Rozszerzanie miar centralności do grup

PIOTR KĘPCZYŃSKI

Dotychczasowa wiedza

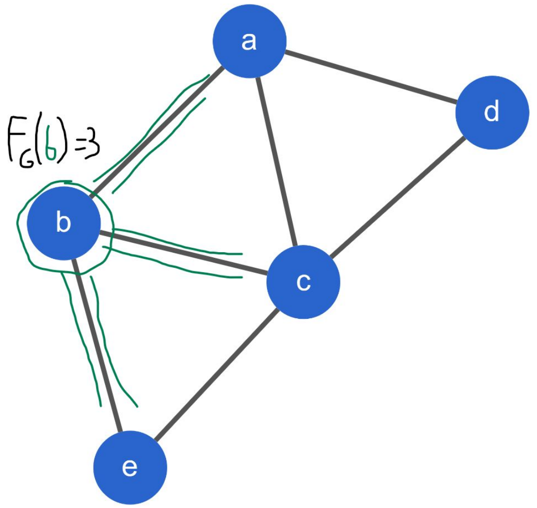
Wcześniej znane sposoby tworzenia grupowych miar centralności z wierzchołkowych.

Miary centralności

WIERZCHOŁKOWE MIARY CENTRALNOŚCI

$F': G \times V \to \mathbb{R}$

Zapisywane jako $F'_G(v)$

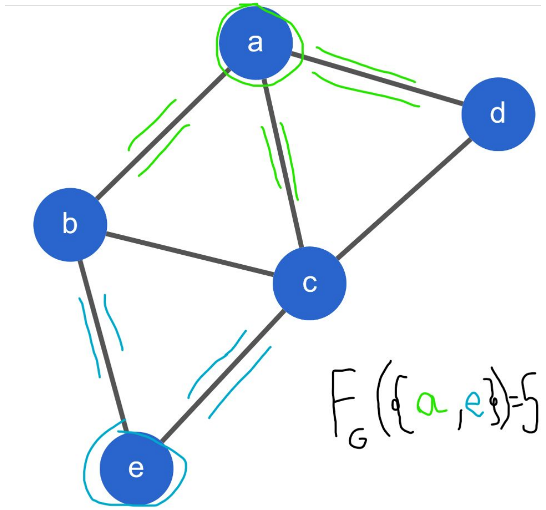


