Cantwell & Newman

Message passing on networks with loops

By Hana Hasan and Hallel Weinberg

$Message\ Passing$ הוא אוסף של שיטות שמשמשות לחישוב ערכים או תכונות על צמתי הרשת ע"י העברת מידע בין השכנים ברשת.

הטכניקה מאפשרת לחלק חישובים גדולים ברשת לחלקים שאפשר לנהל ולפתור בצורה אנליטית או מספרית.

השיטה הובילה לתוצאות חשובות בתחומים רבים וביניהם : פיזיקה, סטטיסטיקה ומדעי המחשב.

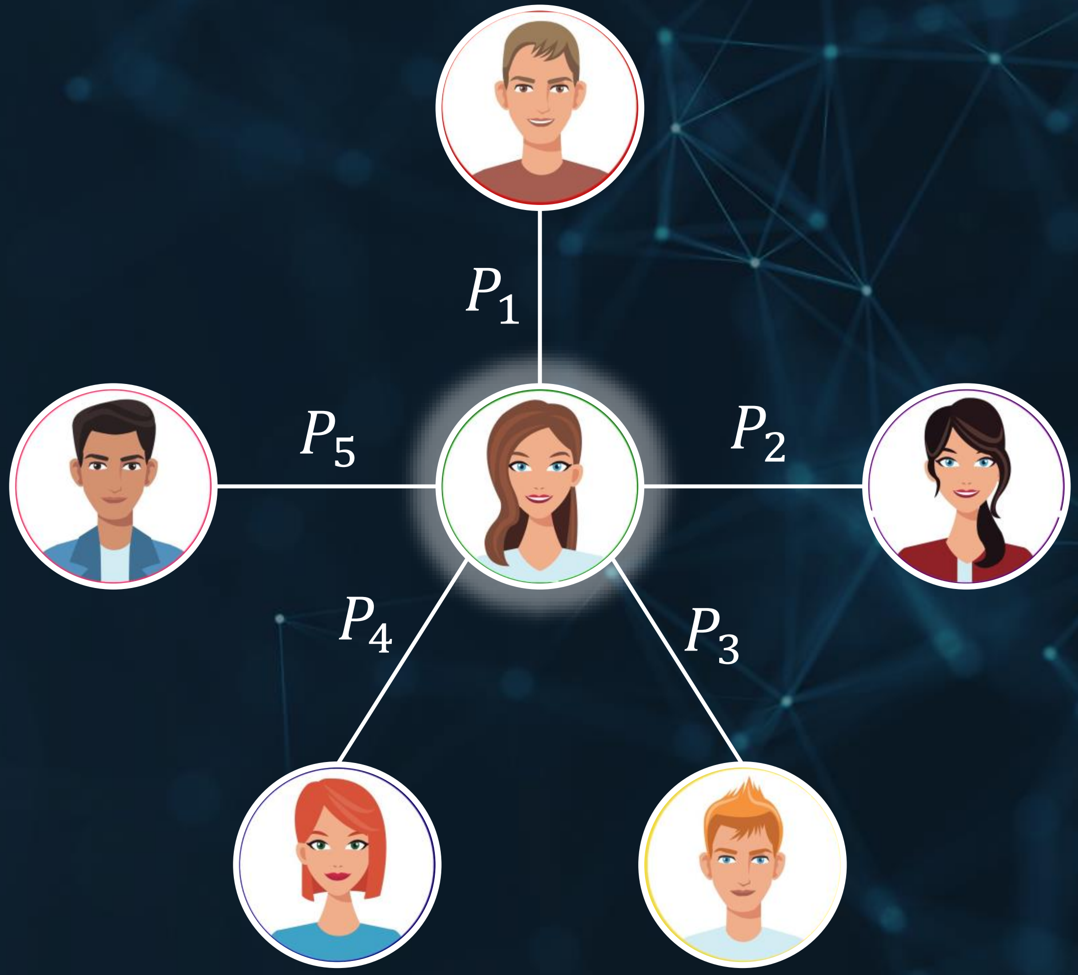
נסתכל על דוגמא של רשת התפשטות מגפה

נרצה לחשב את פוטנציאל ההידבקות של אליס ברשת מסוימת.

אליס

אליס נחשפת במהלך היום ל5 אנשים שונים

נסמן את ההסתברות שאליס נדבקה מהאדם ה-i כ-P_i:



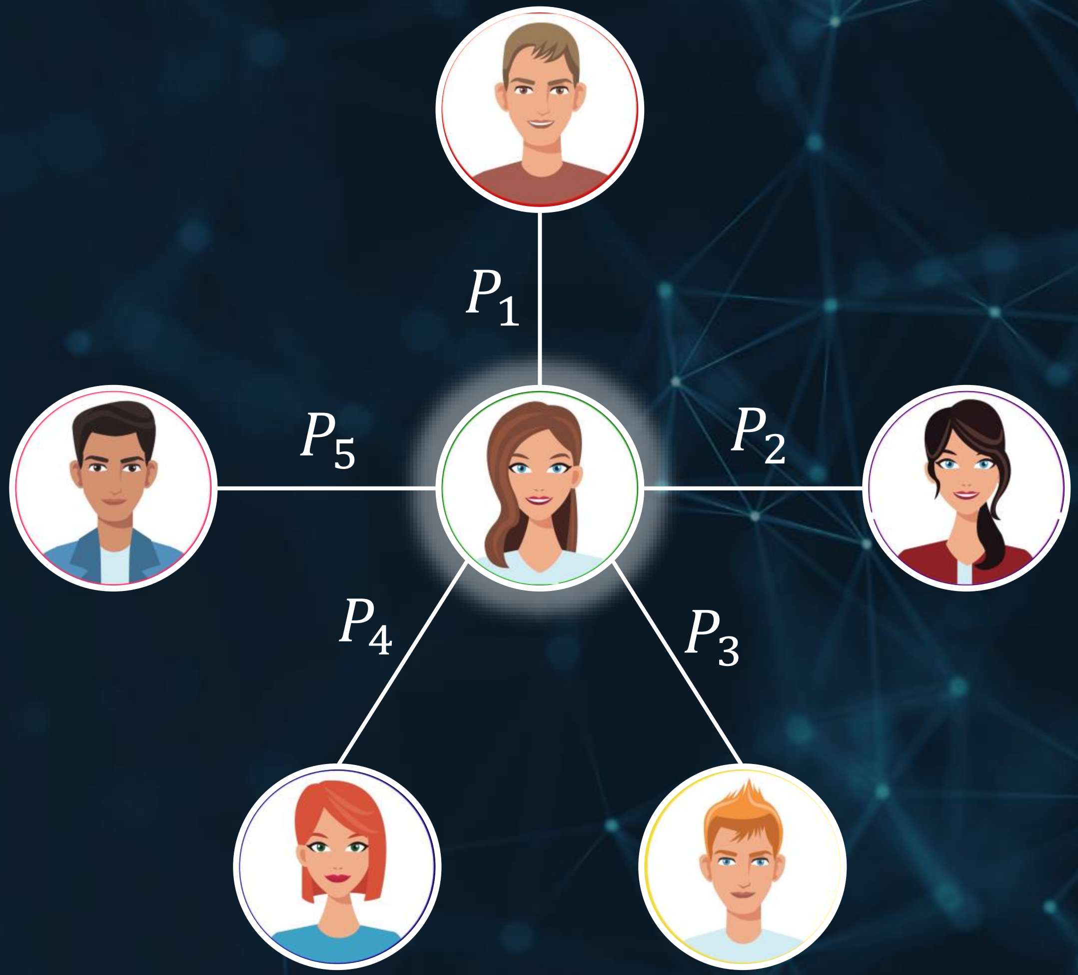
נסמן את ההסתברות שאליס נדבקה מהאדם הi כP_i:

$\boxed{P_i}$ =

ההסתברות שהאדם הi ידבק במחלה

·

ההסתברות שהאדם הi **הוא זה** שידביק את אליס במחלה



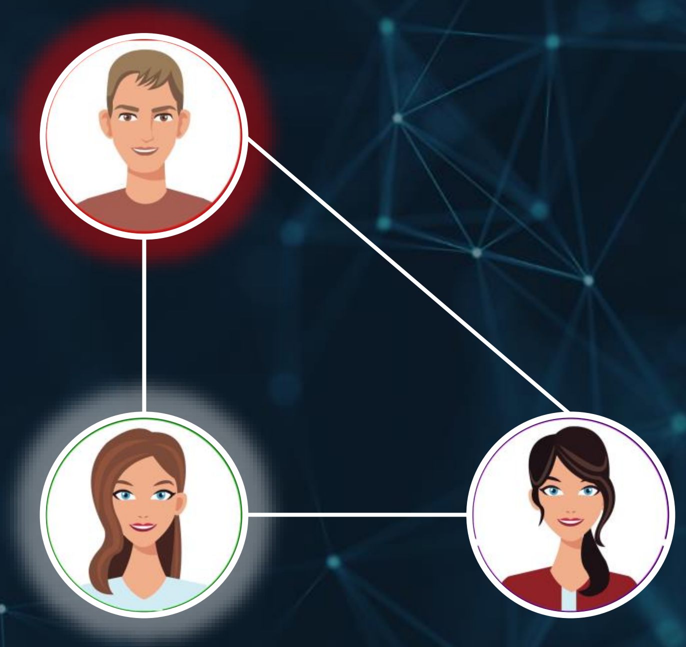
$$P(Alice\ will\ catch\ the\ disease) = P_1 + P_2 + P_3 + P_4 + P_5$$

נבחין כי בחישוב זה השתמשנו בהנחה סמויה, לפיה ההסתברויות $P_1, P_2, \dots, P_5$ הינן **בלתי תלויות**.

כעת, נסתכל על מצב בו הנחה זו אינה מתקיימת.

מה יקרה אם נוסיף את הקשת הבאה?

עכשיו, התנאי מהמצב הקודם לא מתקיים, יש תלות בין P_1 ל-P_2.



קיימת קורלציה:

הסיכוי של אליס לחלות עולה ביחס לפוטנציאל ההידבקות המקורי

קרול עלולה להידבק מבוב

בוב חולה

באופן כללי:

פוטנציאל ההידבקות של צומת עולה אם אחד מהשכנים חולה ומחובר לשכן אחר.

זוהי בעיה, מכיוון שבהינתן צומת x עם $1,2,\ldots,n$ שכנים עם הסתברויות $P_1,P_2,\ldots,P_n$ להדבקה, אין לנו דרך לחשב את ההסתברות שx יחלה.

נבחין כי הבעיה נוצרת כאשר ישנו מעגל ברשת:

בעולם האמיתי- רוב הרשתות הן ככה.

