

# TEORIA DEI GIOCHI COOPERATIVI E VALUTAZIONE DELLA CENTRALITÀ DEI NODI IN UN NETWORK BIOLOGICO

Relatrice:  
Anna Torre

Candidata:  
Giulia Cesari

Correlatore:  
Stefano Moretti

Introduzione

Network e  
centralità:  
misure  
classiche

Teoria dei  
giochi

Teoria dei  
giochi e  
centralità in  
un network

Nuovi  
modelli di  
centralità in  
un network  
biologico

Applicazione  
a dati reali

Conclusioni

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

Applicazione a dati reali

Conclusioni



Stefano Moretti - IST - Epidemiologia molecolare  
LAMSADE - Università Paris-Dauphine

Stefano Bonassi - IST - Epidemiologia molecolare  
IRCCS San Raffaele Pisana - Roma

IST - Istituto Nazionale per la Ricerca sul Cancro -  
Oncopatologia Traslazionale - Genova



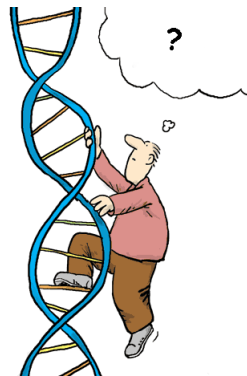
Associazione Italiana per la Lotta al Neuroblastoma



## Scopo del lavoro

Analisi di alcuni approcci con la teoria dei giochi cooperativi al problema dell'**individuazione di geni rilevanti all'interno di una condizione biologica.**

- Moretti et al. (2007): *microarray games*
- Moretti et al. (2010): approccio con i giochi di coalizione alla centralità di geni in un network biologico
- Nuovi modelli di centralità di geni in un network biologico





## Network di coespressione

- i nodi del network rappresentano i geni
- la connessione tra nodi è determinata dalla coespressione dei geni nei campioni in esame (coefficiente di correlazione di Pearson )

## Centralità in un network di coespressione

Lo studio della *centralità* di elementi in un network rappresenta un importante strumento per l'interpretazione dell'interazione tra geni all'interno di network di coespressione.

- Jeong et al. (2001): "Lethality and centrality in protein networks."
- Carlson et al. (2006): "Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks."

## Network di coespressione

- i nodi del network rappresentano i geni
- la connessione tra nodi è determinata dalla coespressione dei geni nei campioni in esame (coefficiente di correlazione di Pearson )

## Centralità in un network di coespressione

Lo studio della *centralità* di elementi in un network rappresenta un importante strumento per l'interpretazione dell'interazione tra geni all'interno di network di coespressione.

- Jeong et al. (2001): "Lethality and centrality in protein networks."
- Carlson et al. (2006): "Gene connectivity, function, and sequence conservation: predictions from modular yeast co-expression networks."







# Centralità in un network: misure classiche

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

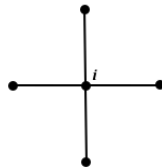
Applicazione a dati reali

Conclusioni

Centralità - Bavelas (1948): studi sulla comunicazione in gruppi di individui

Idea di base: il fulcro di una stella è la posizione più centrale possibile.

- $i$  può comunicare con molti altri nodi.
- vi sono molti altri nodi che necessitano di  $i$  come intermediario nelle loro comunicazioni.
- $i$  è vicino a molti altri nodi.



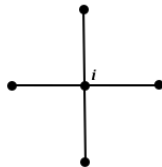
⇒ *degree centrality*  
⇒ *betweenness centrality*  
⇒ *closeness centrality*



## Centralità - Bavelas (1948): studi sulla comunicazione in gruppi di individui

Idea di base: il fulcro di una stella è la posizione più centrale possibile.

- $i$  può comunicare con molti altri nodi.
- vi sono molti altri nodi che necessitano di  $i$  come intermediario nelle loro comunicazioni.
- $i$  è vicino a molti altri nodi.