Dla grafu $G = (V, E)$

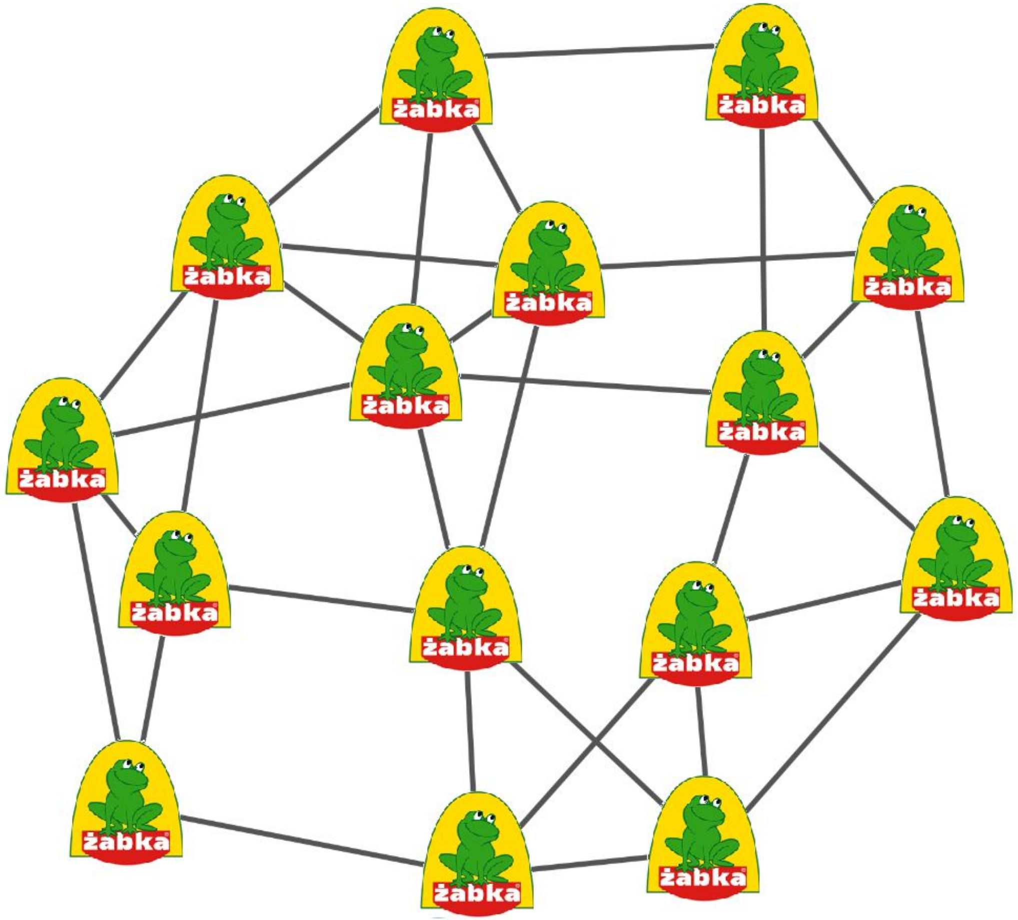
GRUPOWE MIARY CENTRALNOŚCI

$F: G \times 2^V \to \mathbb{R}$

Zapisywane jako $F_G(S)$

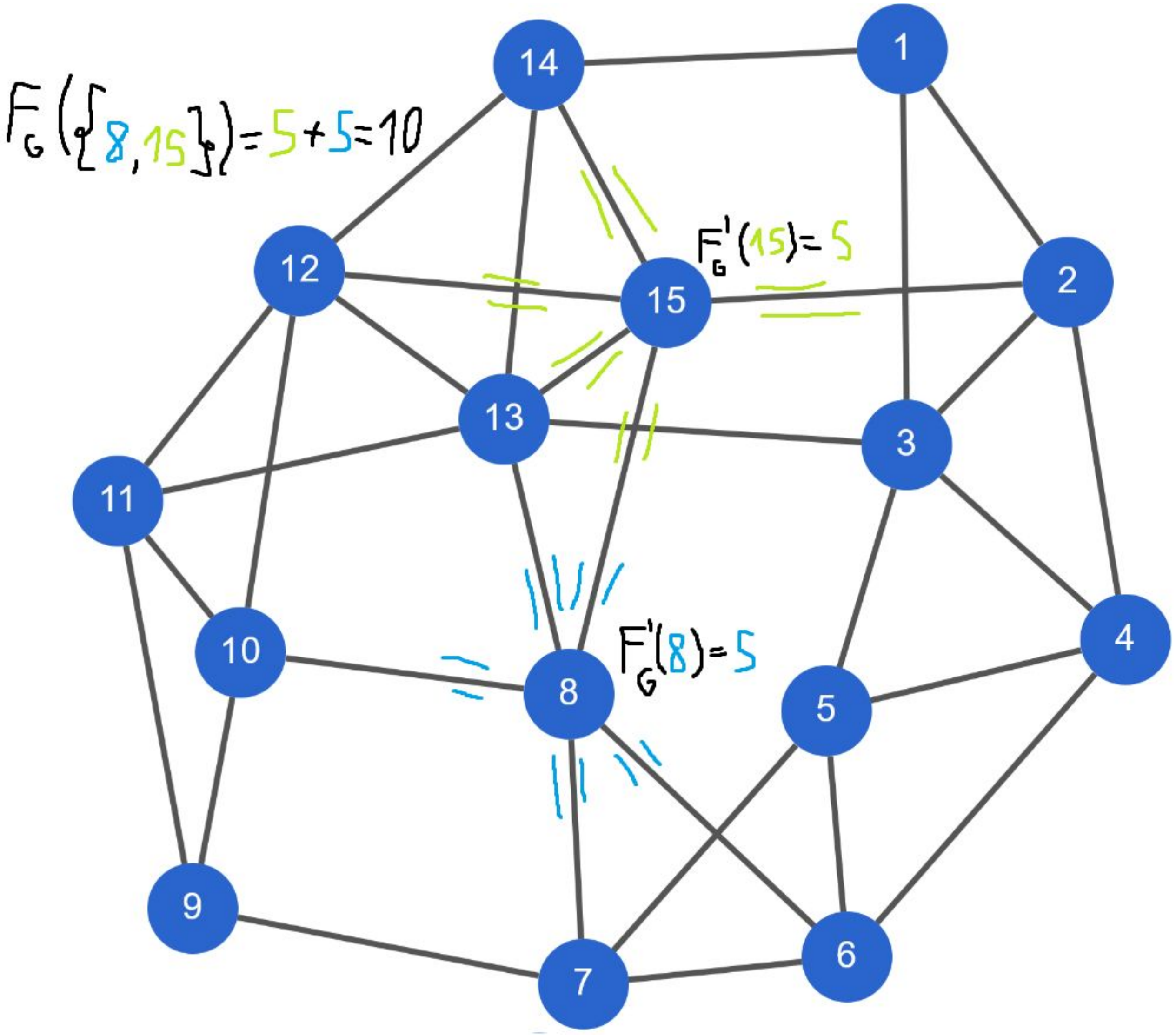


żabka ?



Metoda sumowania

$$F_G(S) = \sum_{v \in S} F'_G(v)$$



$$F_G(S) = \sum_{v \in S} F'_G(v)$$

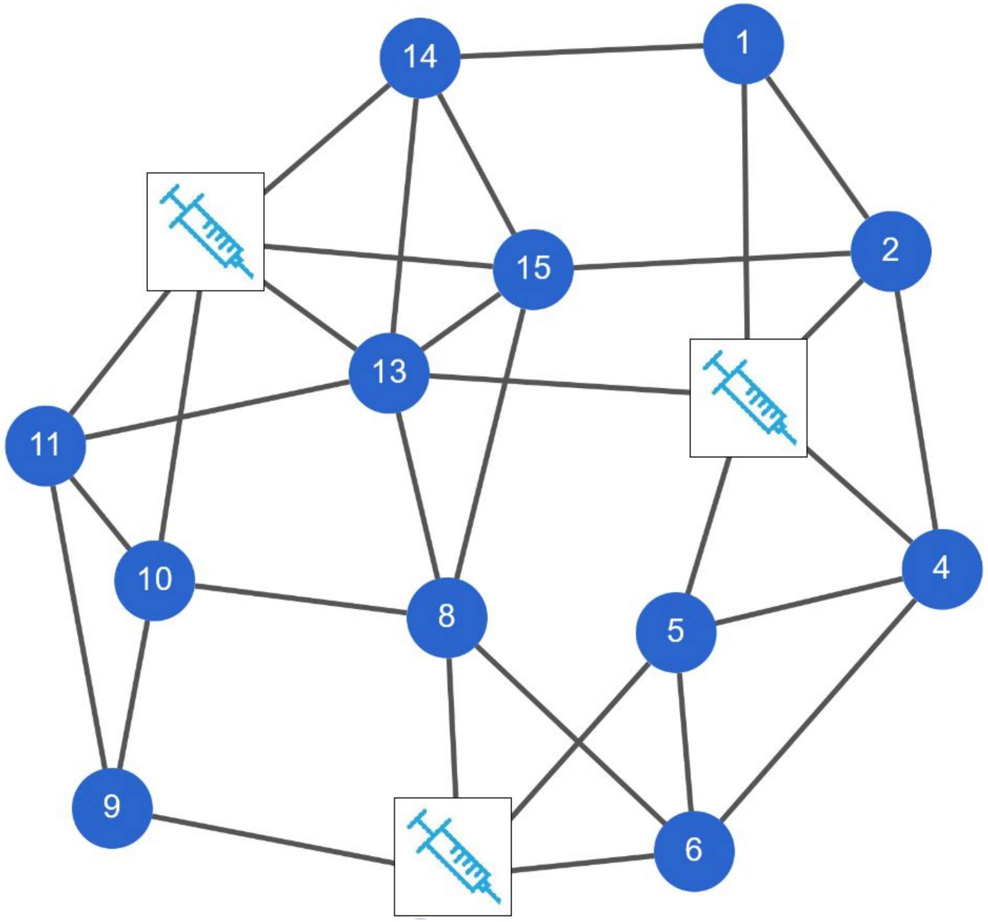
x3 ?

