

# Influential spreaders in multilayer networks

Supervisore  
Alberto Montresor

Laureando  
Gabriele Masina



Dipartimento di Ingegneria e Scienza dell'Informazione

Corso di Laurea in Informatica

Trento, Luglio 2020

# Le reti nel mondo reale

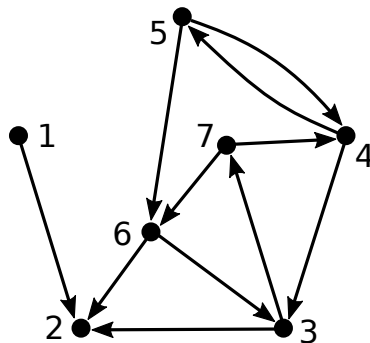
Molti sistemi reali possono essere rappresentati con una rete.

## Alcuni esempi

- 1 Reti sociali
- 2 Reti dei trasporti
- 3 Reti biologiche

## Problema

Un singolo **grafo** permette di rappresentare un unico tipo di relazione tra le entità



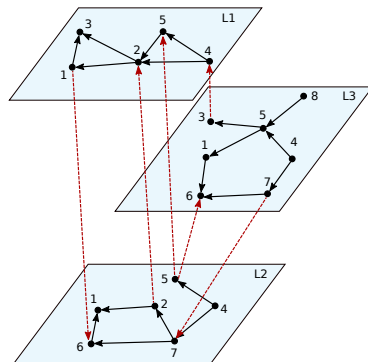
# Multilayer network

## Definizione

Una **multilayer network** è formata da un insieme di grafi detti **layer** ed un insieme di **inter-conessioni** tra nodi appartenenti a layer differenti

## Alcuni esempi

- 1 Profili online in diversi social network
- 2 Stazioni di reti dei trasporti differenti



# Influential spreaders

## Processi di diffusione

Lo studio degli **spreading process** in una rete permette di comprendere le dinamiche di diversi scenari

## Esempi

- 1 Fake news in social network
- 2 Notizie in reti di contatti
- 3 Ritardi tra le stazioni di una città

## Definizione

Gli **influential spreaders** sono quei nodi che in uno spreading process riescono a raggiungere la porzione di rete maggiore

# Algoritmi di centralità

## Obiettivo

Simulare gli spreading process è computazionalmente costoso. Si è cercato di individuare gli influential spreaders usando algoritmi di centralità

### Sul grafo aggregato

- PageRank
- $k$ -core
- Degree Centrality

### Separatamente su ogni layer

- Additive PageRank
- Sum  $k$ -core

### Sull'intera struttura multilayer

- Versatility PageRank
- Versatility Betweenness Centrality
- Multiplex  $k$ -core
- Minimal-layers PCI
- All-layers PCI
- Layer-symmetric PCI
- Layer-agnostic PCI

# Algoritmi di centralità

## Power Community Index

$mlPCI_n(v, l)$  è il massimo numero  $k$  tale che esistono almeno  $k$  vicini di  $v$  nel layer  $l$  con numero di vicini in almeno  $n$  layer maggiore o uguale a  $k$

### minimal-layers PCI

$$\sum_{n=1}^L mlPCI_n(v, l)$$

### all-layers PCI

$$n = L$$

### layer-symmetric PCI

$$n = k$$

### layer-agnostic PCI

L'informazione dei layer viene ignorata

# Dataset

## Dataset utilizzati

Sono state analizzate le prestazioni dei vari algoritmi su reti di vario tipo:

- Reti sociali
- Reti biologiche
- Reti generate a partire da reti applicazioni peer-to-peer

# Spreading process

## Probabilità di contagio

Un valore importante da impostare per riconoscere efficacemente gli influential spreaders è la **epidemic probability**  $\lambda$

## Critical epidemic point

Il **critical epidemic point** (c.e.p.) in un grafo è il valore  $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$

Sono state impostati:

- $\lambda_{ij}$  al c.e.p. del layer  $i$
- $\lambda_{ij}$  al c.e.p. del grafo aggregato

## Spreading process

Per ogni rete sono state fatte 100 simulazioni di spreading process partendo da ogni nodo



# Spreading process

## Probabilità di contagio

Un valore importante da impostare per riconoscere efficacemente gli influential spreaders è la **epidemic probability**  $\lambda$

## Critical epidemic point

Il **critical epidemic point** (c.e.p.) in un grafo è il valore  $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$

Sono state impostati:

- $\lambda_{ij}$  al c.e.p. del layer  $i$
- $\lambda_{ij}$  al c.e.p. del grafo aggregato