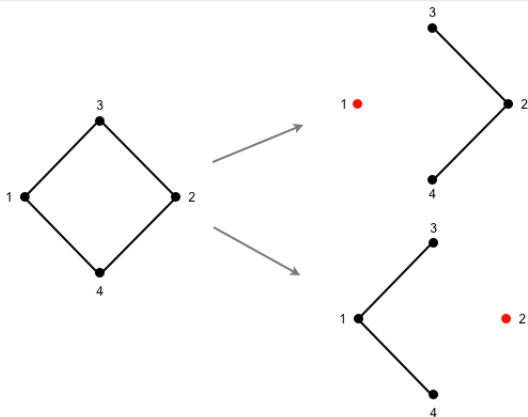


⇒ *degree centrality*  
⇒ *betweenness centrality*  
⇒ *closeness centrality*

# Misure classiche di centralità: limitazioni

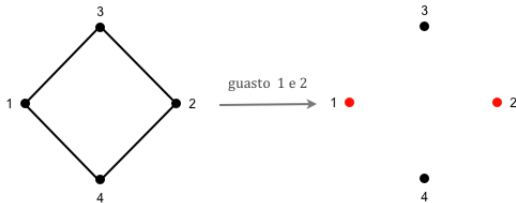
Le misure di centralità classiche possono rivelarsi inefficaci nel riflettere il ruolo delle coppie di nodi o più in generale di sottoinsiemi di nodi.

*Esempio:* rete di comunicazione (guasto di un nodo)



# Misure classiche di centralità: limitazioni

*Esempio:* rete di comunicazione (guasto di una coppia di nodi)



## *Game theoretic network centrality*

Un approccio all'analisi di centralità con la Teoria dei giochi (*game theoretic network centrality*) è stato proposto come quadro teorico per affrontare tali limitazioni.

Introduzione

Network e  
centralità:  
misure  
classiche

Teoria dei  
giochi

Teoria dei  
giochi e  
centralità in  
un network

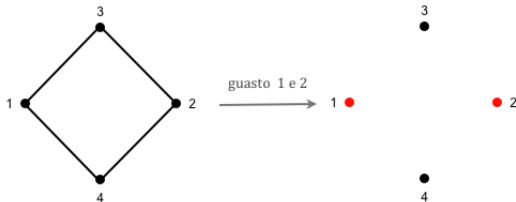
Nuovi  
modelli di  
centralità in  
un network  
biologico

Applicazione  
a dati reali

Conclusioni

# Misure classiche di centralità: limitazioni

*Esempio:* rete di comunicazione (guasto di una coppia di nodi)



## *Game theoretic network centrality*

Un approccio all'analisi di centralità con la Teoria dei giochi (*game theoretic network centrality*) è stato proposto come quadro teorico per affrontare tali limitazioni.

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

Applicazione a dati reali

Conclusioni

*"Theory of Games and Economic Behavior"* von Neumann e Morgenstern (1944)

La *teoria dei giochi* è una disciplina che si occupa di situazioni di interazione strategica fra decisori.

Giochi cooperativi VS Giochi non cooperativi

Giochi cooperativi ad utilità trasferibile (TU-games)

Un *TU-game* è una coppia  $(N, v)$ :

- $N$  insieme finito di giocatori
- $v : \mathcal{P}(N) \rightarrow \mathbb{R}$  *funzione caratteristica* del gioco, tale che  $v(\emptyset) = 0$ .

# Il valore Shapley e altri indice di potere

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

Applicazione a dati reali

Conclusioni

*Come convertire l'informazione su quanto ogni coalizione di giocatori è in grado di procurarsi in un attributo associato ad ogni giocatore?*

Il valore Shapley (Shapley, 1953)

$$\Phi_i(v) = \frac{1}{n!} \sum_{\sigma \in \Sigma_N} m_i^\sigma(v) \quad \text{per ogni } i \in N$$

dove  $m_i^\sigma(v)$  è il contributo marginale di  $i$  alla coalizione  $\{\sigma(1), \sigma(2), \dots, \sigma(j-1)\}$ , e  $j$  è tale che  $\sigma(j) = i$ .

