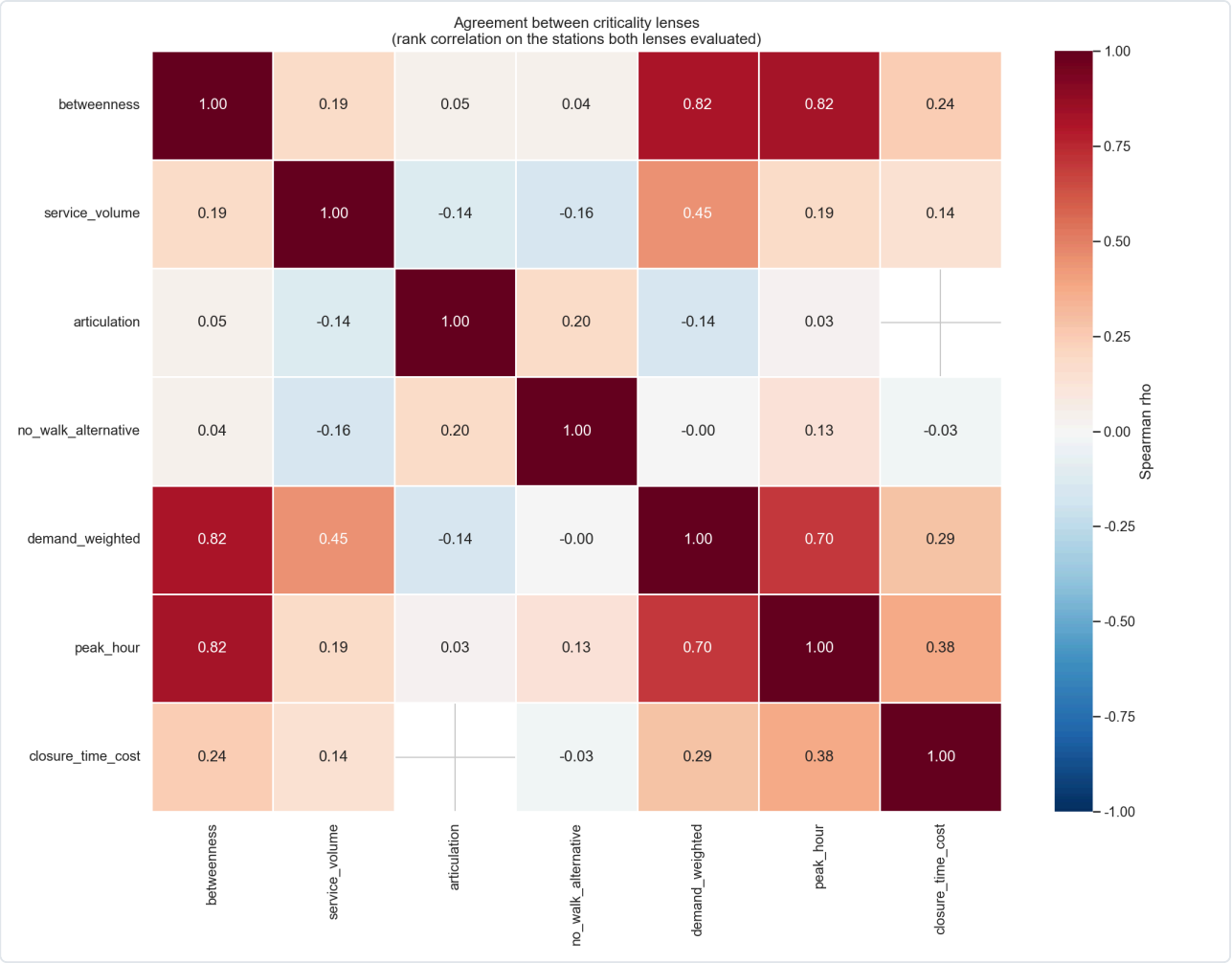

**איור 16.** AUC של חיזוי-קשרים תחת negatives אקראיים (תכלת) מול negatives קשים במרחק שתי קפיצות (כחול). כל השיטות קורסות אל מתחת לקו האקראי 0.5 — Node2Vec יורד מ-0.83 ל-0.24. ה-AUC ה"מרשים" נבע מקלות ה-negatives, לא ממבנה נלמד.




**איור 17.** דירוג טופולוגי (betweenness) מול דירוג לפי עלות-זמן-נוסע בסגירת תחנה, ל-28 מוקדים. הפיזור הרחב מהאלכסון מוכיח שמרכזיות טופולוגית אינה חוֹזֶה נאמנה לעלות הסגירה בפועל ($\rho = 0.19$, לא-מובהק).