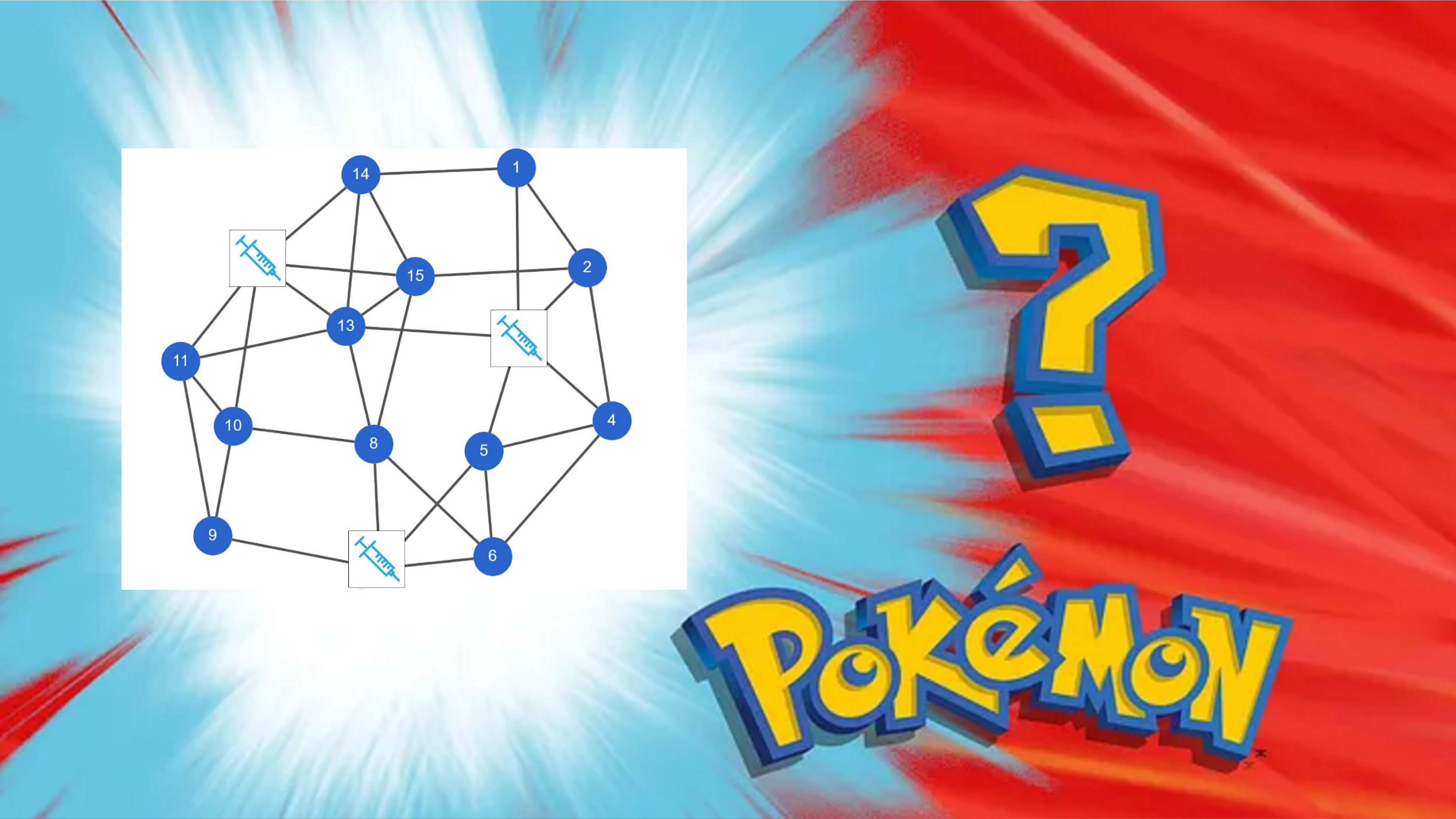
Wierzchołkowa centralność: Closeness

$$F'_G(v) = \frac{1}{\sum_{u \in V} dist(u, v)}$$

Metoda sumowania daje niesatysfakcjonujące wyniki

Czy moglibyśmy dostać coś takiego?



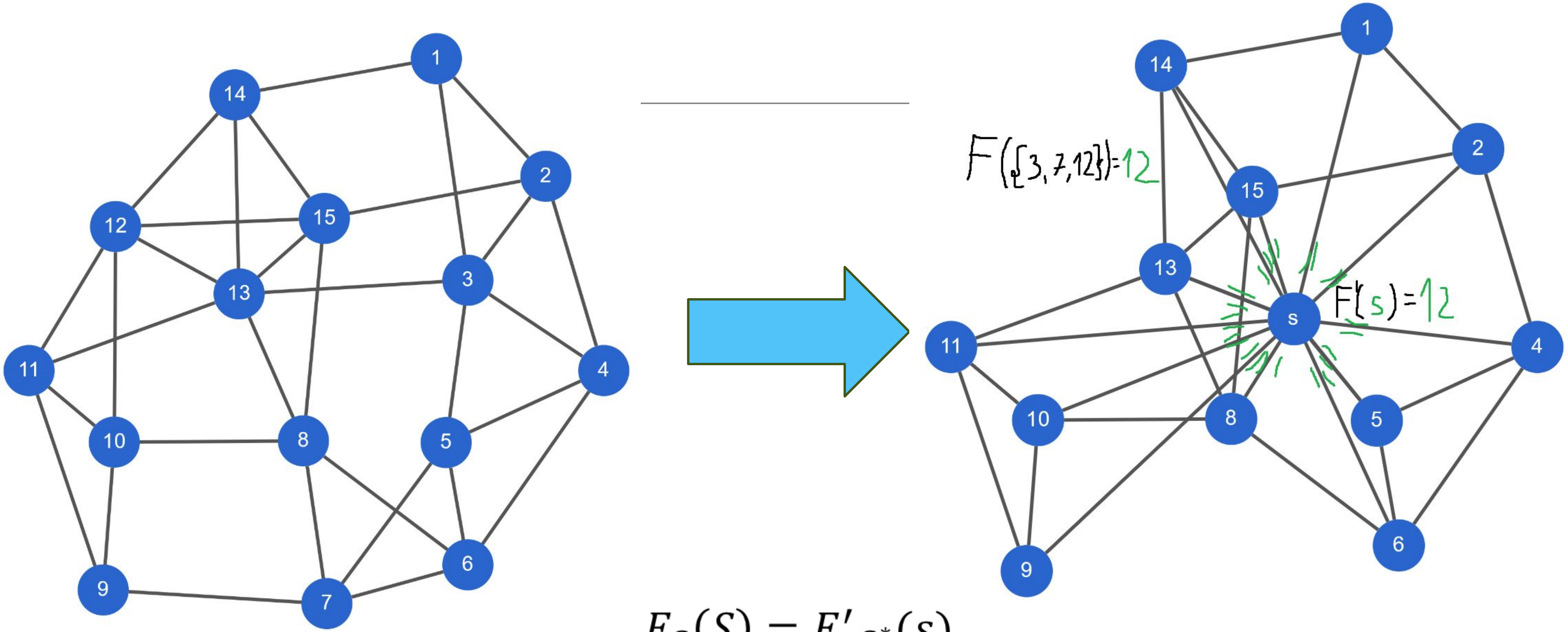

14
1
2
15
13
11
10
8
5
4
9
6
?
Pokémon

Metoda łączenia

$$F_G(S) = F'_{G^*}(s)$$

G^* to graf z sklejonymi wierzchołkami z grupy.

s to wierzchołek powstały z sklejonych wierzchołków grupy.



$$F_G(S) = F'_{G^*}(s)$$

G^* to graf z sklejonymi wierzchołkami z grupy.

s to wierzchołek powstały z sklejonych wierzchołków grupy.

x2

Wierzchołkowa centralność: Betweenness

$$F'_G(v) = \sum_{s,t \in V \setminus \{v\}} \frac{|p \in \Pi_s(s,t) : v \in p|}{|\Pi_s(s,t)|}$$

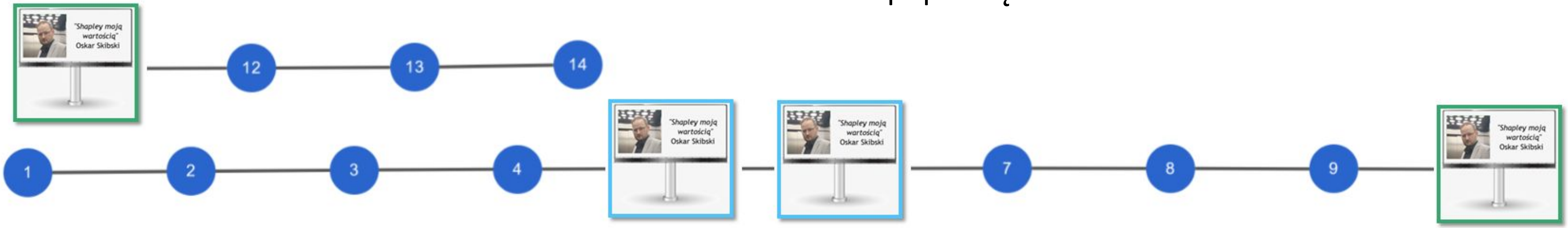
Gdzie $\Pi_s(s,t)$ to zbiór najkrótszych ścieżek z s do t.

Który zestaw jest lepszy - zielony czy niebieski?

Metoda łączenia: zielony > niebieski.

Metoda sumowania (i ja): to nie ma sensu.

Pani w cukierni: poproszę ciasto.