## Spreading process

Per ogni rete sono state fatte 100 simulazioni di spreading process partendo da ogni nodo

# Metodo di valutazione

## Confronto dei risultati

I vettori  $S$  (simulazioni) ed  $A$  (algoritmo) ottenuti sono stati confrontati usando il coefficiente **Kendall's Tau**:

$$\tau = \frac{n_c - n_d}{n(n-1)}$$

- $n_c$  è il numero di coppie *concordanti* nei due vettori ( $A_i > A_j$  e  $S_i > S_j$ ) oppure ( $A_i < A_j$  e  $S_i < S_j$ )
- $n_d$  è il numero di coppie *discordanti* nei due vettori ( $A_i > A_j$  e  $S_i < S_j$ ) oppure ( $A_i < A_j$  e  $S_i > S_j$ )
- Se  $A_i = A_j$  oppure  $S_i = S_j$  la coppia  $i, j$  non viene contata.

# Risultati

	DLN <sub>0.3, 0.3</sub>	DLN <sub>0.3, 0.8</sub>	DLN <sub>0.8, 0.3</sub>	DLN <sub>0.8, 0.8</sub>	DLN <sub>0.8, 0.3</sub>	DLN <sub>0.8, 0.8</sub>	DLN <sub>0.8, 0.3</sub>	DLN <sub>0.8, 0.8</sub>
addPR	0.4885	0.4844	0.4888	0.4860	0.4872	0.4874	0.4853	0.4848
aggCore	0.5575	0.5376	0.5552	0.5379	0.5596	0.5391	0.5578	0.5416
aggDeg	0.6620	0.6402	0.6607	0.6401	0.6274	0.6201	0.6297	0.6179
aggPR	0.5316	0.4958	0.5296	0.4972	0.5214	0.4879	0.5195	0.4869
alPCI	0.6945	0.7036	0.7043	0.7109	0.6565	0.6694	0.6776	0.6800
laPCI	0.7145	0.7157	0.7113	0.7158	0.6759	0.6910	0.6783	0.6868
lsPCI	0.6313	0.6285	0.6354	0.6335	0.6840	0.6848	0.6879	0.6895
mlPCI	0.7344	0.7406	0.7305	0.7427	0.6980	0.7114	0.7027	0.7095
multiCore	0.3551	0.3574	0.3554	0.3576	0.3551	0.3549	0.3553	0.3550
sumCore	0.5225	0.5350	0.5259	0.5346	0.5279	0.5350	0.5259	0.5334
verBC	0.6586	0.5605	0.6610	0.5639	0.6275	0.5404	0.6309	0.5441
verPR	0.5275	0.5141	0.5261	0.5136	0.5195	0.5078	0.5162	0.5064

Tabella: Kendall's Tau dei vari algoritmi in reti DLN

# Risultati

	$SLN_{0.3, 0.3}$	$SLN_{0.3, 0.8}$	$SLN_{0.8, 0.3}$	$SLN_{0.8, 0.8}$	$SLN_{0.8, 0.3}$	$SLN_{0.8, 0.8}$	$SLN_{0.8, 0.3}$	$SLN_{0.8, 0.8}$
addPR	0.4225	0.4293	0.4216	0.4295	0.4198	0.4309	0.4178	0.4287
aggCore	0.4066	0.3684	0.4158	0.3731	0.4010	0.3795	0.4117	0.3841
aggDeg	0.6073	0.5520	0.6009	0.5460	0.6147	0.5490	0.6134	0.5412
aggPR	0.4275	0.3619	0.4313	0.3695	0.4229	0.3713	0.4294	0.3779
alPCI	0.6550	0.6152	0.6379	0.5971	0.6675	0.6300	0.6555	0.6203
laPCI	0.6243	0.5761	0.6195	0.5687	0.6685	0.6208	0.6662	0.6092
lsPCI	0.5901	0.5681	0.5728	0.5479	0.5962	0.5678	0.5897	0.5584
mlPCI	0.6265	0.5810	0.6222	0.5771	0.6660	0.6169	0.6627	0.6057
multiCore	0.3481	0.3560	0.3498	0.3576	0.3506	0.3535	0.3472	0.3513
sumCore	0.4209	0.4288	0.4211	0.4273	0.4164	0.4291	0.4171	0.4267
verBC	0.5589	0.4528	0.5480	0.4506	0.5868	0.4842	0.5739	0.4796
verPR	0.3650	0.3479	0.3697	0.3579	0.3200	0.3297	0.3317	0.3374