השיטה הנאיבית שתיארנו בהתחלה עובדת על:

01 **עצים**

02 **מעגלים ארוכים**

03 **רשתות עם אזורים של "עצים לוקאליים"**

השיטה שנציג במהלך ההרצאה תרחיב את הפתרון גם לרשתות עם מעגלים קצרים.

נציג מספר הגדרות שישמשו אותנו במהלך הפתרון:

נגדיר את הקבוצה $N_i^{(r)}$ הבאה:

$$N_i^{(r)} = i\text{'s edges and all length } 1, \dots, r \text{ paths between } i\text{'s neighbors}$$

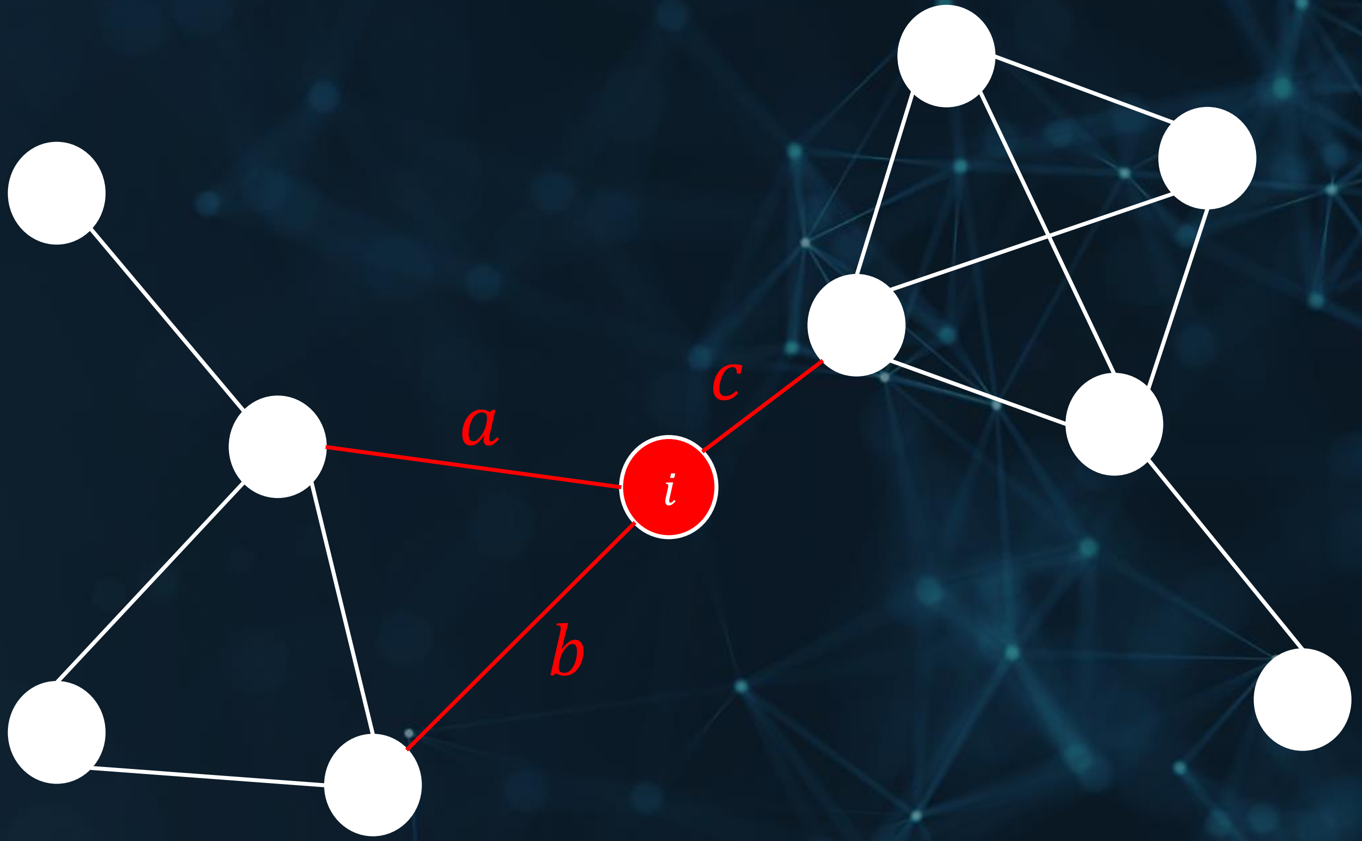
כלומר, הקבוצה $N_i^{(r)}$ תוגדר להיות כל הקשתות שמחוברות לצומת i יחד עם כל הקשתות שנמצאות על מסלולים בין שני שכנים שונים מאורך $\leq$ r.

הדגמה

נסתכל על הרשת הבאה :

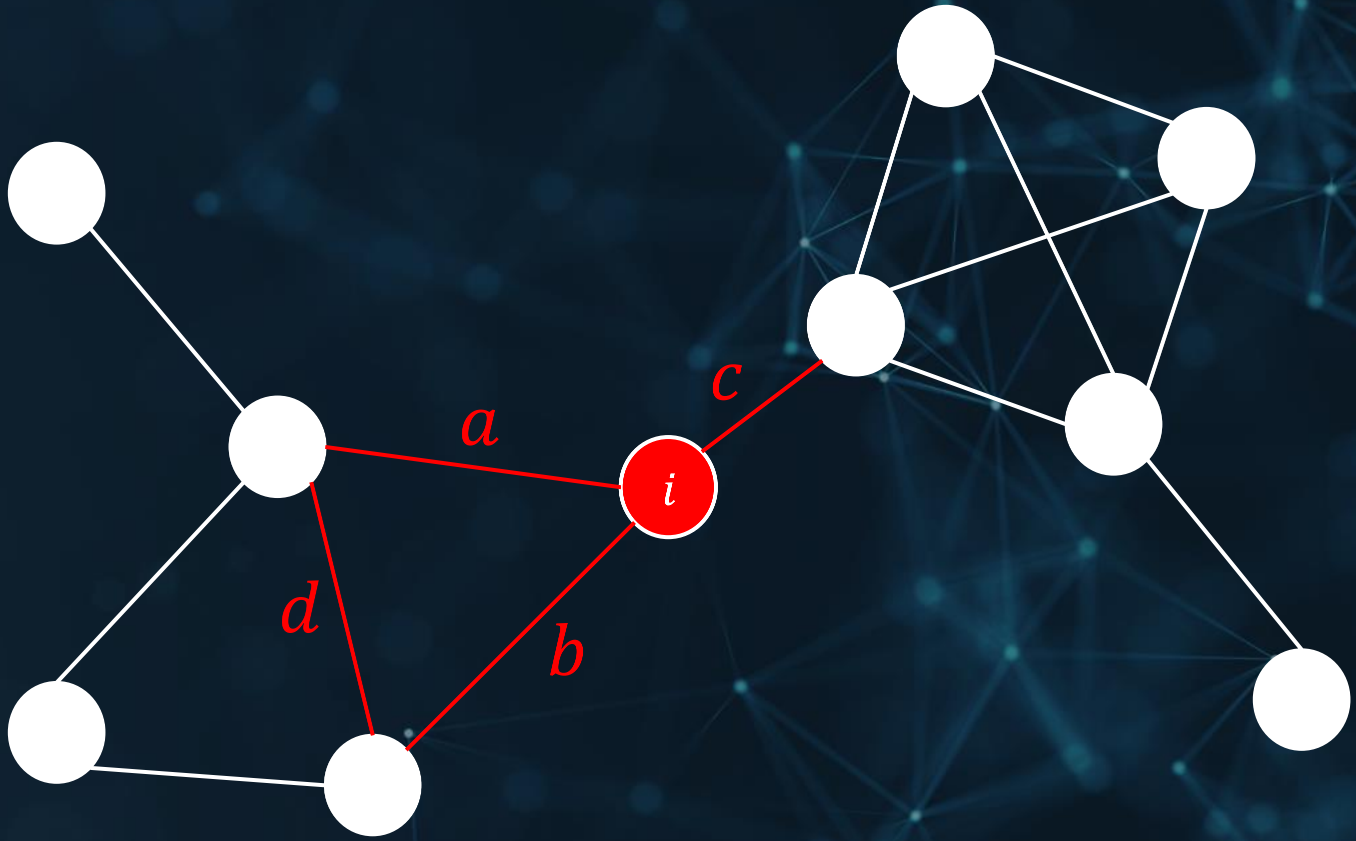

i

עבור $r = 0$:



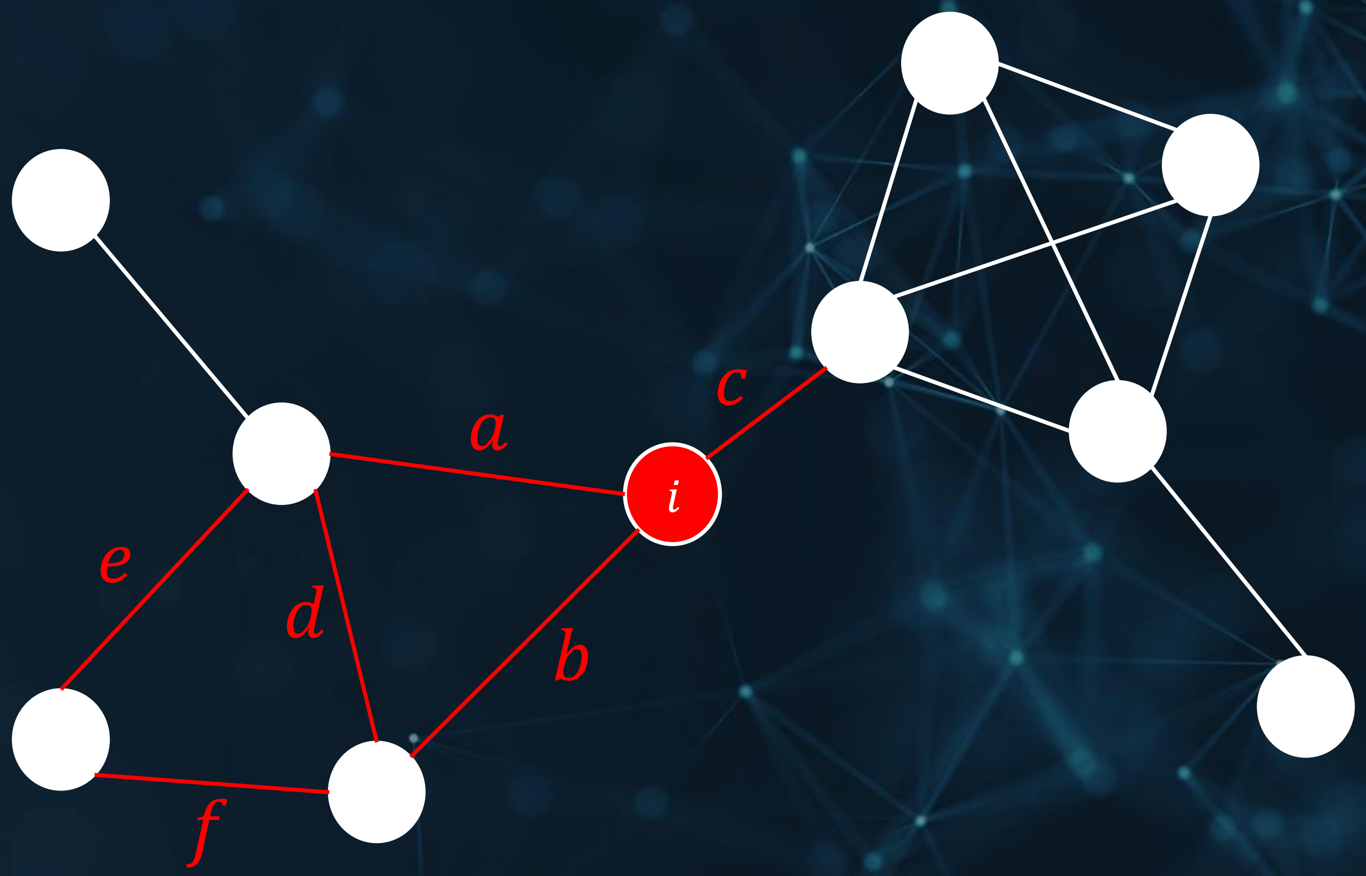
$$N_i^{(0)} = \{a, b, c\}$$

עבור $r = 1$:



$$N_i^{(1)} = \{a, b, c, d\}$$

עבור $r = 2$:



$$N_i^{(2)} = \{a, b, c, d, e, f\}$$

באופן רקורסיבי נוכל להגדיר את השכונה N כך:

$$N_i^{(r)} = N_i^{(r-1)} \cup \{paths\ of\ length\ r\ between\ neighbours\ of\ i\}$$

כאשר $N_i^{(0)} = i's\ edges$

הגדרה

מעגל בו יש לפחות צומת אחת שאינה הכניסה והיציאה של מעגל קצר יותר הוא **מעגל פרימיטיבי**.

לדוגמא:

מעגל פרימיטיבי

לא מעגל פרימיטיבי

שיטה חדשה

האלגוריתם מתבצע באופן איטרטיבי עד שהוא מגיע להתכנסות.

נגדיר רצף של r אומדנים כך שהאומדן ה-r-י יהיה תוצאה מדויקת על רשתות שמכילות מעגלים פרימיטיביים באורך מקסימלי של $r + 2$.

ברשתות שמכילות מעגלים פרימיטיביים ארוכים יותר, האומדן שיתקבל יכול להיות טוב אבל לא בהכרח מדויק.

bond percolation

ה*bond percolation* הוא תהליך מחיקת קשתות.

bond percolation

ה*bond percolation* הוא תהליך מחיקת קשתות.

bond percolation

ה*bond percolation* הוא תהליך מחיקת קשתות.

bond percolation

ה*bond percolation* הוא תהליך מחיקת קשתות.

נסמן בφ את ההסתברות שקשת היא *occupied* ברשת נתונה.

דוגמאות

נוצרים רכיבי קשירות קטנים מנותקים.

$\varphi = 0.2$

$\varphi = 0.5$

$\varphi = 1$

עבור φ הסתברות *occupation* קטנה, יש מעט קשתות *occupied*.

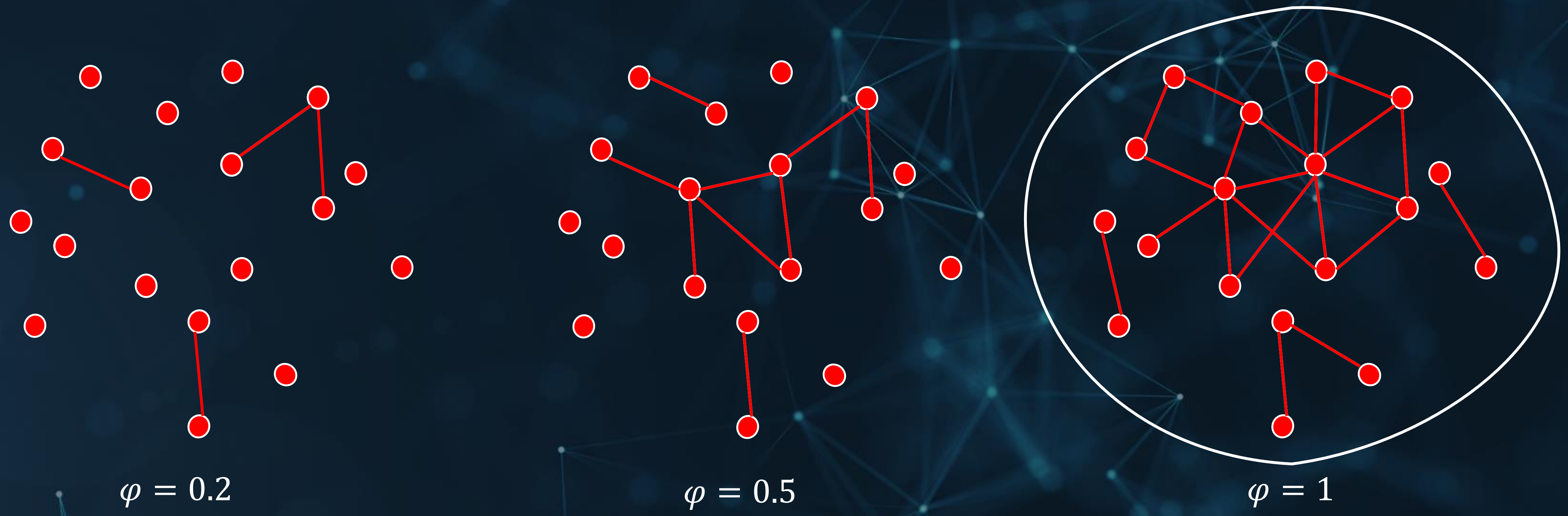
דוגמאות

נוצרים רכיבי קשירות קטנים מנותקים.

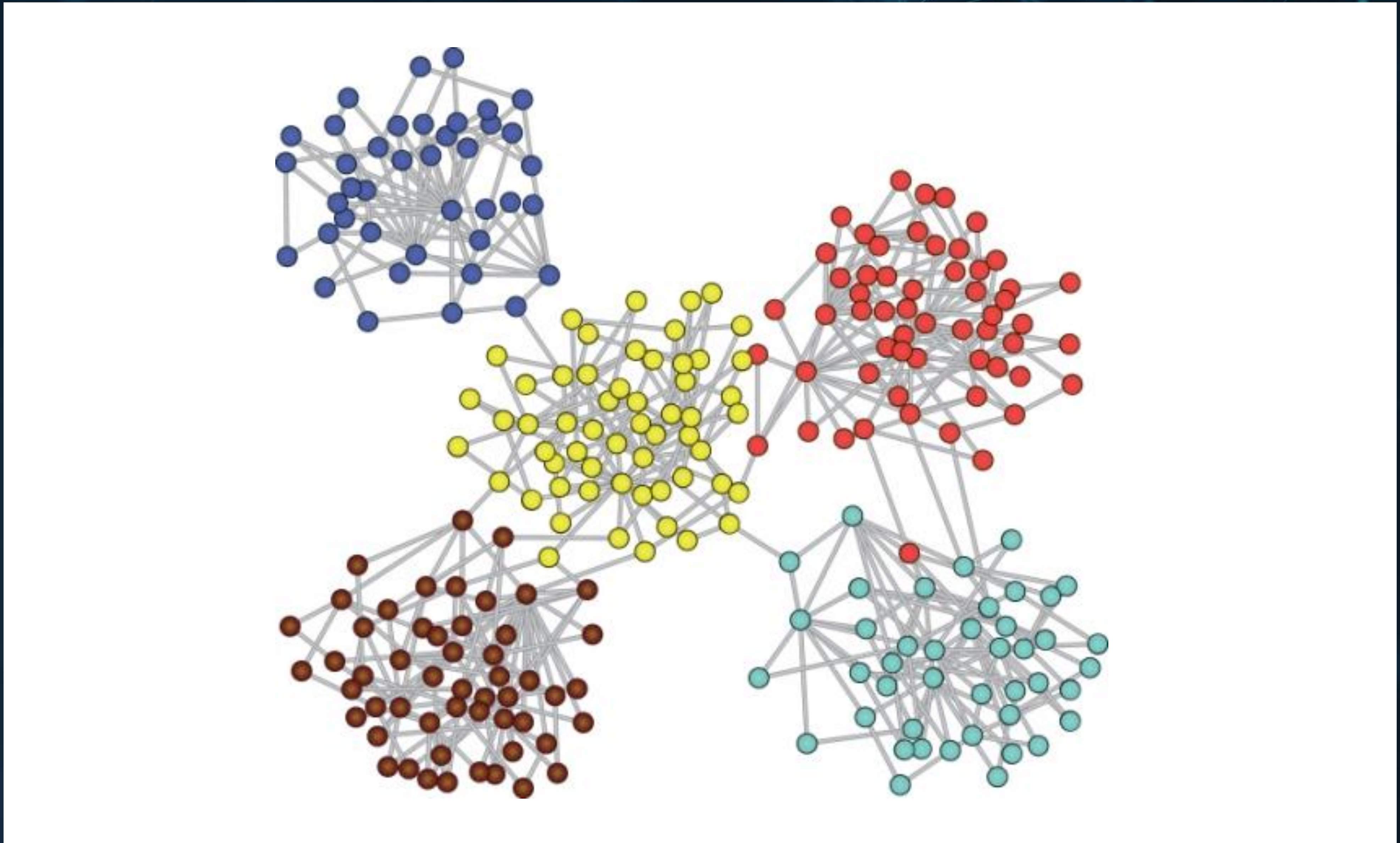
כדי להיות עקביים עם המאמר, נקרא לרכיבי הקשירות $cluster$-ים.

$\varphi = 0.2$ $\varphi = 0.5$ $\varphi = 1$

דוגמאות



ככל ש-φ גדל, מגיעה נקודה (***the percolation transition***) שבה רכיבי הקשירות המנותקים גדלים מספיק כדי להצטרף יחד וליצור *giant cluster*.



נסביר את האלגוריתם באמצעות אפליקציה

נסתכל על תהליך $bond\ percolation$ עבור רשת לא מכוונת בגודל n כאשר כל קשת היא $occupied$ בהסתברות p באופן בלתי תלוי.

נרצה לדעת מה ההתפלגות של גדלי ה$cluster$ים והאם קיים $giant\ cluster$.