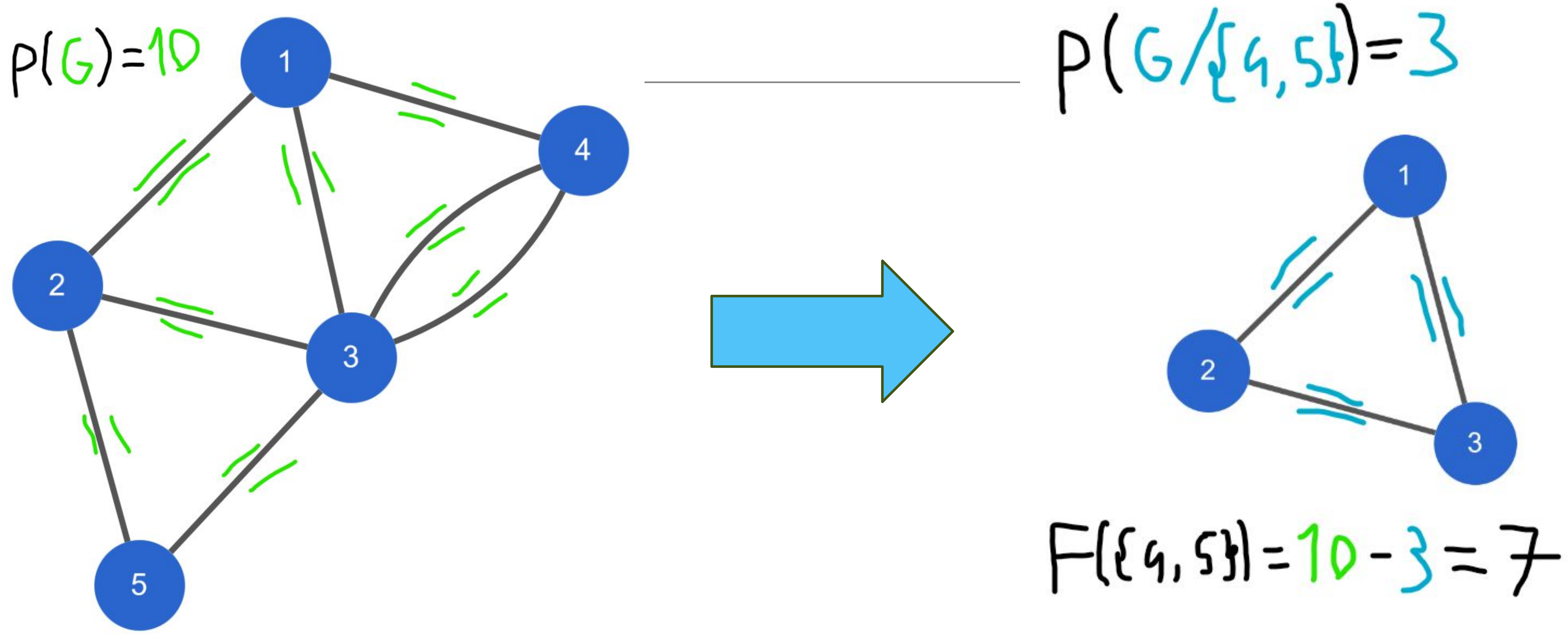


Metoda totalna

Tylko dla wierzchołkowych miar centralności vitality!

Miary centralności vitality:
$\exists_{p:G\to\mathbb{R}} \forall_G \; F'_G(v) = p(G) - p(G \setminus \{v\})$

Wtedy
$F_G(S) = p(G) - p(G \setminus S)$

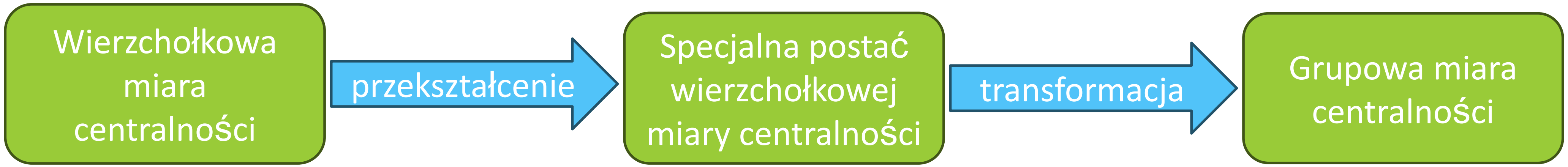


$$F_G(S) = p(G) - p(G \setminus S)$$

Nowy pomysł

Nowy schemat do transformacji wierzchołkowych miar centralności w grupowe.

Schemat



Przekształcenie:
- Niejednoznaczne
- Nie zawsze się sensownie da
- Trzeba podjąć pewne wybory

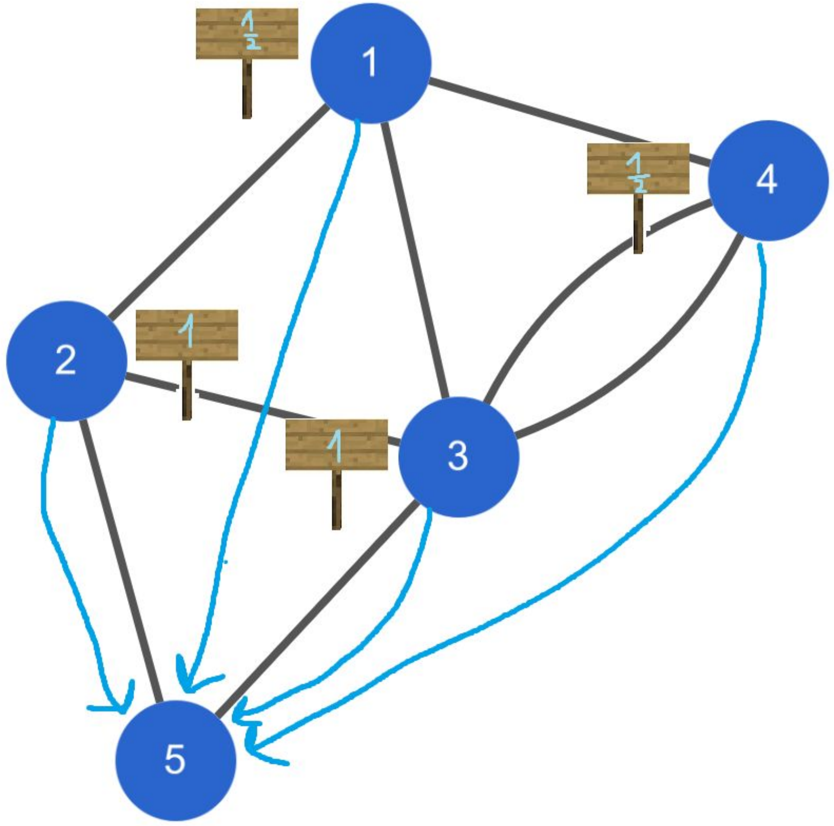
Transformacja:
- Jednoznaczna
- Elegancka
- Naturalna


transformacja
przekształcenie

Specjalna postać

Harmonic:

$F'_G(v) = \sum_{u \in V} \frac{1}{dist(u,v)}$



Specjalna postać*

$$F'_G(v) = \sum_{u \in V} f_G\,(u, v)$$

- Funkcja oceniająca $f: G \times V \times V \to \mathbb{R}$

Problemy:

Betweenness

$$F'_G(v) = \sum_{s,t \in V \setminus \{v\}} \frac{|p \in \Pi_s(s,t) \colon \mathrm{v} \in p|}{|\Pi_s(s,t)|}$$

Attachment

$$F'_G(v) = \sum_{U \subseteq V \setminus \{v\}} \frac{|U|!\,(|V| - |U| - |1|)!}{|V|!}\, 2(|K(G[U])| - |K(G[U \cup \{v\}])| + 1)$$

*prawie

Specjalna postać*

$$F'_G(v) = \sum_{U \in \mathcal{F}^W} f_G(U, v)$$

- Funkcja oceniająca $f\colon G \times 2^V \times V \to \mathbb{R}$
- Klasa wierzchołków $\mathcal{F}^W \subseteq 2^W, W \subseteq V$

Problemy:

Closeness

$$F'_G(v) = \frac{1}{\sum_{u \in V} dist(u, v)}$$

*prawie

Specjalna postać

$$F'_G(v) = g\left(\sum_{U \in \mathcal{F}^W} f_G(U, v)\right)$$

- Funkcja oceniająca $f: G \times 2^V \times V \rightarrow \mathbb{R}$
- Klasa wierzchołków $\mathcal{F}: W \rightarrow 2^{2^V}, W \subseteq V$
- Funkcja normalizująca $g: \mathbb{R} \rightarrow \mathbb{R}$, rosnąca

Transformacja

$$F'_G(v) = g(\sum_{U \in \mathcal{F}^W} f(U,v))$$

sum

$$F_G(S) = g(\sum_{U \in \mathcal{F}^W} \sum_{v \in S} f_G(U,v))$$

max

$$F_G(S) = g(\sum_{U \in \mathcal{F}^W} \sum_{v \in S} f_G(U,v))$$

min

$$F_G(S) = g(\sum_{U \in \mathcal{F}^W} \sum_{v \in S} f_G(U,v))$$

proportional

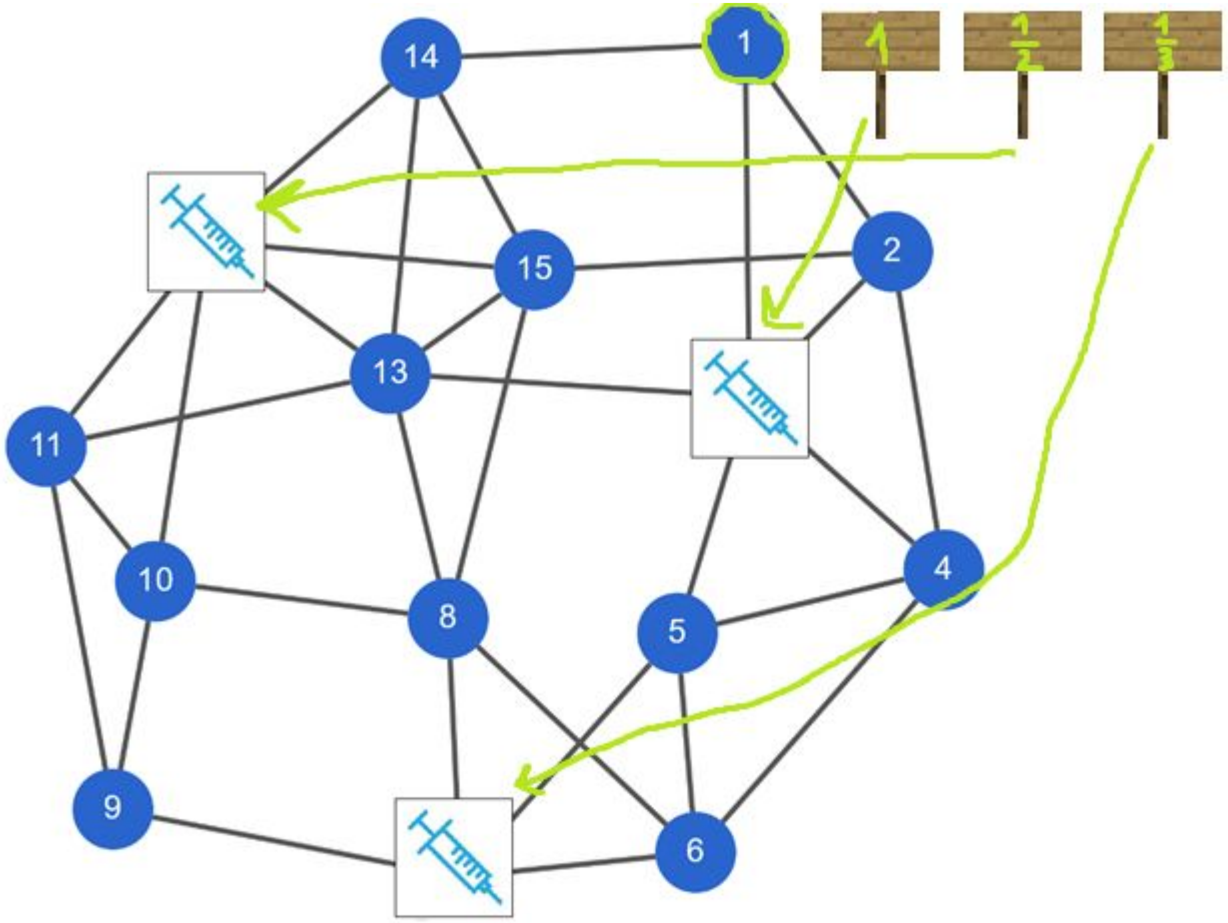
$$F_G(S) = g(\sum_{U \in \mathcal{F}^W} \sum_{v \in S} f_G(U,v))$$

$pav: \mathbb{R}^n \rightarrow \mathbb{R}$

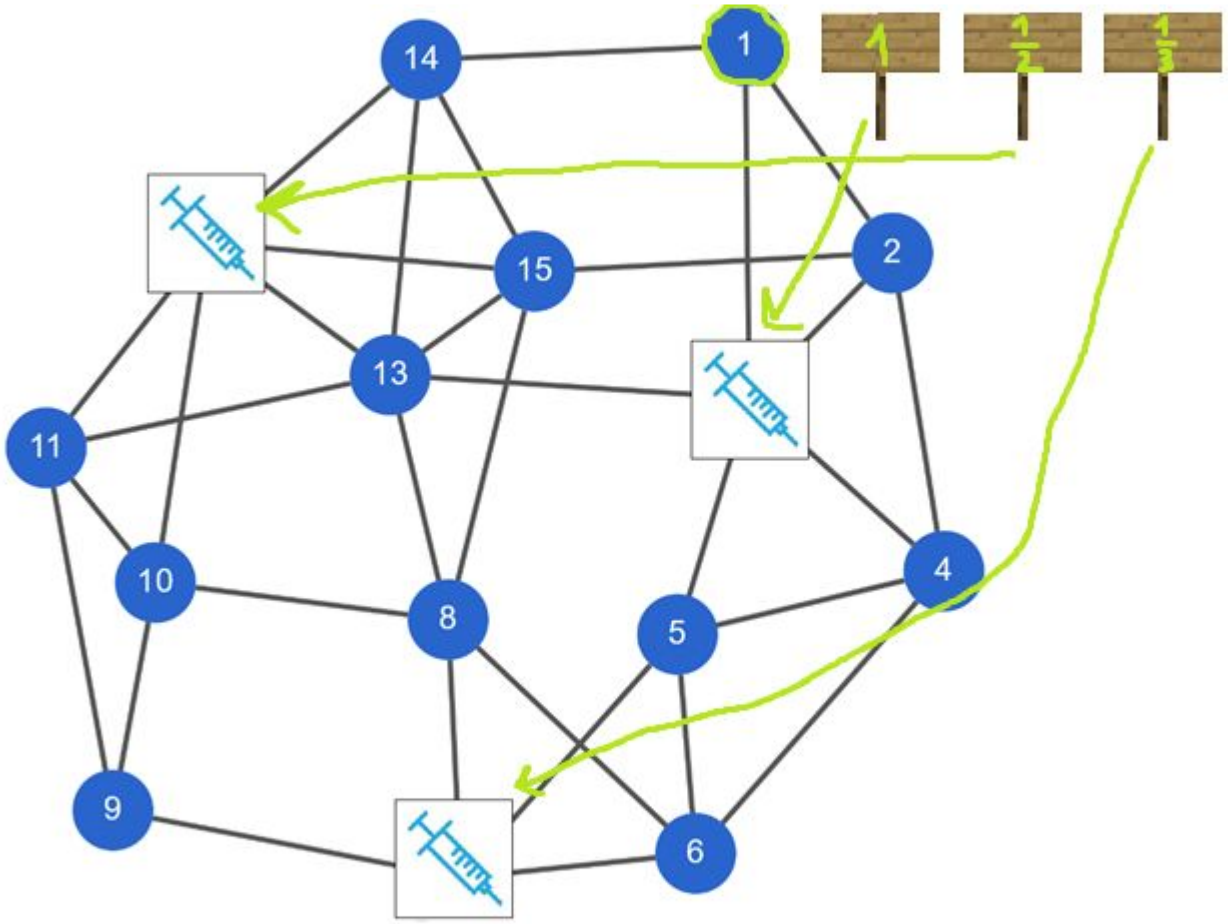
$pav(A_{1\ldots n}) = \sum_{i=1}^{n} \frac{1}{i} A_i$