Tabella: Kendall's Tau dei vari algoritmi in reti SLN

# Risultati

	Drosophila	Homo	MA2013	NYCM2014	SacchCere	SacchPomb
addPR	0.0398	0.3424	0.0059	-0.1313	0.3300	0.3410
aggCore	0.0646	0.4132	0.0572	-0.0453	0.4384	0.1046
aggDeg	0.7355	0.7096	0.5711	0.6150	0.6886	0.7656
aggPR	0.0417	0.3857	0.0164	-0.0771	0.3944	0.2562
alPCI	0.3682	0.1588	0.0493	0.1124	0.0455	0.4502
laPCI	0.6040	0.6859	0.5534	0.6178	0.6980	0.6853
lsPCI	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
mlPCI	0.6947	0.7191	0.5532	0.6028	0.7073	0.7729
multiCore	0.0497	0.1962	0.0428	0.0230	0.1925	-0.0170
sumCore	0.0658	0.4575	0.0616	-0.0487	0.4546	0.1000
verBC	0.5243	0.5653	0.1726	0.2045	0.5666	0.6550
verPR	-0.3505	-0.0606	-0.2707	-0.4107	0.0004	-0.4201

**Tabella:** Kendall's Tau dei vari algoritmi in reti multiplex

# Risultati

	Reti DLN										Reti DLN										Reti multiplex					
addPR	0.4885	0.4844	0.4888	0.4860	0.4872	0.4874	0.4853	0.4848	0.4225	0.4293	0.4216	0.4295	0.4198	0.4309	0.4178	0.4287	0.0398	0.3424	0.0059	-0.1313	0.3300	0.3410				
aggCore	0.5575	0.5376	0.5552	0.5379	0.5596	0.5391	0.5578	0.5416	0.4066	0.3684	0.4158	0.3731	0.4010	0.3795	0.4117	0.3841	0.0646	0.4132	0.0572	-0.0453	0.4384	0.1046				
aggDeg	0.6620	0.6402	0.6607	0.6401	0.6274	0.6201	0.6297	0.6179	0.6073	0.5520	0.6009	0.5460	0.6147	0.5490	0.6134	0.5412	0.7355	0.7096	0.5711	0.6150	0.6886	0.7656				
aggPR	0.5316	0.4958	0.5296	0.4972	0.5214	0.4879	0.5195	0.4869	0.4275	0.3619	0.4313	0.3695	0.4229	0.3713	0.4294	0.3779	0.0417	0.3857	0.0164	-0.0771	0.3944	0.2562				
alPCI	0.6945	0.7036	0.7043	0.7109	0.6565	0.6694	0.6776	0.6800	0.6550	0.6152	0.6379	0.5971	0.6675	0.6300	0.6555	0.6203	0.3682	0.1588	0.0493	0.1124	0.0455	0.4502				
laPCI	0.7145	0.7157	0.7113	0.7158	0.6759	0.6910	0.6783	0.6868	0.6243	0.5761	0.6195	0.5687	0.6685	0.6208	0.6662	0.6092	0.6040	0.6859	0.5534	0.6178	0.6980	0.6853				
lsPCI	0.6313	0.6285	0.6354	0.6335	0.6840	0.6848	0.6879	0.6895	0.5901	0.5681	0.5728	0.5479	0.5962	0.5678	0.5897	0.5584	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000				
mlPCI	0.7344	0.7406	0.7305	0.7427	0.6980	0.7114	0.7027	0.7095	0.6265	0.5810	0.6222	0.5771	0.6660	0.6169	0.6627	0.6057	0.6947	0.7191	0.5532	0.6028	0.7073	0.7729				
multiCore	0.3551	0.3574	0.3554	0.3576	0.3551	0.3549	0.3553	0.3550	0.3481	0.3560	0.3498	0.3576	0.3506	0.3535	0.3472	0.3513	0.0497	0.1962	0.0428	0.0230	0.1925	-0.0170				
sumCore	0.5225	0.5350	0.5259	0.5346	0.5279	0.5350	0.5259	0.5334	0.4209	0.4288	0.4211	0.4273	0.4164	0.4291	0.4171	0.4267	0.0658	0.4575	0.0616	-0.0487	0.4546	0.1000				
verBC	0.6586	0.5605	0.6610	0.5639	0.6275	0.5404	0.6309	0.5441	0.5589	0.4528	0.5480	0.4506	0.5868	0.4842	0.5739	0.4796	0.5243	0.5653	0.1726	0.2045	0.5666	0.6550				
verPR	0.5275	0.5141	0.5261	0.5136	0.5195	0.5078	0.5162	0.5064	0.3650	0.3479	0.3697	0.3579	0.3200	0.3297	0.3317	0.3374	-0.3505	-0.0606	-0.2707	-0.4107	0.0004	-0.4201				