תזכורת:

$$N_i^{(r)} = i\text{'s } edges\ and\ all\ length\ 1, \dots, r\ paths\ between\ i's\ neighbors$$

נגדיר כעת משתנה מקרי Γ_i

נסתכל על הרשת הבאה:

קו רציף-
קשתות
occupied

קו מקוקו-
קשתות
empty

קשתות אדומות-
קשתות השכונה
$N_i^{(r)}$ כאשר $r = 3$.

המשתנה המקרי Γ_i יהיה בדיוק קבוצת הצמתים מתוך $N_i^{(r)}$ שאפשר להגיע מ i אליהם ע״י $occupied\ edges$:

המטרה הראשונה היא לחשב את $\pi_i(s)$:

$\pi_i(s)$: ההסתברות שצומת i שייך ל-cluster בגודל s.

נעשה זאת בשני שלבים:

1. נחשב את ההסתברות המותנית $\pi_i(s|\Gamma_i)$:

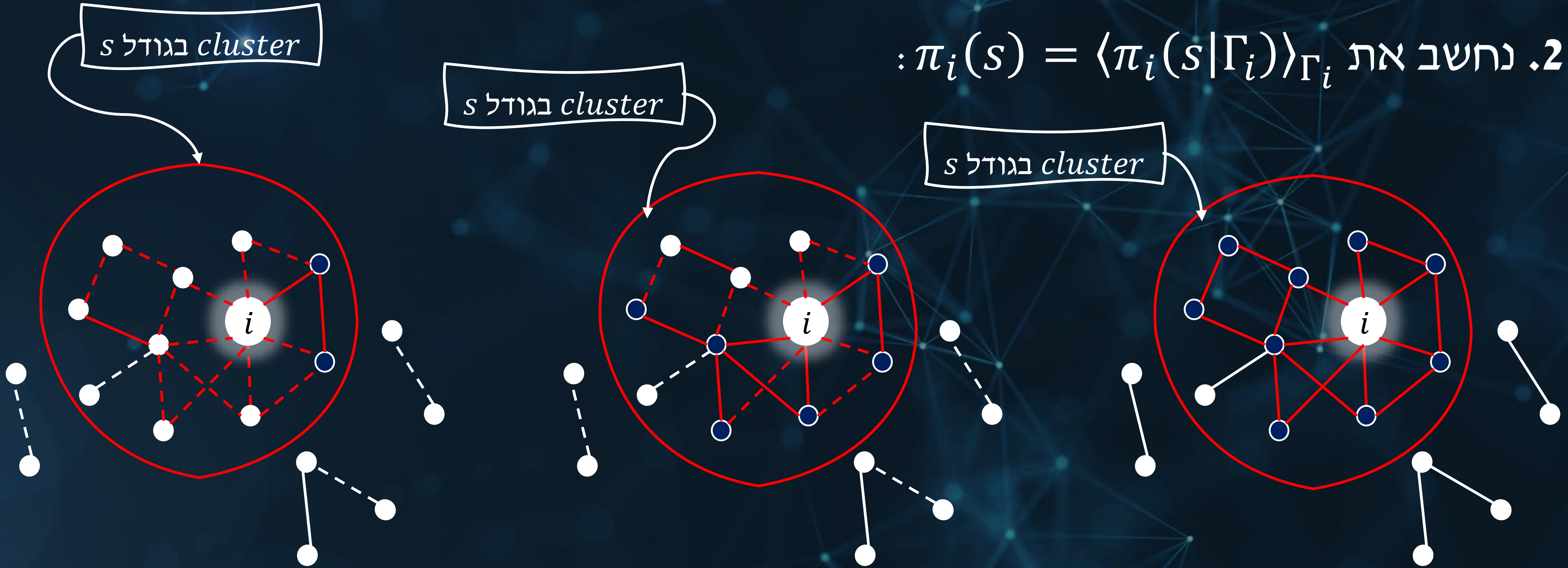
נעשה זאת בשני שלבים:

1. נחשב את ההסתברות המותנית $\pi_i(s|\Gamma_i)$:

$\boldsymbol{\pi_i(s|\Gamma_i)}$**:** ההסתברות שצומת i שייך ל$cluster$ בגודל s בהינתן קבוצת הצמתים הישיגים.

2. נחשב את $\pi_i(s) = \langle \pi_i(s|\Gamma_i) \rangle_{\Gamma_i}$:

2. נחשב את $\langle \pi_i(s|\Gamma_i) \rangle_{\Gamma_i} = \pi_i(s)$:



$\boldsymbol{\pi_i(s)}$**:** הממוצע המשוקלל של $\pi_i(s|\Gamma_i)$ מעל Γ_i-ים שונים (בכל שלב של תהליך ה*bond percolation*).

ההסתברות של כל קונפיגורציית קשתות הינה:

$$p^k(1-p)^{m-k}$$

p- ההסתברות שקשת היא $occupied$.

k- מספר הקשתות ה$occupied$.

m- גודל השכונה $N_i^{(r)}$.

נניח שצומת i שייך ל$cluster$ בגודל s והרשת לא מכילה מעגלים פרימיטיביים באורך גדול מ$r+2$

אז אם נמחק את כל קשתות השכונה $N_i^{(r)}$ צמתי Γ_i יהפכו להיות מנותקים אחד מהשני.

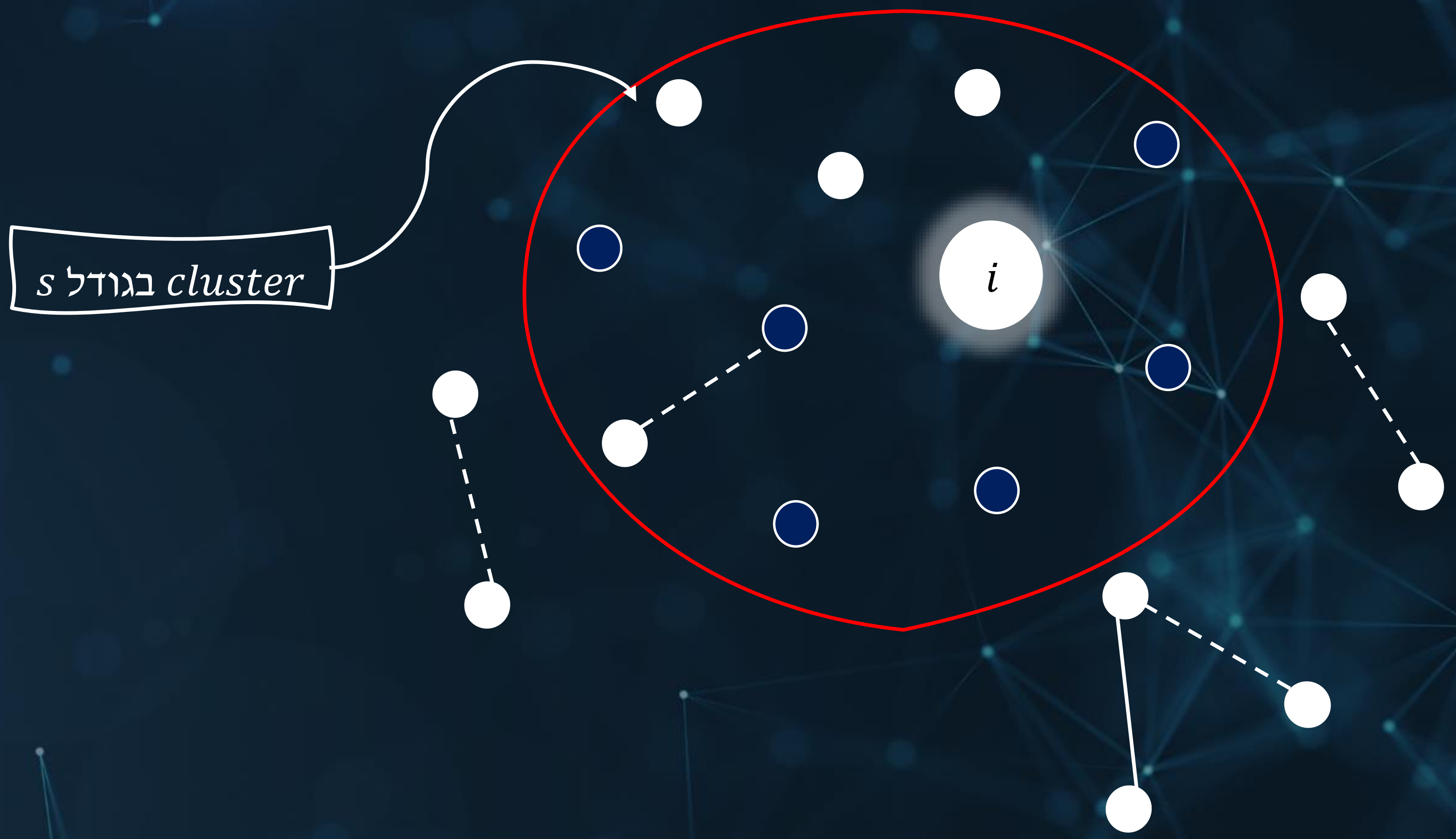
כעת, נמחק את קשתות השכונה.

נניח שצומת i שייך ל$cluster$ בגודל s והרשת לא מכילה מעגלים פרימיטיביים באורך גדול מ$r+2$

אז אם נמחק את כל קשתות השכונה $N_i^{(r)}$ צמתי Γ_i יהפכו להיות מנותקים אחד מהשני.