- Shapley (1953): caratterizzazione assiomatica sulla classe dei TU-games.
- Shapley e Shubik (1954): valore Shapley come *indice di potere*.



*Come convertire l'informazione su quanto ogni coalizione di giocatori è in grado di procurarsi in un attributo associato ad ogni giocatore?*

## Il valore Shapley (Shapley, 1953)

$$\Phi_i(v) = \frac{1}{n!} \sum_{\sigma \in \Sigma_N} m_i^\sigma(v) \quad \text{per ogni } i \in N$$

dove  $m_i^\sigma(v)$  è il contributo marginale di  $i$  alla coalizione  $\{\sigma(1), \sigma(2), \dots, \sigma(j-1)\}$ , e  $j$  è tale che  $\sigma(j) = i$ .

- Shapley (1953): caratterizzazione assiomatica sulla classe dei TU-games.
- Shapley e Shubik (1954): valore Shapley come *indice di potere*.

Interpretazione di un indice di potere:

- misura dell'influenza di un giocatore sull'esito del gioco
- valutazione a priori della probabilità di un giocatore di giocare un "ruolo rilevante"

Regular semivalues

$$\Phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} p_i^S (v(S \cup \{i\}) - v(S)) \quad \forall i \in N$$

dove la distribuzione di probabilità  $\{p_i^S\}_S$  dipende solo dalla cardinalità di una coalizione ed è strettamente positiva.

- $p_i^S = \frac{s!(n-s-1)!}{n!} \Rightarrow$  valore Shapley
- $p_i^S = \frac{1}{2^{n-1}} \Rightarrow$  indice di Banzhaf

Interpretazione di un indice di potere:

- misura dell'influenza di un giocatore sull'esito del gioco
- valutazione a priori della probabilità di un giocatore di giocare un "ruolo rilevante"

## Regular semivalues

$$\Phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} p_i^s (v(S \cup \{i\}) - v(S)) \quad \forall i \in N$$

dove la distribuzione di probabilità  $\{p_i^s\}_s$  dipende solo dalla cardinalità di una coalizione ed è strettamente positiva.

- $p_i^s = \frac{s!(n-s-1)!}{n!} \Rightarrow$  valore Shapley
- $p_i^s = \frac{1}{2^{n-1}} \Rightarrow$  indice di Banzhaf

## Problema

Individuazione di geni che rivestono un ruolo importante all'interno di una condizione biologica, tipicamente l'insorgenza di una malattia.

## Moretti et al. (2010)

- *Association game*:  $(N, v)$ , dove per ogni  $S \subseteq N$ ,  $v(S)$  è il numero di *geni chiave* che interagiscono solo con geni in  $S$ , e tale che  $v(\emptyset) = 0$ .
- *Graph-restricted game*:  $(N, w_\Gamma^v)$ , per ogni  $S \subseteq N$ ,  $v(S)$  è pari alla somma dei valori assegnati dal gioco  $v$  alle componenti connesse del network ristretto alla coalizione  $S$ .

## Misura di centralità $\gamma$

$$\gamma_i(v, \Gamma) = \phi_i(w_\Gamma^v) - \phi_i(v) \quad \text{per ogni } i \in N$$

## Problema

Individuazione di geni che rivestono un ruolo importante all'interno di una condizione biologica, tipicamente l'insorgenza di una malattia.

## Moretti et al. (2010)

- *Association game*:  $(N, v)$ , dove per ogni  $S \subseteq N$ ,  $v(S)$  è il numero di *geni chiave* che interagiscono solo con geni in  $S$ , e tale che  $v(\emptyset) = 0$ .
- *Graph-restricted game*:  $(N, w_\Gamma^v)$ , per ogni  $S \subseteq N$ ,  $v(S)$  è pari alla somma dei valori assegnati dal gioco  $v$  alle componenti connesse del network ristretto alla coalizione  $S$ .