(A jest posortowane malejąco)

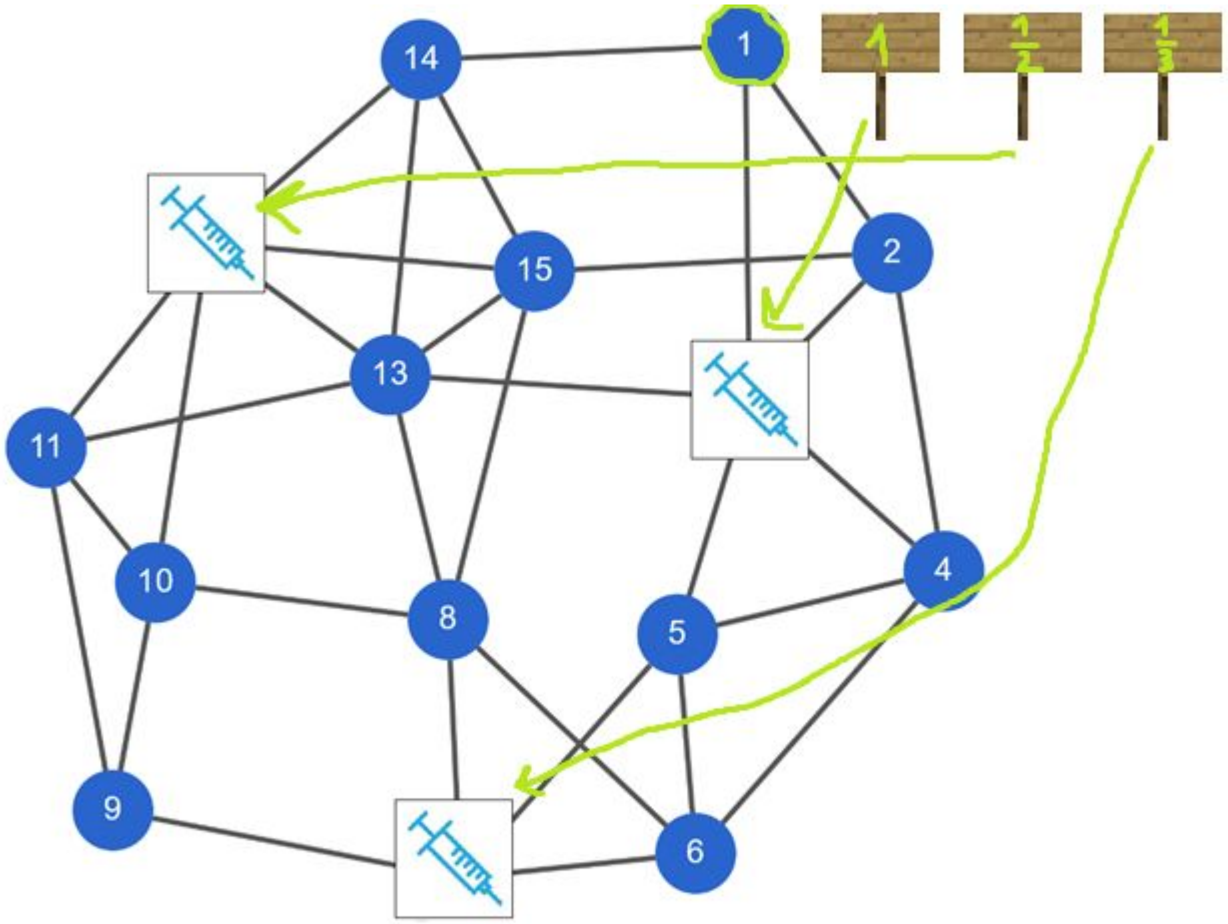
Intuicja



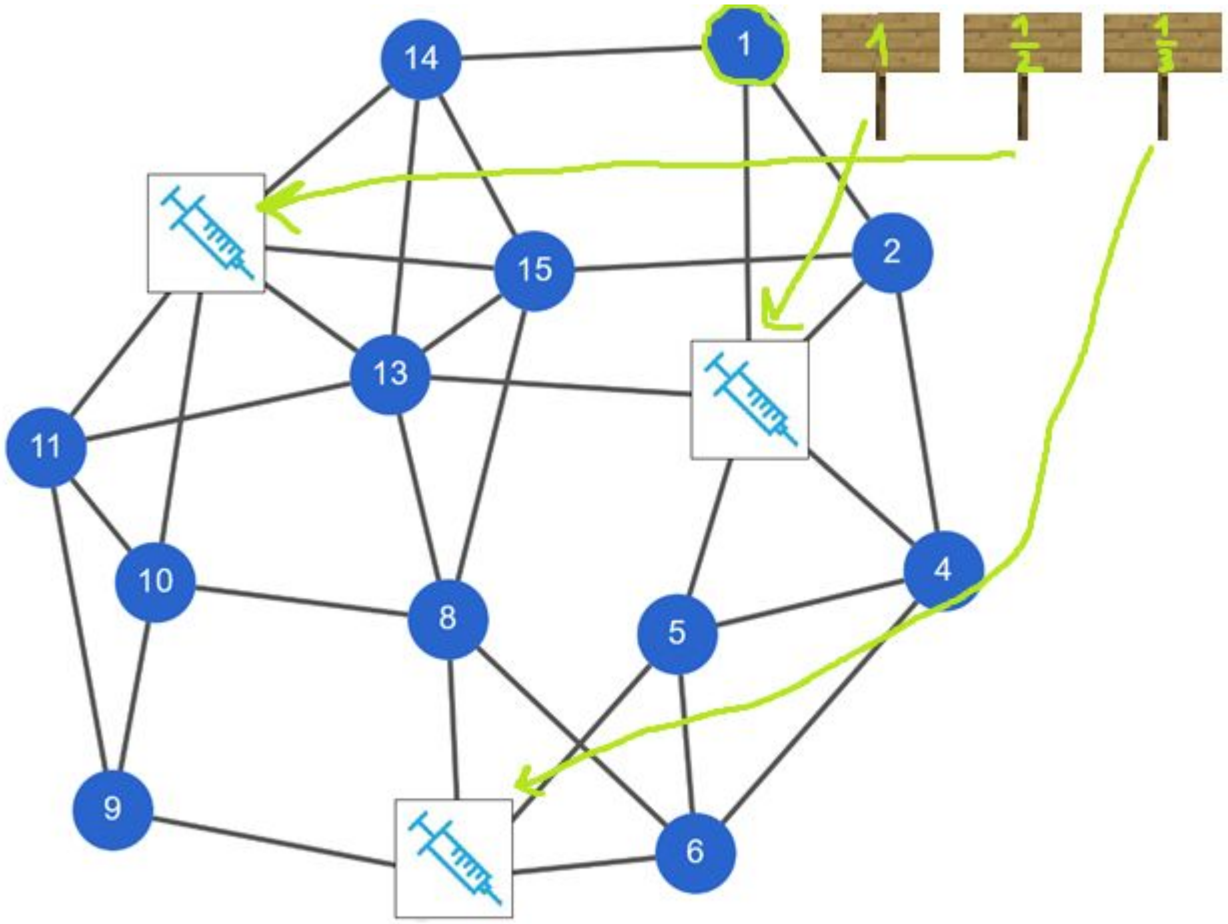
Intuicja - sum



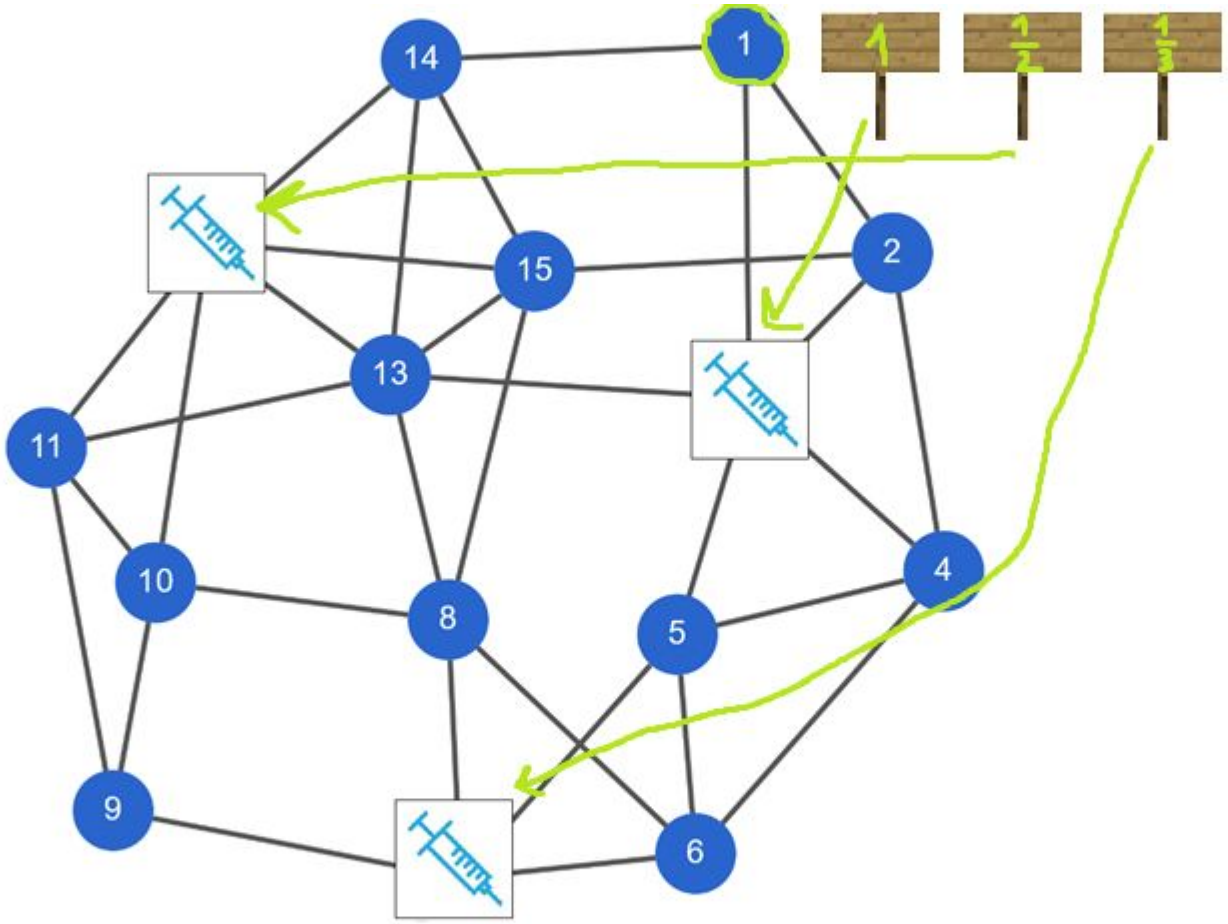
Intuicja - max



Intuicja - min



Intuicja - proportional



Problem

$$CB_v(G) = \frac{1}{(n-1)(n-2)} \sum_{s,t\in V\setminus\{v\}} \frac{1}{2}(-|b_{s,t}(v)| + \sum_{e:v\in e} |x(e)|)$$

$$B_v(G) = \sum_{s,t\in V\setminus\{v\}} \frac{|p\in\Pi_s(s,t):v\in p\}|}{|\Pi_s(s,t)|}$$

$$A_v(G) = \sum_{S\subseteq V\setminus\{v\}} \frac{|S|!(|V|-|S|-1)!}{|V|!} 2(|K(G[S])| - |K(G[S\cup\{v\}])| + 1)$$

$$C_v(G) = \sum_{u\in V\setminus\{v\}} \frac{1}{dist(u,v)}$$

$$\frac{1}{\sum_{u\in V\setminus\{v\}} dist(u,v)}$$

$$PR_v(G) = a \cdot \left(\sum_{u\in N_G(v)} \frac{PR_u(G)}{D_u(G)}\right) + b_v$$

$$Y_v(G) = \sum_{u\in V\setminus\{v\}} \delta^{dist(u,v)}$$

$$FB_v(G) = \sum_{s,t\in V\setminus\{v\}} flow_{s,t}(G) - flow_{s,t}(G[V\setminus v])$$