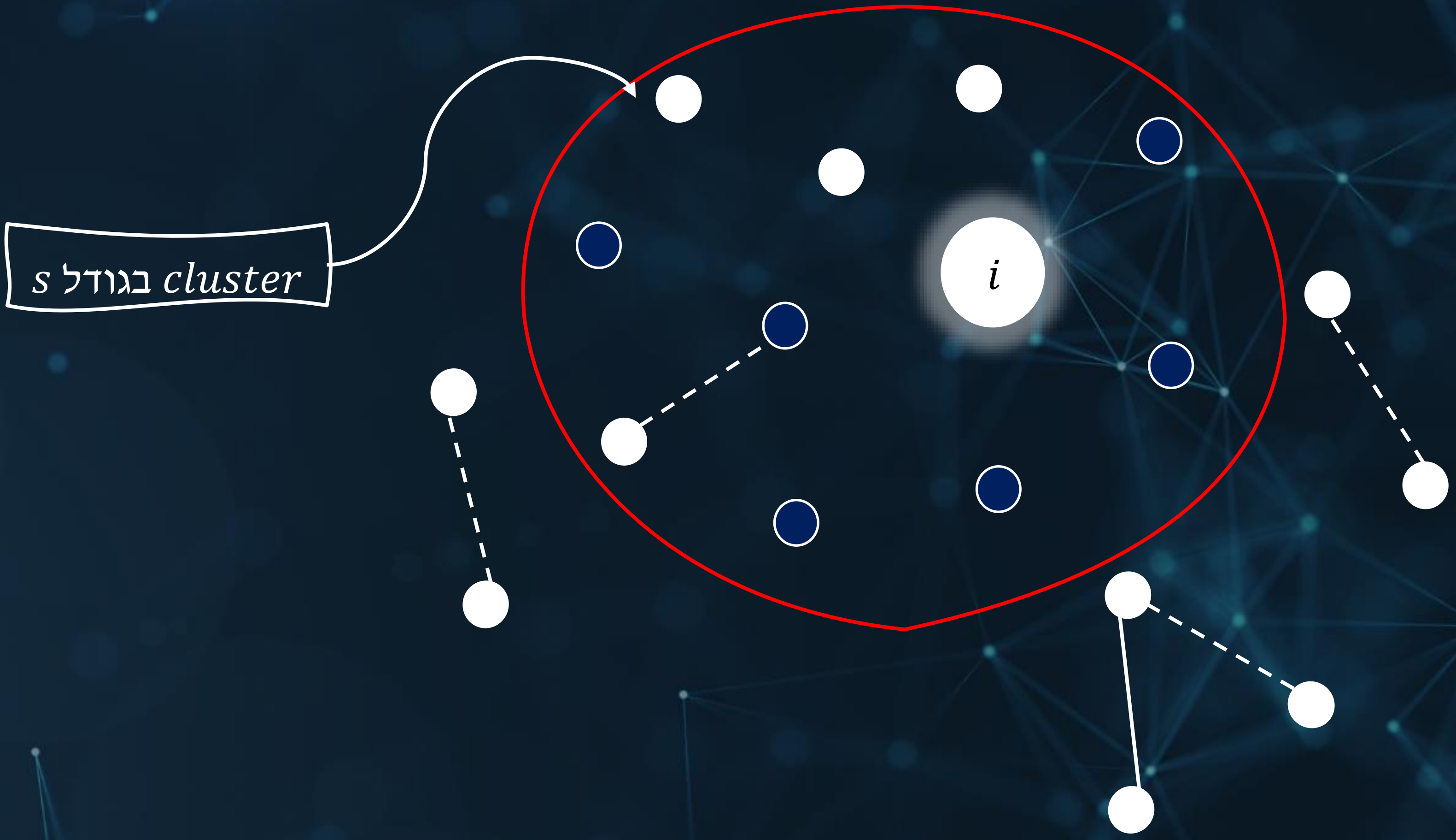
אחרי מחיקת קשתות השכונה, צמתי Γ_i (הצמתים הכחולים) מנותקים אחד מהשני.

הוכחה



נניח בשלילה כי מחיקת קשתות השכונה לא מנתקת את צמתי Γ_i זה מזה.

הוכחה

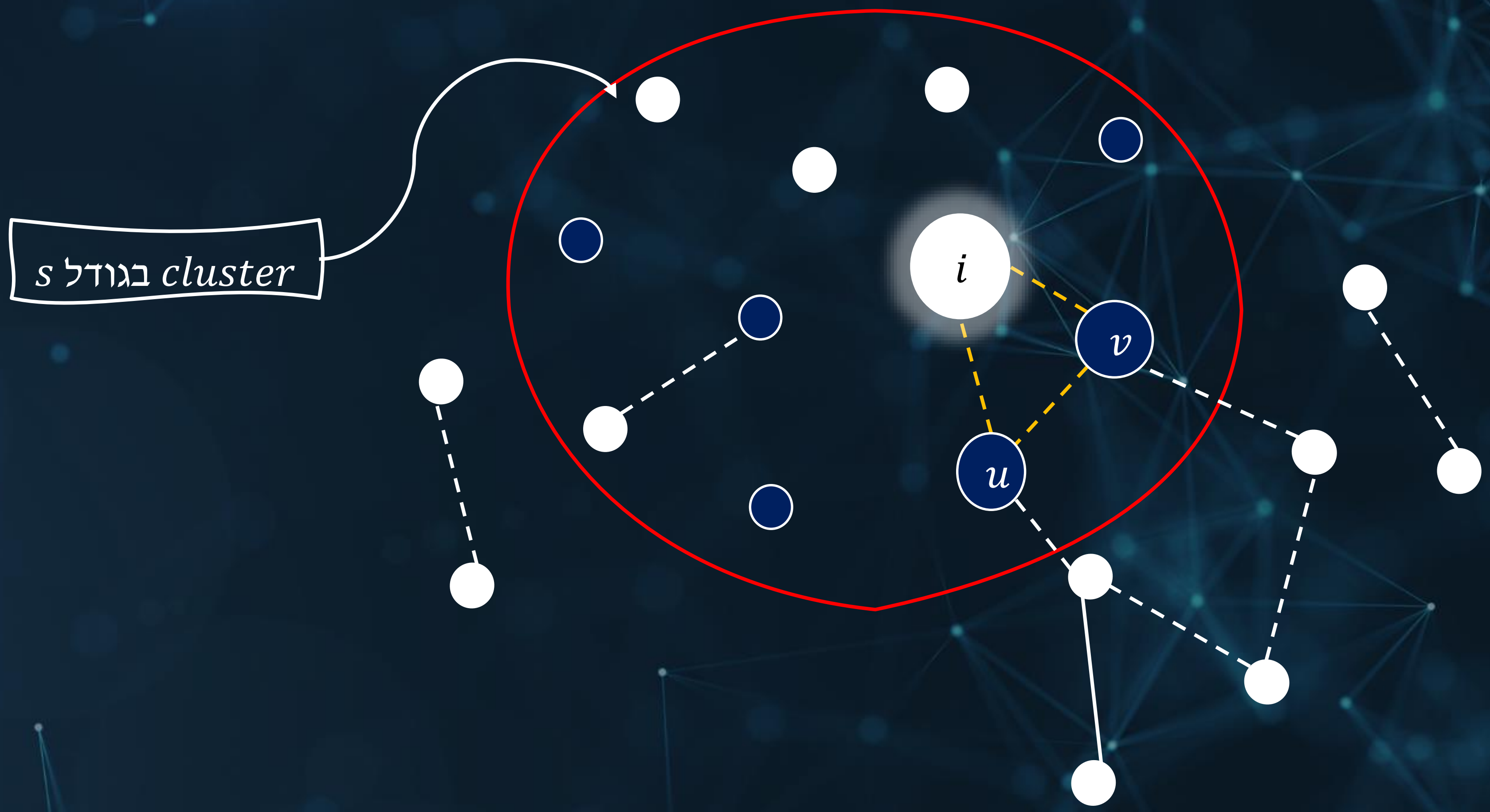


אבל, מחיקת קשתות השכונה מוחקת כל מסלול "פנימי" בין צמתי Γ_i.

הוכחה

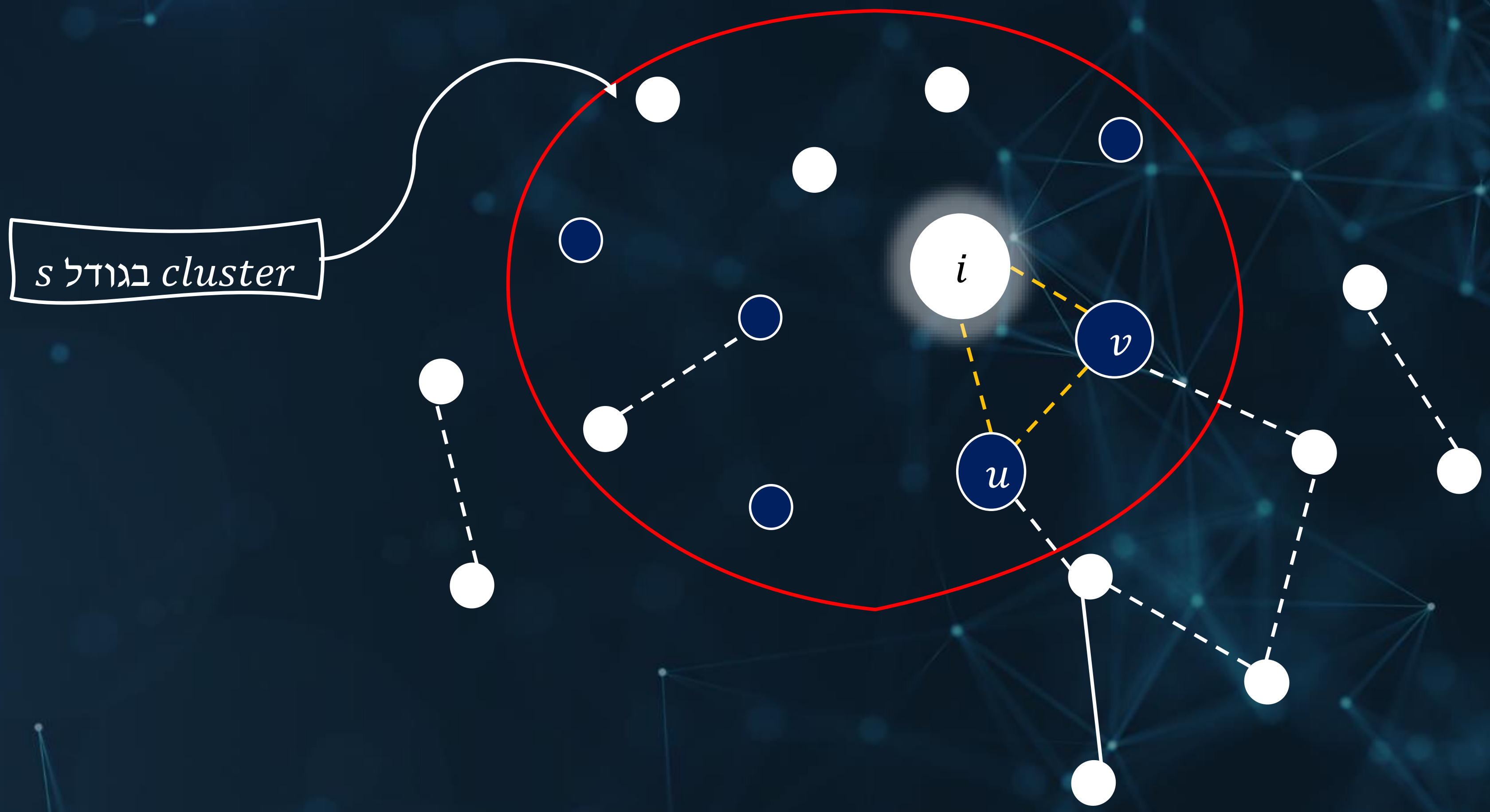
לכן, נניח ששני צמתים $u,v \in \Gamma_i$ מחוברים ע״י מסלול ״חיצוני״.