## Misura di centralità $\gamma$

$$\gamma_i(v, \Gamma) = \phi_i(w_\Gamma^v) - \phi_i(v) \quad \text{per ogni } i \in N$$

## Problema

Individuazione di geni che rivestono un ruolo importante all'interno di una condizione biologica, tipicamente l'insorgenza di una malattia.

## Moretti et al. (2010)

- *Association game*:  $(N, v)$ , dove per ogni  $S \subseteq N$ ,  $v(S)$  è il numero di *geni chiave* che interagiscono solo con geni in  $S$ , e tale che  $v(\emptyset) = 0$ .
- *Graph-restricted game*:  $(N, w_\Gamma^v)$ , per ogni  $S \subseteq N$ ,  $v(S)$  è pari alla somma dei valori assegnati dal gioco  $v$  alle componenti connesse del network ristretto alla coalizione  $S$ .

## Misura di centralità $\gamma$

$$\gamma_i(v, \Gamma) = \phi_i(w_\Gamma^v) - \phi_i(v) \quad \text{per ogni } i \in N$$

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

**Nuovi modelli di centralità in un network biologico**

Applicazione a dati reali

Conclusioni

## Problema

Come valutare l'importanza di geni in un network, **senza la conoscenza a priori di geni chiave** all'interno di tale network?

## Nuovi modelli di centralità

- definizione di un opportuno gioco di coalizione
- confronto tra due condizioni, ad esempio: campioni malati VS sani
- selezione dei geni con centralità positiva più alta: geni candidati a rivestire un ruolo centrale nella condizione di malattia
- selezione dei geni con centralità negativa più bassa: geni candidati a rivestire un ruolo centrale nella condizione di salute

## Problema

Come valutare l'importanza di geni in un network, **senza la conoscenza a priori di geni chiave** all'interno di tale network?

## Nuovi modelli di centralità

- definizione di un opportuno gioco di coalizione
- confronto tra due condizioni, ad esempio: campioni malati VS sani
- selezione dei geni con centralità positiva più alta: geni candidati a rivestire un ruolo centrale nella condizione di malattia
- selezione dei geni con centralità negativa più bassa: geni candidati a rivestire un ruolo centrale nella condizione di salute



# Nuovi modelli di centralità in un network biologico

A partire da dati di espressione genica, costruiamo un grafo pesato  $\langle N, E, w \rangle$ .

Modello 1 (somma dei pesi dei lati interni)

- $v_1(S) = \sum_{e \in E_S} w(e)$ , dove  $E_S = \{e = \{i, j\} \in E : i, j \in S\}$ .

Modello 2 (somma dei pesi dei lati adiacenti)

- $v_2(S) = \sum_{e \in E_N} w(e)$ , dove  $E_N = \{e = \{i, j\} \in E : i \in S\}$ .

Dato  $\langle N, E, w \rangle$ , costruiamo un grafo non pesato  $\langle N, \Gamma \rangle$ , tramite un *cutoff* sui pesi dei lati.

Modello 3 ( numero dei nodi adiacenti)

- $v_3(S) = |\text{fringe}(S)|$ , dove *fringe*(*S*) è l'insieme dei nodi di *S* e dei nodi raggiungibili da *S* con al massimo un salto.
- Gioco introdotto da Suri e Narahari (2008) nel contesto dei network di co-autorship per modellizzare il problema dei *top-k nodes*.
- In Aadithya et al. (2010) viene fornita una formula esatta per il valore Shapley di tale gioco.

# Nuovi modelli di centralità in un network biologico

A partire da dati di espressione genica, costruiamo un grafo pesato  $\langle N, E, w \rangle$ .

## Modello 1 (somma dei pesi dei lati interni)

- $v_1(S) = \sum_{e \in E_S} w(e)$ , dove  $E_S = \{e = \{i, j\} \in E : i, j \in S\}$ .