Możliwe rozwiązania:

1. Oceniana grupa nie ocenia swoich wierzchołków nawzajem.
2. Wierzchołki nie oceniają same siebie.
3. Wierzchołki oceniają same siebie.

A. Graf bloków mieszkalnych i stawianie w nich żabek.
B. Modelowanie zarażanej populacji – która grupa osób zarażonych na początku zarazi największą część społeczności.

Wcześniej znane rozwiązania w nowym schemacie

Metoda sumowania	sum, V
Metoda łączenia	max, V \ S* (nie zawsze tożsame)
Metoda totalna	nie da się
	sum, V \ S, $\mathcal{F}$ = {V\S}*** (bardzo oszukane)

Ekspresywność - przekształcenie

Eccentricity

$$C_v(G) = \frac{1}{\max_{u \in V} dist(v,u)}$$

$$F'_v(G) = g\left(\sum_{U \in \mathcal{F}^V} f_G(U,v)\right)$$

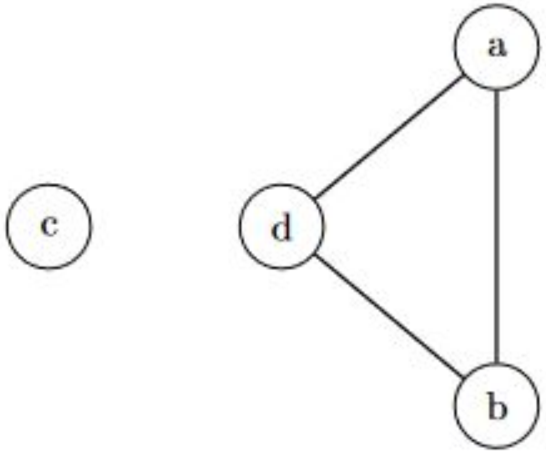
$$g(x) = x$$

$$f_G(U,v) = \begin{cases} F'_v(G) & U = \{v\} \\ 0 & \text{oth.} \end{cases}$$

$$\mathcal{F}^V = 2^V$$

Ekspresywność – osiągalne grupowe centralności (V)