הוכחה



נבחין כי לפי הגדרה, השכונה $N_i^{(r)}$ מגדירה מעגל פרימיטיבי מקסימלי באורך $r+2$.

הוכחה



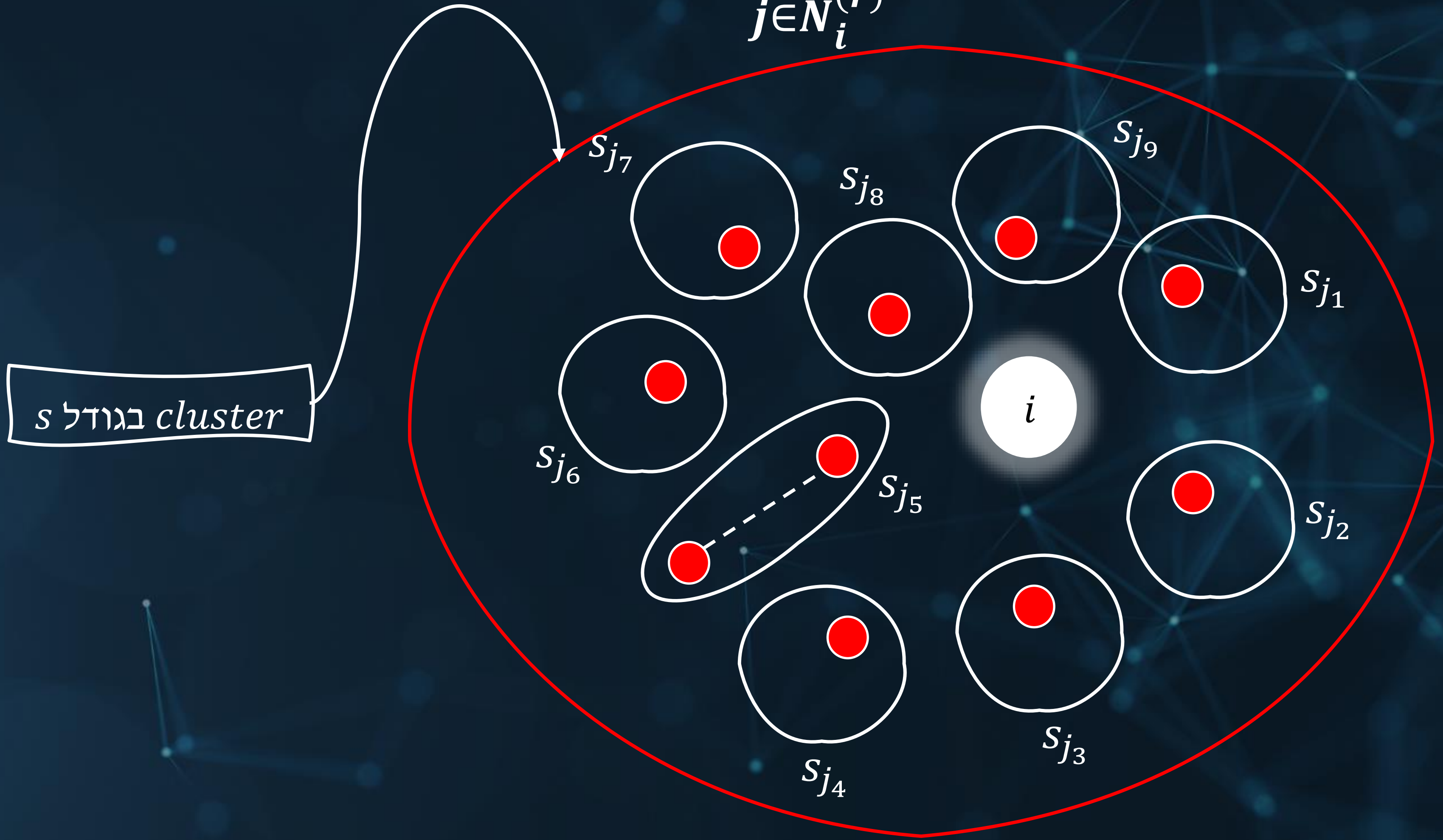
אבל, במצב כזה, היינו מקבלים ברשת מעגל פרימיטיבי ארוך מ$2r+$, בסתירה להנחה ולכן הנחת השלילה אינה מתקיימת.

יהי צומת j שעבורו קיים צומת u ברשת (לפני מחיקת הקשתות) כך ש $(u,j) \in N_i^{(r)}$. נגדיר את s_j

להיות גודל ה$cluster$ שj נמצא בו אחרי מחיקת קשתות $N_i^{(r)}$.

אזי:

$$\sum_{j\in N_i^{(r)}} s_j = s - 1$$

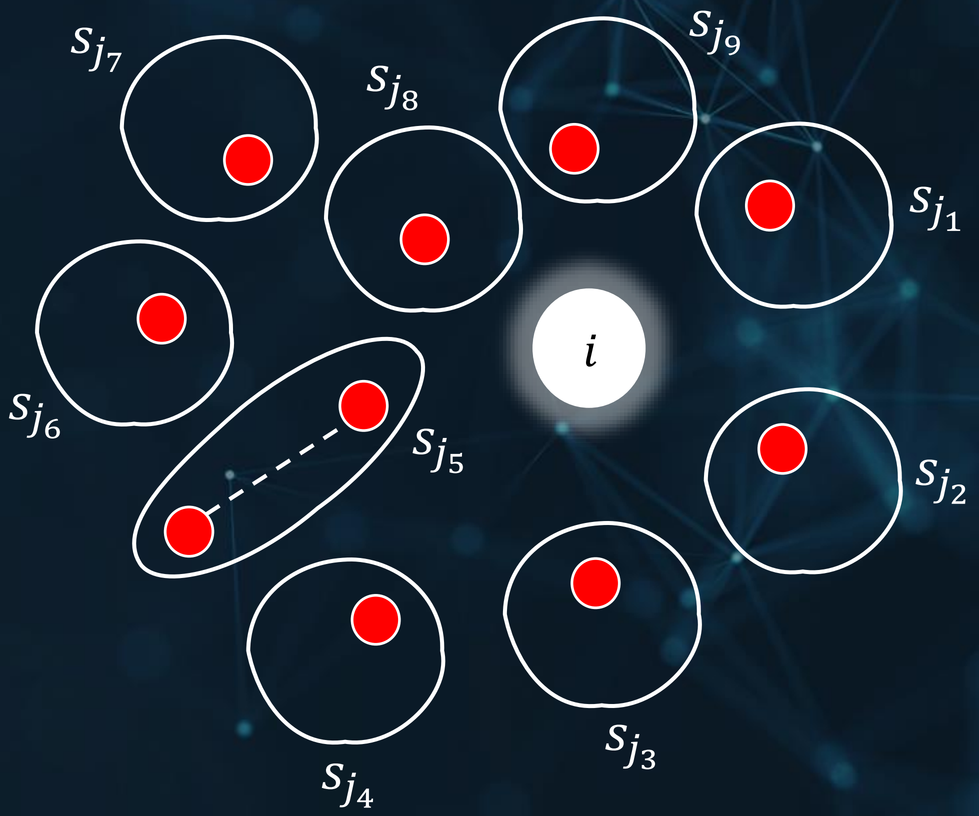


חישוב $\pi_i(s|\Gamma_i)$

נסתכל על רשת אחרי מחיקת השכונה של i :

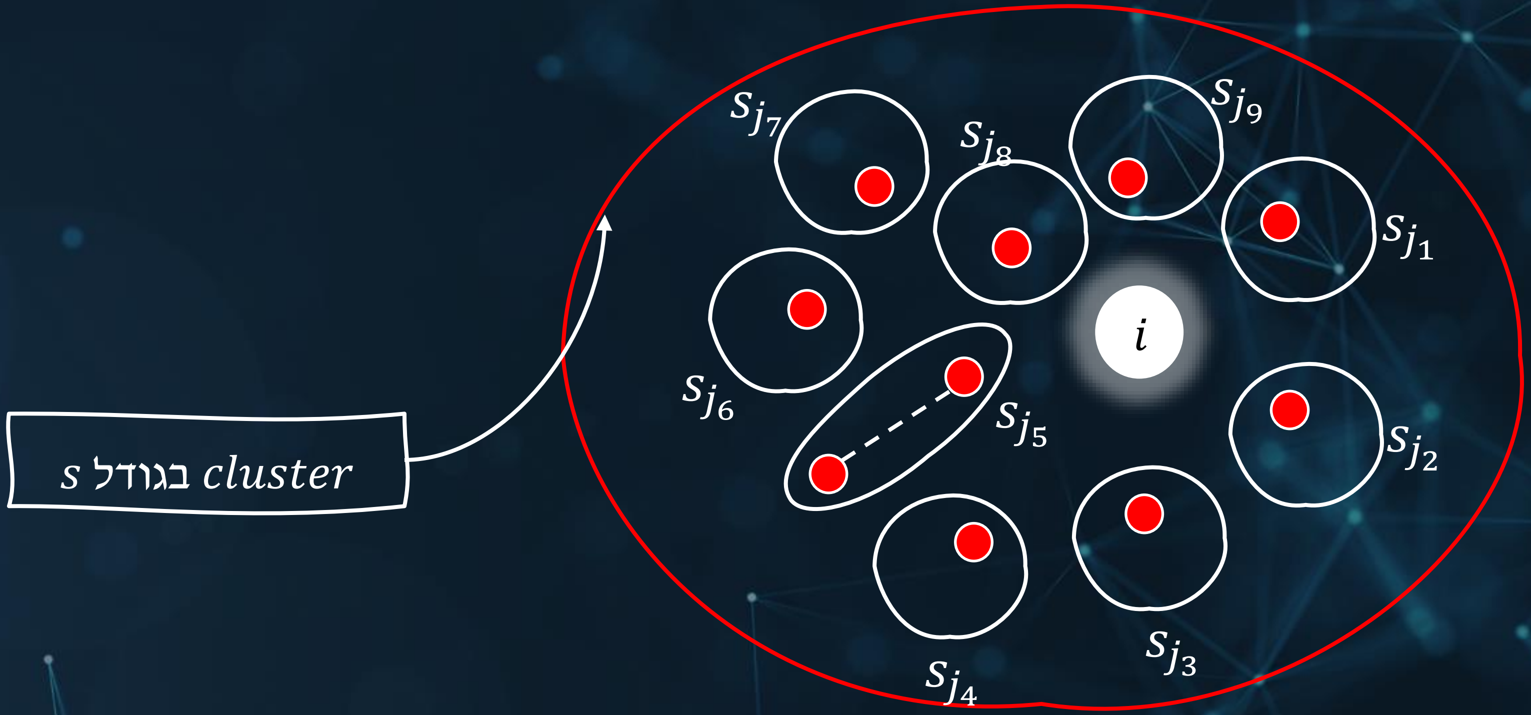
חישוב $\pi_i(s|\Gamma_i)$

נסתכל על גדלי ה- *clusters* שנוצרו בשכונה :



חישוב $\pi_i(s|\Gamma_i)$

על מנת לחשב את ההסתברות שצומת i שייך ל$cluster$ בגודל s לפני המחיקה, נרצה לחשב את ההסתברות שאחרי המחיקה נוצר $cluster$ בגודל 1 ואת ההסתברות שנוצר $cluster$ בגודל 2 וכך הלאה...