## Modello 2 (somma dei pesi dei lati adiacenti)

- $v_2(S) = \sum_{e \in E_N} w(e)$ , dove  $E_N = \{e = \{i, j\} \in E : i \in S\}$ .

Dato  $\langle N, E, w \rangle$ , costruiamo un grafo non pesato  $\langle N, \Gamma \rangle$ , tramite un *cutoff* sui pesi dei lati.

## Modello 3 (numero dei nodi adiacenti)

- $v_3(S) = |\text{fringe}(S)|$ , dove  $\text{fringe}(S)$  è l'insieme dei nodi di  $S$  e dei nodi raggiungibili da  $S$  con al massimo un salto.
- Gioco introdotto da Suri e Narahari (2008) nel contesto dei network di co-autorship per modellizzare il problema dei *top-k nodes*.
- In Aadithya et al. (2010) viene fornita una formula esatta per il valore Shapley di tale gioco.

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

Applicazione a dati reali

Conclusioni

# Nuovi modelli di centralità in un network biologico

A partire da dati di espressione genica, costruiamo un grafo pesato  $\langle N, E, w \rangle$ .

## Modello 1 (somma dei pesi dei lati interni)

- $v_1(S) = \sum_{e \in E_S} w(e)$ , dove  $E_S = \{e = \{i, j\} \in E : i, j \in S\}$ .

## Modello 2 (somma dei pesi dei lati adiacenti)

- $v_2(S) = \sum_{e \in E_N} w(e)$ , dove  $E_N = \{e = \{i, j\} \in E : i \in S\}$ .

Dato  $\langle N, E, w \rangle$ , costruiamo un grafo non pesato  $\langle N, \Gamma \rangle$ , tramite un *cutoff* sui pesi dei lati.

## Modello 3 ( numero dei nodi adiacenti)

- $v_3(S) = |\text{fringe}(S)|$ , dove  $\text{fringe}(S)$  è l'insieme dei nodi di  $S$  e dei nodi raggiungibili da  $S$  con al massimo un salto.
- Gioco introdotto da Suri e Narahari (2008) nel contesto dei network di co-autorship per modellizzare il problema dei *top-k nodes*.
- In Aadithya et al. (2010) viene fornita una formula esatta per il valore Shapley di tale gioco.

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

Applicazione a dati reali

Conclusioni

# Nuovi modelli di centralità in un network biologico

A partire da dati di espressione genica, costruiamo un grafo pesato  $\langle N, E, w \rangle$ .

## Modello 1 (somma dei pesi dei lati interni)

- $v_1(S) = \sum_{e \in E_S} w(e)$ , dove  $E_S = \{e = \{i, j\} \in E : i, j \in S\}$ .

## Modello 2 (somma dei pesi dei lati adiacenti)

- $v_2(S) = \sum_{e \in E_N} w(e)$ , dove  $E_N = \{e = \{i, j\} \in E : i \in S\}$ .

Dato  $\langle N, E, w \rangle$ , costruiamo un grafo non pesato  $\langle N, \Gamma \rangle$ , tramite un *cutoff* sui pesi dei lati.

## Modello 3 ( numero dei nodi adiacenti)

- $v_3(S) = |\text{fringe}(S)|$ , dove  $\text{fringe}(S)$  è l'insieme dei nodi di  $S$  e dei nodi raggiungibili da  $S$  con al massimo un salto.
- Gioco introdotto da Suri e Narahari (2008) nel contesto dei network di co-autorship per modellizzare il problema dei *top-k nodes*.
- In Aadithya et al. (2010) viene fornita una formula esatta per il valore Shapley di tale gioco.

# Nuovi modelli di centralità in un network biologico

A partire da dati di espressione genica, costruiamo un grafo pesato  $\langle N, E, w \rangle$ .

## Modello 1 (somma dei pesi dei lati interni)

- $v_1(S) = \sum_{e \in E_S} w(e)$ , dove  $E_S = \{e = \{i, j\} \in E : i, j \in S\}$ .