Dla konkretnej miary vitality:
$p(G) = $ [G ma krawędzie]



- min
It cannot be expressed in our framework with min, since adding nodes to the group cannot make the measure bigger in the V variant, and $F_{\{a,b\}} > F_{\{a\}}(G)$
- sum
Let's assume it's possible to express the group measure using a sum variant in our framework. Then

$$g\left(\sum_{U \in \mathcal{F}^V} f_G(U,a)\right) = g\left(\sum_{U \in \mathcal{F}^V} f_G(U,c)\right) = g\left(\sum_{U \in \mathcal{F}^V} f_G(U,a) + \sum_{U \in \mathcal{F}^V} f_G(U,c)\right) = 0$$

Since g is increasing it implies

$$\sum_{U \in \mathcal{F}^V} f_G(U,a) = \sum_{U \in \mathcal{F}^V} f_G(U,c) = 0, g(0) = 0$$

For analogous reasons

$$\sum_{U \in \mathcal{F}^V} f_G(U,b) = \sum_{U \in \mathcal{F}^V} f_G(U,c) = 0$$

But then

$$F_{\{a,b\}}(G) = g\left(\sum_{U \in \mathcal{F}^V} f_G(U,a) + f_G(U,b)\right) = g\left(\sum_{U \in \mathcal{F}^V} f_G(U,a) + \sum_{U \in \mathcal{F}^V} f_G(U,b)\right) = g(0+0) = 0 \neq 1$$

Which is a contradiction.
- max
Let's assume it's possible to express the group measure using a max variant in our framework. Then

$$g\left(\sum_{U \in \mathcal{F}^V} f_G(U,a)\right) = g\left(\sum_{U \in \mathcal{F}^V} f_G(U,c)\right) = g\left(\sum_{U \in \mathcal{F}^V} \max(f_G(U,a), f_G(U,c))\right) = 0$$

This implies

$$\forall_{U \in \mathcal{F}^V} f_G(U,a) = f_G(U,c)$$

For analogous reasons

$$\forall_{U \in \mathcal{F}^V} f_G(U,b) = f_G(U,c)$$

Which implies

$$\forall_{U \in \mathcal{F}^V} f_G(U,a) = f_G(U,b)$$

Which implies

$$F_{\{a,b\}}(G) = g\left(\sum_{U \in \mathcal{F}^V} \max(f_G(U,a), f_G(U,b))\right) = g\left(\sum_{U \in \mathcal{F}^V} f_G(U,a)\right) = 0 \neq 1$$

Which is a contradiction

Ekspresywność – osiągalne grupowe centralności (V \ S)

$$g\left(\sum_{U\in\mathcal{F}^{V\setminus S}}\sum_{v\in S} f_G(U,v)\right)$$

$$g(x) = x$$

$$f_G(U,v) = \frac{F_{V\setminus U}(G)}{|V\setminus U|}$$

$$\mathcal{F}^W = \{W\}$$

$$\sum_{U\in\mathcal{F}^{V\setminus S}}\sum_{v\in S} f_G(U,v) = \sum_{v\in S} f_G(V\setminus S, v) = \sum_{v\in S}\frac{F_{V\setminus(V\setminus S)}(G)}{|V\setminus(V\setminus S)|} = \sum_{v\in S}\frac{F_S(G)}{|S|} = |S|\frac{F_S(G)}{|S|} = F_S(G)$$

$$F'_G(v) = g\left(\sum_{U \in \mathcal{F}^W} f_G(U, v)\right)$$

Przekształcanie klas miar – medial

LOOSE DEFINITION

A centrality measure is a *medial centrality* if there exists a function $\Delta_v^{s,t}(G)$ that evaluates the role of v in connecting s, t in G such that

$$F_v(G) = \sum_{s,t \in V} \Delta_v^{s,t}(G)\,.$$

Betweenness

$$B_v(G) = \sum_{s,t \in V \setminus \{v\}} \frac{|p \in \Pi_s(s,t) : v \in p\}|}{|\Pi_s(s,t)|}$$

Stress

$$S_v(G) = \sum_{s,t \in V \setminus \{v\}} |p \in \Pi_s(s,t) : v \in p|$$

Flow Betweenness

$$FB_v(G) = \sum_{s,t \in V \setminus \{v\}} flow_{s,t}(G) - flow_{s,t}(G[V \setminus v])\,,$$

Przekształcanie klas miar – distance-based

$$F'_G(v) = g(\sum_{U \in \mathcal{F}^W} f_G(U, v))$$

DEFINITION

A centrality measure is a *distance-based centrality* if there exists a function $f: \mathbb{R}^n \to \mathbb{R}$ such that

$$F_v(G) = f\left(\left(dist_{u,v}(G)\right)_{u \in V}\right)$$

Distance-based Centralities

$$\sum_{u \in V \setminus \{v\}} a_{dist_{u,v}(G)}$$

Alternatives:

ADDITIVE

		a_1	a_2	a_3	$a_{i>3}$
Harmonic [Rochat 2009]	$H_v(G) = \sum_u 1/dist_{u,v}(G)$	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{i}$
Decay [Jackson 2005]	$Y_v(G) = \sum_u \delta^{dist_{u,v}(G)}$ where $\delta \in (0,1)$	δ^0	δ^1	δ^2	δ^i
Degree	$D_v(G) = \lvert\{u \in V \setminus \{v\} : (u,v) \in V\}\rvert$	1	0	0	0
k-Step Reach	$R_v^k(G) = \left\lvert\{u \in V \setminus \{v\} : dist_{u,v}(G) \leq k\}\right\rvert$	1	1	1	$[i \leq k]$
Eccentricity	$EC_v(G) = \max_{u \in V}\{dist_{u,v}(G)\}$				

Przekształcanie klas miar – distance-based

$$F'_G(v) = g\left(\sum_{U \in \mathcal{F}^W} f_G(U,v)\right)$$

DEFINITION

A centrality measure is a *distance-based centrality* if there exists a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$F_v(G) = f\left(\left(dist_{u,v}(G)\right)_{u \in V}\right)$$

$$\sum_{u \in V \setminus \{v\}} a_{dist_{u,v}(G)}$$

- Closeness

$$C_v(G) = \frac{1}{\sum_u dist_{u,v}(G)}$$

$$F'_G(v) = g\Big(\sum_{U \in \mathcal{F}W} f_G(U, v)\Big)$$

Przekształcanie klas miar – vitality

DEFINITION

A centrality measure is a *vitality index* if there exists a function $f : \mathcal{G} \to \mathbb{R}$ such that

$$F_v(G) = f(G) - f(G - v)$$

THEOREM

A centrality measure is a vitality index if and only if it is a Shapley-value based induced game-theoretic centralities.

$$F'_G(v) = g(\sum_{U \in \mathcal{F}^W} f_G(U, v))$$

Przekształcanie klas miar – vitality

DEFINITION

A centrality measure is a *vitality index* if there exists a function $f: \mathcal{G} \to \mathbb{R}$ such that

$$F_v(G) = f(G) - f(G - v)$$

$$\sum_{U \subseteq V, U \neq \emptyset} f(U, v))$$

$$f(U, v) = \frac{(|U| - 1)!(|V| - |U|)!}{|V|!}(F_\Sigma(U) - F_\Sigma(U \setminus \{v\}))$$

$F_\Sigma(U)$ is the sum of single node centralities of all nodes in the subgraph induced by U.

Źródła:

Axioms4Centalities

https://centrality.mimuw.edu.pl/

Oskar Skibski - Presentation of MsC topics.

https://aiecon.mimuw.edu.pl/wp-content/uploads/2023/10/Centrality-Measures-SEM.pdf