$\boldsymbol{\pi_{i\leftarrow j}(s)}$: ההסתברות שצומת j נמצא ב$cluster$ בגודל s אחרי שמוחקים את קשתות השכונה $N_i^{(r)}$.

עבור גודל $cluster$: s', ההסתברות שאחרי המחיקה נוצר $cluster$ בגודל s' שווה ל: $\prod_{j\in\Gamma_i}\pi_{i\leftarrow j}(s')$

נוסחה לחישוב $\pi_i(s|\Gamma_i)$

נחשב את ההסתברות שאחרי המחיקה נוצר
cluster בגודל 1 ואת ההסתברות שנוצר
cluster בגודל 2 וכך הלאה... ונחשב את סכומן

$i \in cluster\ of\ size\ S \Rightarrow \sum_{j\in\Gamma_i} s_j = s-1$

$$\boldsymbol{\pi_i(s|\Gamma_i)} = \begin{cases} \displaystyle\sum_{\{s_j:j\in\Gamma_i\}} \left[\prod_{j\in\Gamma_i} \pi_{i\leftarrow j}(s_j)\right] & \displaystyle\sum_{j\in\Gamma_i} s_j = s-1 \\ 0 & else \end{cases}$$

$\boldsymbol{\pi_{i\leftarrow j}(s)}$: ההסתברות שצומת j נמצא ב*cluster* בגודל s אחרי שמוחקים את קשתות השכונה $N_i^{(r)}$.

נגדיר פונקציה יוצרת עבור גדלי ה$clusters$ $H_i(z)$

נגדיר את המשתנה המקרי S להיות גדלי ה- $clusters$,

נרצה למצוא ל- S פונקציה יוצרת-הסתברות $H_i(z)$:

$$H_i(z) = \sum_s \pi_i(s) z^s$$

נגדיר פונקציה יוצרת עבור גדלי ה$clusters$ $H_i(z)$

מוטיבציה:

לחשב את הגודל
הממוצע של
$cluster$ שצומת
i שייך אליו:
$\langle s_i \rangle = H_i'(1)$.

לחשב את ההסתברות
שצומת i שייך
ל$cluster$ כלשהו:

$$\sum_s \pi_i(s) = H_i(1)$$

חישוב $H_i(z|\Gamma_\mathrm{i})$

כמו שהגדרנו את $\pi_i(s)$ להיות $\langle \pi_i(s|\Gamma_i)\rangle_{\Gamma_i}$, נגדיר את $H_i(z)$ ולכן נגדיר קודם את $H_i(z|\Gamma_i)$:

($\boldsymbol{H_i(z|\Gamma_\mathrm{i})}$ היא פונקציה יוצרת-הסתברות עבור גדלי ה- $clusters$ בשכונה תחת קונפיגורציה ספציפית)

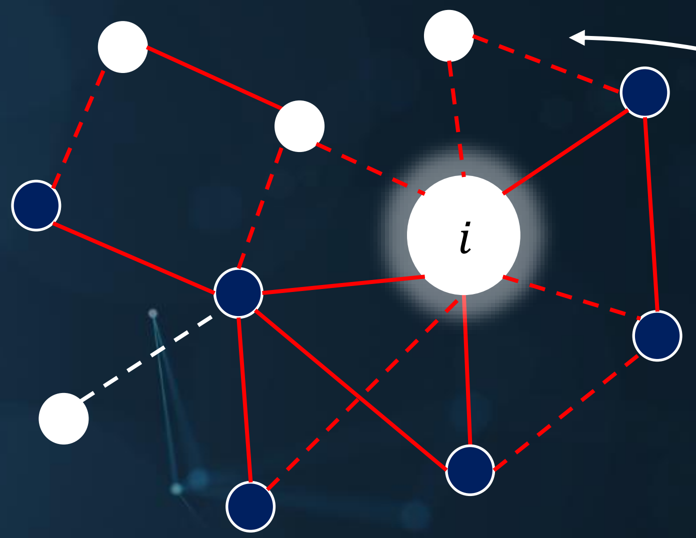
$$\boldsymbol{H_i(z|\Gamma_i)} = \sum_{s=1}^{\infty} \pi_i(s|\Gamma_i) z^s = \cdots = \boldsymbol{z \prod_{j\in\Gamma_i} \sum_{s_j=1}^{\infty} z^{s_j} \pi_{i\leftarrow j}(s_j)}$$

חישוב $H_i(z)$ על ידי $H_i(z|\Gamma_i)$

הגדרות:

- עבור צומת i וצומת $j \in N_i^{(r)}$ נגדיר אינדיקטור w_{ij}:

$$w_{ij} = \begin{cases} 1 & j \in \Gamma_i \\ 0 & else \end{cases}$$



בשכונה הזאת, למשל, הצמתים הלבנים
מקבלים את הערך 0 והצמתים הכחולים
מקבלים את הערך 1

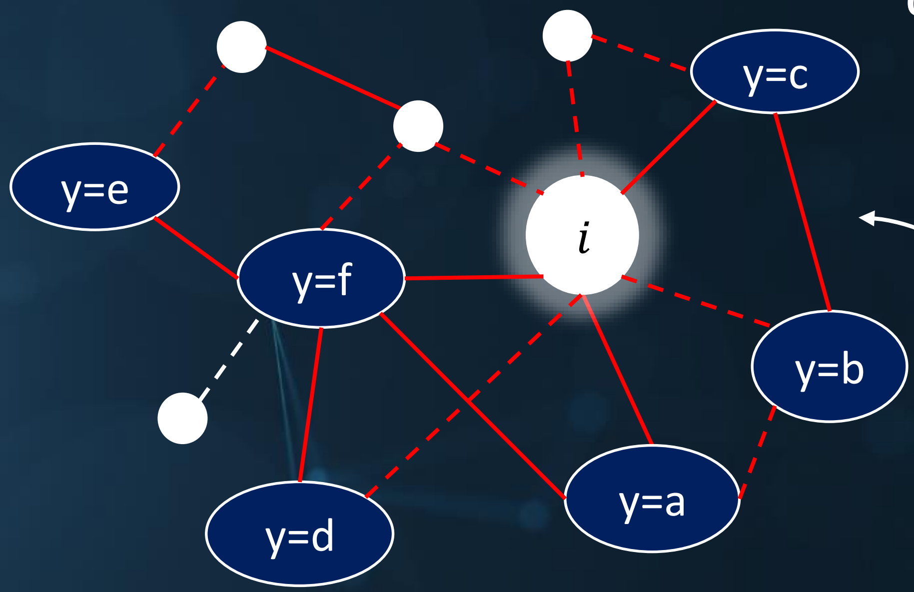
חישוב $H_i(z)$ על ידי $H_i(z|\Gamma_i)$

הגדרות:

- עבור עבור צומת i ומשתנה y נגדיר את הערך $G_i(y)$ להיות:

$$G_i(y) = \left\langle \prod_{j \in N_i^{(r)}} y_j^{w_{ij}} \right\rangle_{\Gamma_i}$$

במילים אחרות, לכל Γ_i: $G_i(y) = \prod_{j \in \Gamma_i} y_j$.



בשכונה הזאת, ועבור קונפיגורצית הקשתות הבאה:

$$G_i(y) = abcdef$$

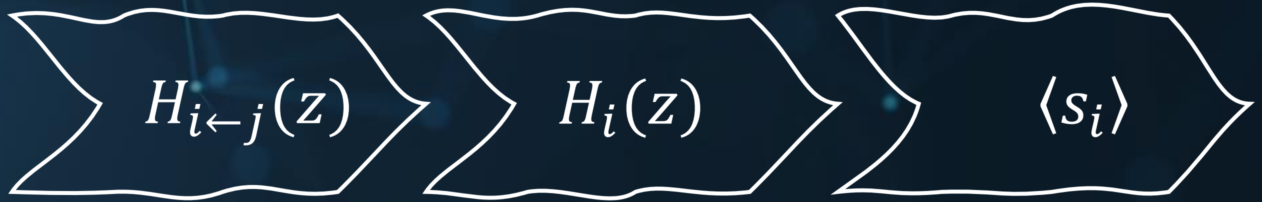
חישוב $H_i(z)$ על ידי $H_i(z|\Gamma_i)$

$$H_i(z) = \sum_{s=1}^{\infty} \pi_i(s) z^s = zG_i\big(H_{i\leftarrow}(z)\big)$$

נפתח את הביטוי $G_i\big(H_{i\leftarrow}(z)\big)$:

$$G_i\big(H_{i\leftarrow}(z)\big) = \left\langle \prod_{j\in\Gamma_i} H_{i\leftarrow j}(z) \right\rangle_{\Gamma_i}$$

כעת, נותר לנו לחשב את $H_{i\leftarrow j}(z)$: (פונקציה יוצרת-הסתברות עבור משתנה מקרי המקבל את ערכי גדלי ה- $clusters$ בשכונה תחת קונפיגורציה ספציפית)



חישוב $H_{i\leftarrow j}(z)$

כמו שחישבנו את $H_i(z)$ על ידי $H_\mathrm{i}(z|\Gamma_\mathrm{i})$, נחשב את $H_{i\leftarrow j}(z)$ על ידי $\boldsymbol{H_{i\leftarrow j}(z|\Gamma_{j\setminus i})}$ (פונקציה יוצרת- הסתברות עבור גדלי ה- $clusters$ אליהם שייך j בשכונה שלו אחרי שמוחקים את השכונה של i תחת קונפיגורציה ספציפית). החישוב זהה לחישוב $H_i(z)$:

$$\boldsymbol{H_{i\leftarrow j}(z|\Gamma_{j\setminus i}) = z \prod_{k\in N_{j\setminus i}^{(r)}} \left[H_{j\leftarrow k}(z)\right]^{w_{j\setminus i,k}}}$$

$$\boldsymbol{H_{i\leftarrow j}(z) = zG_{i\leftarrow j}\left(H_{j\leftarrow}(z)\right)}$$