## Modello 2 (somma dei pesi dei lati adiacenti)

- $v_2(S) = \sum_{e \in E_N} w(e)$ , dove  $E_N = \{e = \{i, j\} \in E : i \in S\}$ .

Dato  $\langle N, E, w \rangle$ , costruiamo un grafo non pesato  $\langle N, \Gamma \rangle$ , tramite un *cutoff* sui pesi dei lati.

## Modello 3 (numero dei nodi adiacenti)

- $v_3(S) = |\text{fringe}(S)|$ , dove *fringe*(*S*) è l'insieme dei nodi di *S* e dei nodi raggiungibili da *S* con al massimo un salto.
- Gioco introdotto da Suri e Narahari (2008) nel contesto dei network di co-autorship per modellizzare il problema dei *top-k nodes*.
- In Aadithya et al. (2010) viene fornita una formula esatta per il valore Shapley di tale gioco.

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

Applicazione a dati reali

Conclusioni

## Modello 1

Una coalizione di geni è tanto più forte, quanto più i geni al suo interno sono correlati tra loro.

**Motivazione biologica:** se la correlazione tra i geni di una coalizione risulta significativamente diversa nelle due condizioni (ad esempio malati e sani), ciò potrebbe indicare un'implicazione dell'interazione tra tali geni nella regolazione della condizione biologica in esame

## Modello 3

Una coalizione è tanto più forte quanto più i geni al suo interno sono correlati agli altri geni del network.

La misura di centralità introdotta a partire da questo modello fornisce una misura di quanto un gene sia responsabile della 'connettività' del network.

## Modello 1

Una coalizione di geni è tanto più forte, quanto più i geni al suo interno sono correlati tra loro.

**Motivazione biologica:** se la correlazione tra i geni di una coalizione risulta significativamente diversa nelle due condizioni (ad esempio malati e sani), ciò potrebbe indicare un'implicazione dell'interazione tra tali geni nella regolazione della condizione biologica in esame

## Modello 3

Una coalizione è tanto più forte quanto più i geni al suo interno sono correlati agli altri geni del network.

La misura di centralità introdotta a partire da questo modello fornisce una misura di quanto un gene sia responsabile della 'connettività' del network.

## Teorema

Per ogni gene  $i$ , il valore Shapley del gioco  $(N, v_1)$  è dato da:

$$\phi_i(v_1) = \frac{1}{2} \sum_{j \in N \setminus \{i\}} w(i, j),$$

dove  $d_i(w) = \sum_{j \in N \setminus \{i\}} w(i, j)$  è la *degree centrality* per grafi pesati.

## Teorema

Per ogni gene  $i$ , l'indice di Banzhaf del gioco  $(N, v_1)$  coincide con il valore Shapley.

## Nuova interpretazione della degree centrality in chiave di gioco

Essa può essere interpretata come contributo marginale atteso nel gioco  $v_1$  rispetto alla distribuzione di probabilità sulla formazione delle coalizioni che dà origine al valore Shapley e all'indice di Banzhaf.



## Teorema

Per ogni gene  $i$ , il valore Shapley del gioco  $(N, v_1)$  è dato da:

$$\phi_i(v_1) = \frac{1}{2} \sum_{j \in N \setminus \{i\}} w(i, j),$$

dove  $d_i(w) = \sum_{j \in N \setminus \{i\}} w(i, j)$  è la *degree centrality* per grafi pesati.

## Teorema

Per ogni gene  $i$ , l'indice di Banzhaf del gioco  $(N, v_1)$  coincide con il valore Shapley.

## Nuova interpretazione della degree centrality in chiave di gioco

Essa può essere interpretata come contributo marginale atteso nel gioco  $v_1$  rispetto alla distribuzione di probabilità sulla formazione delle coalizioni che dà origine al valore Shapley e all'indice di Banzhaf.