## 11. סינתזה: שבע הגדרות של קריטיות

לסיום ריכזנו את כל ההגדרות. דירגנו את התחנות לפי שבע הגדרות קריטיות — betweenness (זרימה), נפח שירות, נקודת-חיתוך (מבנה), היעדר-חלופה (מרחב), משוקלל-ביקוש, שעת-שיא, ועלות-סגירה — ובחנו את ההסכמה ביניהן. ההסכמה הממוצעת על פני 21 זוגות ההגדרות היא $\rho = 0.21$ בלבד (איור 18). הגדרות השבירות המבנית (נקודת-חיתוך, היעדר-חלופה) מציגות מתאם אפסי עד שלילי מול הגדרת העומס (−0.16 עד −0.14), בעוד betweenness, ביקוש ושעת-שיא מתלכדות (עד 0.82). מתוך 231 התחנות שסומנו קריטיות ולו בהגדרה אחת, רק 26 מסכימות על שלוש הגדרות או יותר, ואף תחנה אינה מגיעה ל-Top-50 ביותר מחמש מתוך שבע ההגדרות.




**איור 18.** מטריצת קורלציית Spearman בין הדירוגים לפי שבע הגדרות של קריטיות. ההסכמה הממוצעת היא $\rho = 0.21$ בלבד; הגדרות השבירות המבנית מציגות מתאם אפסי-עד-שלילי מול הגדרת העומס, בעוד betweenness/ביקוש/שעת-שיא מתלכדות. זהות התחנות ה"קריטיות" תלויה באופן מכריע בהגדרה הנבחרת — הממצא המרכזי של העבודה.

## 12. מסקנות

1. **אין הגדרה יחידה ל"תחנה קריטית".** שבע הגדרות מדורגות מסכימות במתאם ממוצע של 0.21 בלבד, וקריטיות תפעולית ומבנית כמעט בלתי-תלויות ($\rho = 0.19$, מול רצפת-רעש 0.96). בחירת המדד קובעת את התשובה.
2. **הרשת היא עולם-קטן שאינו scale-free.** מודל-אפס משמר-דרגות מוכיח שההתקבצות היא חתימה מרחבית (פי 726), והזנב דק מכדי להיות scale-free — בניגוד לציפייה מרשת מבוססת-מרכזים.
3. **חסינה-אך-שבירה.** עמידה מאוד לכשל אקראי, פגיעה לתקיפה מכוונת; ופגיעות עצמה מתפצלת — הצפון מוביל בצוואריבקבוק, ירושלים במבנה עצי דל-יתירות.
4. **שוויוניות היא מקומית ותלוית-מדד.** פרדוקס Simpson מהפך את סימן הקשר הסוציו-אקונומי בין קהילות, ורק 8 מ-57 שורדות תיקון FDR — המתאם הארצי אינו מייצג אף קהילה.
5. **המבנה הארצי הוא רשת אוטובוס.** אך המודים הקטנים חסרי-יתירות (רכבת קלה: 100% גשרים), והמעבר הרב-אופני הוא עובדה מרחבית שמודל הנתונים מסתיר.
6. **קפיצות אינן זמן.** ניתוב לפי קפיצות תואם ניתוב לפי זמן ב-0.33% מהמסלולים — סייג מהותי על כל ניתוח טופולוגי, כולל זה שלנו.

מגבלות עיקריות: משקלים סופרים שירות מתוכנן ולא ביקוש בפועל; betweenness מקורב (רצפת-רעש 0.96); תחליפיות מרחבית בלבד; אין איחוד רציפים; תמונת-מצב של חודש אחד. הרחבות טבעיות — שילוב נתוני כרטוס אמיתיים ומודל זמן-המתנה מלא — מפורטות במחברת המסכמת (24).

## מקורות


1. Newman, M. E. J. (2010). *Networks: An Introduction*. Oxford University Press.
2. Easley, D., & Kleinberg, J. (2010). *Networks, Crowds, and Markets*. Cambridge University Press.
3. Brandes, U. (2001). A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology*, 25(2), 163–177.
4. Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech*.
5. Grover, A., & Leskovec, J. (2016). node2vec: Scalable Feature Learning for Networks. *KDD*.
6. Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate. *JRSS-B*, 57(1), 289–300.
7. MobilityData/ (2026). *General Transit Feed Specification Reference*. https://gtfs.org
8. הלשכה המרכזית לסטטיסטיקה (2021). *אפיון יחידות גאוגרפיות ודירוגן לפי הרמה החברתית-כלכלית*.