חישוב $H_{i\leftarrow j}(z)$

$$H_{i\leftarrow j}\left(z\middle|\Gamma_{j\setminus i}\right) = z\prod_{k\in N_{j\setminus i}^{(r)}}\left[H_{j\leftarrow k}(z)\right]^{w_{j\setminus i,k}}$$

$$H_{i\leftarrow j}(z) = zG_{i\leftarrow j}\left(H_{j\leftarrow}(z)\right)$$

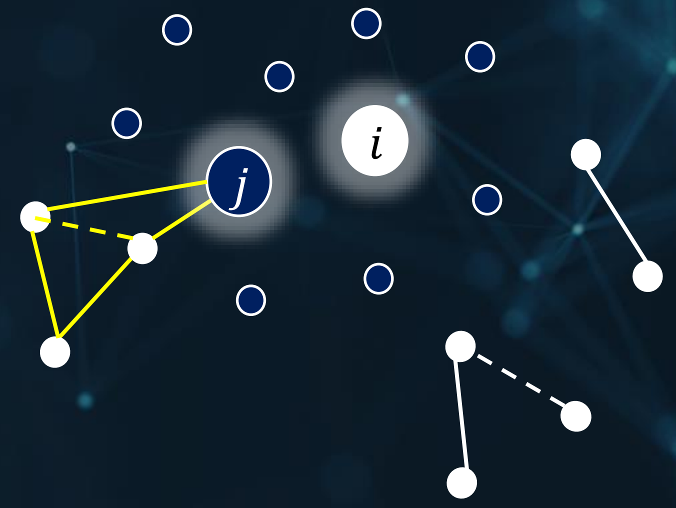
חישוב $H_{i\leftarrow j}(z)$

$$H_{i\leftarrow j}(z|\Gamma_{j\setminus i}) = z\prod_{k\in N_{j\setminus i}^{(r)}} \left[H_{j\leftarrow k}(z)\right]^{w_{j\setminus i,k}}$$

$$H_{i\leftarrow j}(z) = zG_{i\leftarrow j}\left(H_{j\leftarrow}(z)\right)$$

הקשתות הצהובות

הן בדיוק $N_{j\setminus i}^{(r)}$

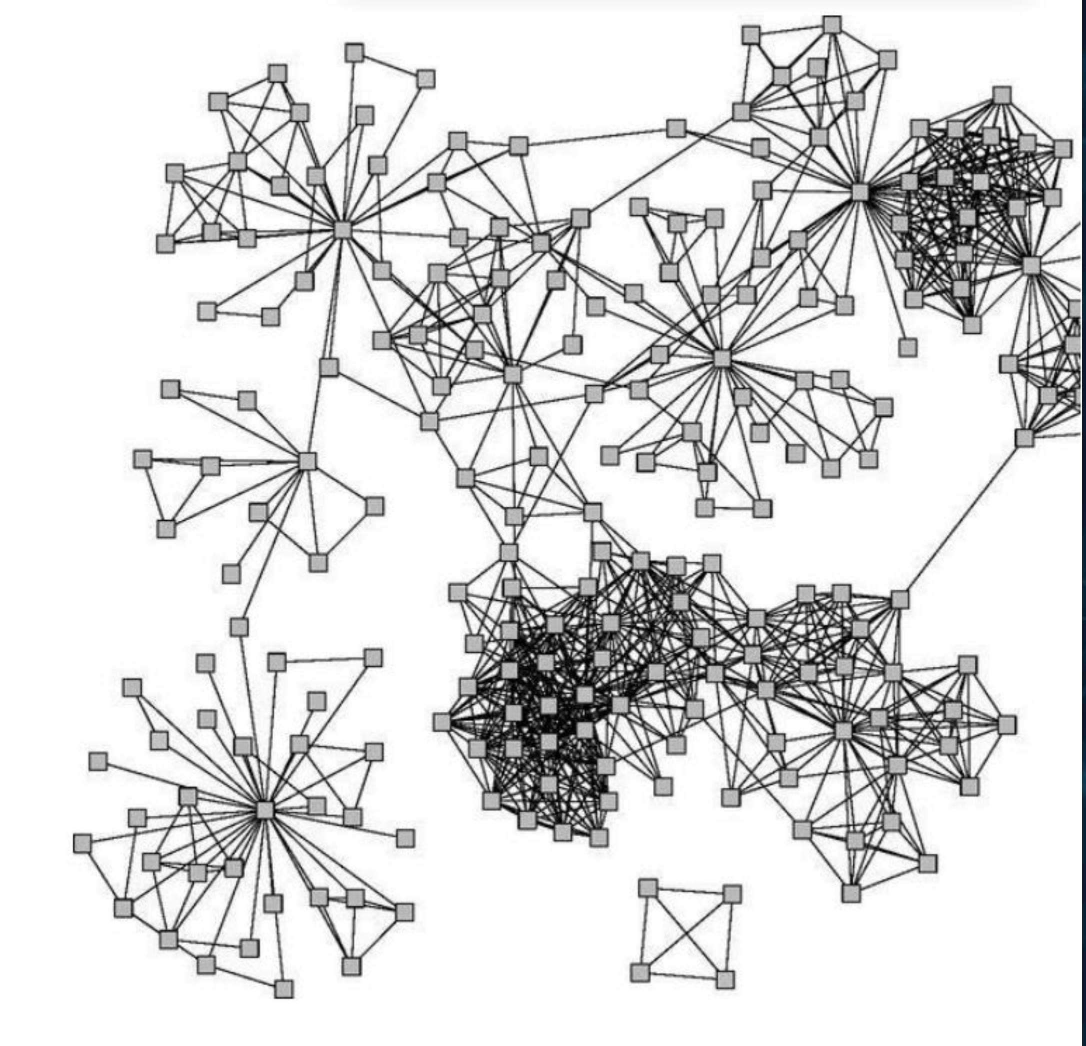


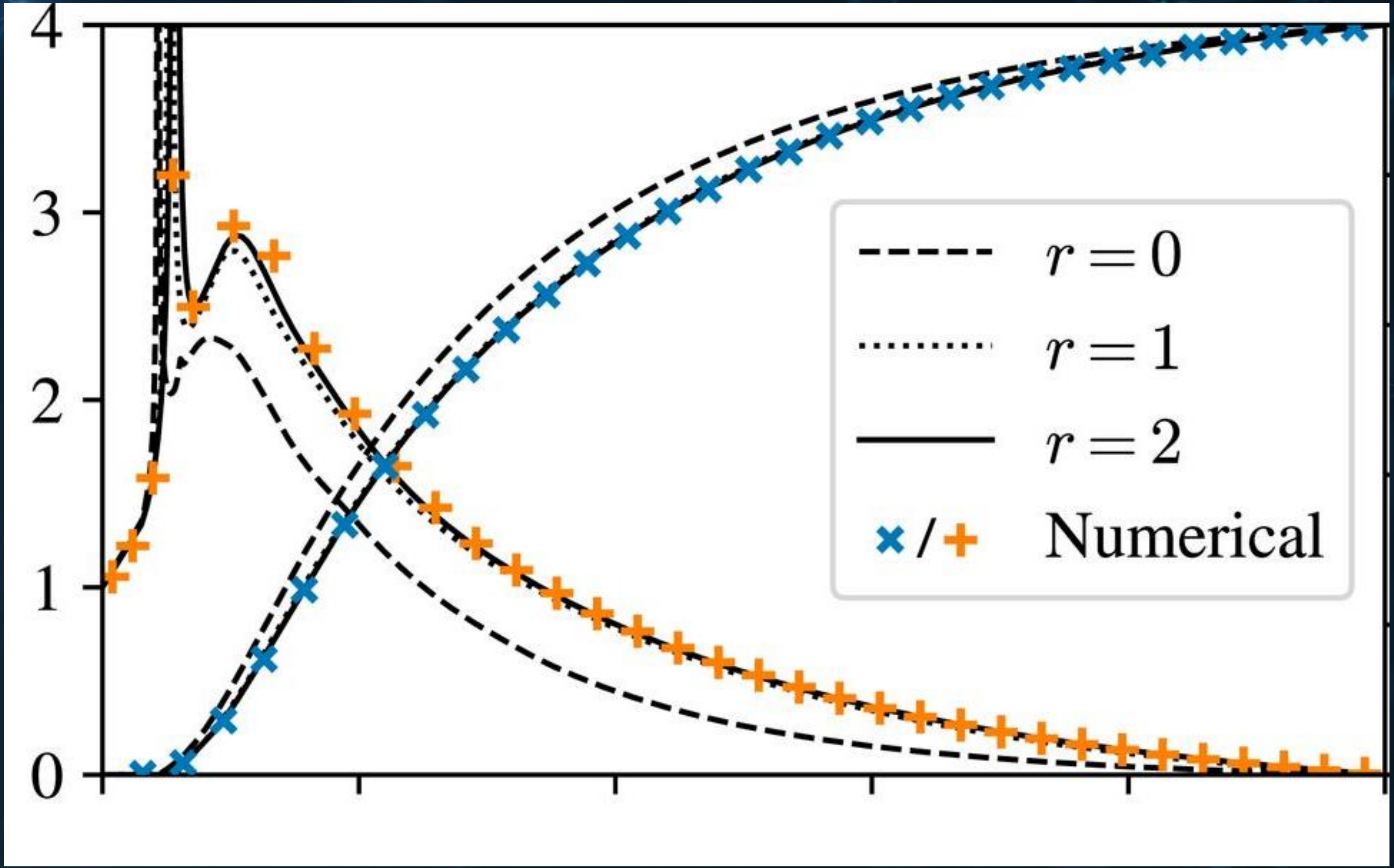
הבחנות

- מהפונקציה היוצרת של $\pi_i(s)$ ‏($H_i(z) = \sum_s \pi_{i\leftarrow j}(s) z^s$) אפשר לחשב את ההסתברות שצומת i שייך ל-$cluster$ כלשהו: $H_i(1) = \sum_s \pi_i(s)$.
- אפשר לחשב את הגודל הממוצע של s_i באופן הבא:

$$\langle s_i \rangle = \sum_{s=1}^{\infty} s * \pi_i(s) = H_i'(1)$$

דוגמא: רשת של שיתופי פעולה בכתיבת מאמרים אקדמיים





4
3
2
1
0
$r=0$
$r=1$
$r=2$
×/+ Numerical

סיכום

האלגוריתם שתיארנו יעיל יחסית, אבל גישת ה*message passing* לא יעילה יותר מהגישות המסורתיות של חישוב מספרי של סימולציות של *percolation*.

סיכום

עם זאת, 2 הגישות מחשבות דברים שונים.

הגישה הנאיבית:

בביצוע אחד מחשבת

עבור מימוש אחד

גישת ה*message passing*:

בביצוע אחד מחשבת את הממוצע

עבור כל מימוש

אם ברצוננו לחשב את הממוצע, בגישה הנאיבית נצטרך לבצע כמה חישובים ואילו בגישת ה*message passing* נצטרך לחשב חישוב בודד בלבד.

לכן, במצבים מסוימים גישת ה*message passing* יעילה יותר.