## Teorema

Per ogni gene  $i$ , il valore Shapley del gioco  $(N, v_1)$  è dato da:

$$\phi_i(v_1) = \frac{1}{2} \sum_{j \in N \setminus \{i\}} w(i, j),$$

dove  $d_i(w) = \sum_{j \in N \setminus \{i\}} w(i, j)$  è la *degree centrality* per grafi pesati.

## Teorema

Per ogni gene  $i$ , l'indice di Banzhaf del gioco  $(N, v_1)$  coincide con il valore Shapley.

## Nuova interpretazione della degree centrality in chiave di gioco

Essa può essere interpretata come contributo marginale atteso nel gioco  $v_1$  rispetto alla distribuzione di probabilità sulla formazione delle coalizioni che dà origine al valore Shapley e all'indice di Banzhaf.

## Teorema

*Tutti gli indici di potere (regular semivalue), costruiti a partire dal gioco  $v_1$ , del tipo:*

$$\Phi_i(v_1) = \sum_{S \subseteq N \setminus \{i\}} p_s (v_1(S \cup \{i\}) - v_1(S))$$

*dove  $p_s$  è la probabilità che una coalizione di cardinalità  $s$  si formi, ordinano i nodi di  $N$  alla stessa maniera, e l'ordinamento coincide con quello determinato dalla degree centrality.*



## Conseguenza

*Nella scelta di un indice di rilevanza di geni in questo contesto, non dobbiamo fare ipotesi sulla probabilità di formazione delle coalizioni.*

## Teorema

*Tutti gli indici di potere (regular semivalue), costruiti a partire dal gioco  $v_1$ , del tipo:*

$$\Phi_i(v_1) = \sum_{S \subseteq N \setminus \{i\}} p_s (v_1(S \cup \{i\}) - v_1(S))$$

*dove  $p_s$  è la probabilità che una coalizione di cardinalità  $s$  si formi, ordinano i nodi di  $N$  alla stessa maniera, e l'ordinamento coincide con quello determinato dalla degree centrality.*



## Conseguenza

*Nella scelta di un indice di rilevanza di geni in questo contesto, non dobbiamo fare ipotesi sulla probabilità di formazione delle coalizioni.*

## Teorema

Per ogni gene  $i$ , il valore Shapley del gioco  $(N, v_3)$  è dato da:

$$\phi_i(v_3) = \sum_{j \in \text{fringe}(i)} \frac{1}{1 + \text{deg}(j)}$$

## Nuova misura di centralità $\mu$

$$\mu_i(v_3) = \phi_i(v_3^M) - \phi_i(v_3^S)$$

dove  $v_3^M$  è il gioco definito a partire dai campioni corrispondenti alla condizione in esame (ad esempio, la malattia) e  $v_3^S$  il gioco definito a partire dai campioni corrispondenti alla condizione di riferimento (i campioni sani).

Introduzione

Network e centralità: misure classiche

Teoria dei giochi

Teoria dei giochi e centralità in un network

Nuovi modelli di centralità in un network biologico

Applicazione a dati reali

Conclusioni

Applicazione del modello a dati di espressione genica pubblicati in van Leeuwen et al. (2008) in cui un'esperimento di microarray è condotto su campioni sanguigni di 47 bambini provenienti da due diverse regioni della Repubblica Ceca: Teplice e Prachaticice.

- Dataset di 20130 geni e 47 campioni.
- 23 campioni dalla regione di Teplice, 24 campioni dalla regione di Prachaticice.
- Selezione di 250 geni con maggiore correlazione con il gene TP53.