## Osservazione

Due algoritmi hanno mostrato una buona capacità di riconoscere gli influential spreaders in tutte le reti analizzate:

- layer-agnostic Power Community Index
- minimal-layers Power Community Index

# Risultati

	Reti DLN										Reti DLN										Reti multiplex					
addPR	0.4885	0.4844	0.4888	0.4860	0.4872	0.4874	0.4853	0.4848			0.4225	0.4293	0.4216	0.4295	0.4198	0.4309	0.4178	0.4287			0.0398	0.3424	0.0059	-0.1313	0.3300	0.3410
aggCore	0.5575	0.5376	0.5552	0.5379	0.5596	0.5391	0.5578	0.5416			0.4066	0.3684	0.4158	0.3731	0.4010	0.3795	0.4117	0.3841			0.0646	0.4132	0.0572	-0.0453	0.4384	0.1046
aggDeg	0.6620	0.6402	0.6607	0.6401	0.6274	0.6201	0.6297	0.6179			0.6073	0.5520	0.6009	0.5460	0.6147	0.5490	0.6134	0.5412			0.7355	0.7096	0.5711	0.6150	0.6886	0.7656
aggPR	0.5316	0.4958	0.5296	0.4972	0.5214	0.4879	0.5195	0.4869			0.4275	0.3619	0.4313	0.3695	0.4229	0.3713	0.4294	0.3779			0.0417	0.3857	0.0164	-0.0771	0.3944	0.2562
alPCI	0.6945	0.7036	0.7043	0.7109	0.6565	0.6694	0.6776	0.6800			0.6550	0.6152	0.6379	0.5971	0.6675	0.6300	0.6555	0.6203			0.3682	0.1588	0.0493	0.1124	0.0455	0.4502
laPCI	0.7145	0.7157	0.7113	0.7158	0.6759	0.6910	0.6783	0.6868			0.6243	0.5761	0.6195	0.5687	0.6685	0.6208	0.6662	0.6092			0.6040	0.6859	0.5534	0.6178	0.6980	0.6853
lSPCI	0.6313	0.6285	0.6354	0.6335	0.6840	0.6848	0.6879	0.6895			0.5901	0.5681	0.5728	0.5479	0.5962	0.5678	0.5897	0.5584			0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
mlPCI	0.7344	0.7406	0.7305	0.7427	0.6980	0.7114	0.7027	0.7095			0.6265	0.5810	0.6222	0.5771	0.6660	0.6169	0.6627	0.6057			0.6947	0.7191	0.5532	0.6028	0.7073	0.7729
multiCore	0.3551	0.3574	0.3554	0.3576	0.3551	0.3549	0.3553	0.3550			0.3481	0.3560	0.3498	0.3576	0.3506	0.3535	0.3472	0.3513			0.0497	0.1962	0.0428	0.0230	0.1925	-0.0170
sumCore	0.5225	0.5350	0.5259	0.5346	0.5279	0.5350	0.5259	0.5334			0.4209	0.4288	0.4211	0.4273	0.4164	0.4291	0.4171	0.4267			0.0658	0.4575	0.0616	-0.0487	0.4546	0.1000
verBC	0.6586	0.5605	0.6610	0.5639	0.6275	0.5404	0.6309	0.5441			0.5589	0.4528	0.5480	0.4506	0.5868	0.4842	0.5739	0.4796			0.5243	0.5653	0.1726	0.2045	0.5666	0.6550
verPR	0.5275	0.5141	0.5261	0.5136	0.5195	0.5078	0.5162	0.5064			0.3650	0.3479	0.3697	0.3579	0.3200	0.3297	0.3317	0.3374			-0.3505	-0.0606	-0.2707	-0.4107	0.0004	-0.4201

## Osservazione

Due algoritmi hanno mostrato una buona capacità di riconoscere gli influential spreaders in tutte le reti analizzate:

- layer-agnostic Power Community Index
- minimal-layers Power Community Index

# Riepilogo

- Sono state selezionate diverse reti multilayer biologiche, sociali e di applicazioni peer-to-peer
- Per ognuna sono stati simulati degli spreading process per individuare i nodi più influenti nella diffusione
- Sono stati selezionati e testati diversi algoritmi di centralità
- Sono stati individuati due algoritmi in grado di riconoscere gli influential spreaders sulle reti del dataset



Grazie per l'attenzione