# Risultati: I 20 geni più centrali secondo la misura di centralità $\mu$

Gene symbol	Gene name	$\mu$
PDCD8 <sup>3</sup>	programmed cell death 8 (apoptosis-inducing factor)	1.4881
DBF4B <sup>3</sup>	DBF4 homolog B ( <i>S. cerevisiae</i> )	1.4252
ATP9A	ATPase, Class II, type 9A	1.3169
ACSL5	acyl-CoA synthetase long-chain family member 5	1.3056
OR2B2 <sup>1,2,3</sup>	olfactory receptor, family 2, subfamily B, member 2	1.2368
FLJ20035 <sup>2</sup>	hypothetical protein FLJ20035	1.1033
MRC2	mannose receptor, C type 2	1.1023
NPAS4	neuronal PAS domain protein 4	1.0693
MICAL-L2	MICAL-like 2	1.0550
MCCC1	methylcrotonoyl-Coenzyme A carboxylase 1 (alpha)	-0.8404
NAPA <sup>3</sup>	N-ethylmaleimide-sensitive factor attachment protein, alpha	-0.8638
SFPQ <sup>2,3</sup>	splicing factor proline/glutamine-rich (polypyrimidine tract binding protein associated)	-0.8974
ELA3A <sup>3</sup>	elastase 3A, pancreatic	-0.9023
NUDT22	nudix (nucleoside diphosphate linked moiety X)-type motif 22	-1.0509
LPXN	leupaxin	-1.0823
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	-1.1990
ZNF395 <sup>3</sup>	zinc finger protein 395	-1.4767
PECI <sup>3</sup>	peroxisomal D3,D2-enoyl-CoA isomerase	-1.5676

I numeri indicano i geni che risultano più centrali secondo la  $\gamma$  centrality (<sup>1</sup>), degree centrality (<sup>2</sup>), betweenness centrality (<sup>3</sup>) e closeness centrality (<sup>4</sup>).

- massimo overlap con betweenness centrality: 8 geni.
- il gene PDCD8, alias AIF (apoptosis-inducing factor), è il più centrale rispetto alla misura  $\mu$ : AIF è una proteina coinvolta nel *caspase-independent apoptotic pathway*, responsabile della degradazione di DNA in una cellula che induce la morte cellulare programmata (apoptosi).

AIF è stato osservato agire in relazione all'inquinamento dell'aria:

K. Andreau, M. Leroux, A. Bouharrou "Health and cellular impacts of air pollutants: from cytoprotection to cytotoxicity.", *Biochemistry Research International* (2012).

- geni in comune alla misura  $\gamma$ : OR2B2 e SFPQ.



## Sviluppi futuri:

### Dal punto di vista teorico

- caratterizzazione assiomatica tramite proprietà con interpretazione biologica

### Dal punto di vista computazionale

- indici più complessi: problematiche di tipo computazionale
- algoritmi di approssimazione/metodi euristici

### Dal punto di vista pratico/biologico (applicazione in sè)

- validazione del modello su altri dati
- validazione su network generati in maniera virtuale
- comparazione più dettagliata con altri indici di centralità
- indagare corrispondenza con betweenness centrality

## Sviluppi futuri:

### Dal punto di vista teorico

- caratterizzazione assiomatica tramite proprietà con interpretazione biologica

### Dal punto di vista computazionale

- indici più complessi: problematiche di tipo computazionale
- algoritmi di approssimazione/metodi euristici

### Dal punto di vista pratico/biologico (applicazione in sè)

- validazione del modello su altri dati
- validazione su network generati in maniera virtuale
- comparazione più dettagliata con altri indici di centralità
- indagare corrispondenza con betweenness centrality

## Sviluppi futuri:

### Dal punto di vista teorico

- caratterizzazione assiomatica tramite proprietà con interpretazione biologica

### Dal punto di vista computazionale

- indici più complessi: problematiche di tipo computazionale
- algoritmi di approssimazione/metodi euristici

### Dal punto di vista pratico/biologico (applicazione in sè)

- validazione del modello su altri dati
- validazione su network generati in maniera virtuale
- comparazione più dettagliata con altri indici di centralità
- indagare corrispondenza con betweenness centrality

Introduzione

Network e  
centralità:  
misure  
classiche

Teoria dei  
giochi

Teoria dei  
giochi e  
centralità in  
un network

Nuovi  
modelli di  
centralità in  
un network  
biologico

Applicazione  
a dati reali

Conclusioni

*Grazie per l'attenzione.*