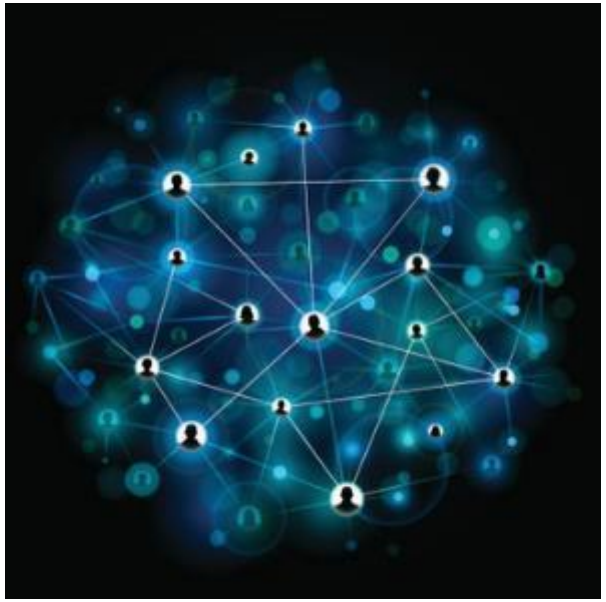
Interplay between Social Influence and Network Centrality: A Comparative Study of Shapley Centrality and Single-Node-Influence Centrality

Wei Chen
Microsoft Research Asia

Shang-Hua Teng
University of Southern California

Network Centrality: Key Concept in Network Science

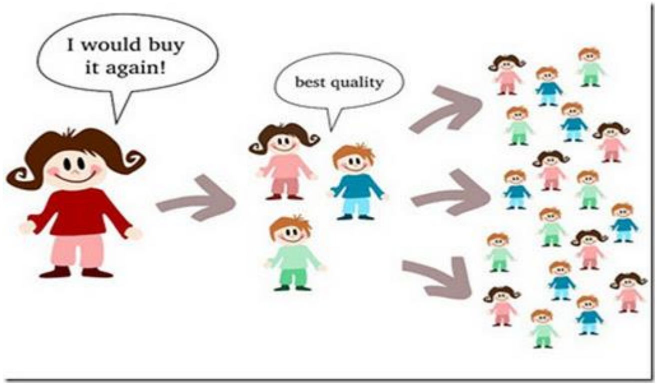
- Key question: who are at the central positions in a network?
- Classical Centrality Measures: Degree, Distance, Betweenness, Eigenvalue (PageRank)
- Issue: Only deal with static network structure, what about the effect of social interaction dynamics on network centrality?



Social Influence: Dynamics on Social Networks

- Social Influence: Social influence is everywhere
  - Adoptions if ideas, innovations, products, opinions
  - Conformity, social pressure, obedience
  - Influences are propagated in the network

- Questions:
  - How to incorporate social influence in centrality measure?
  - How to systematically study influence-based centrality measures?



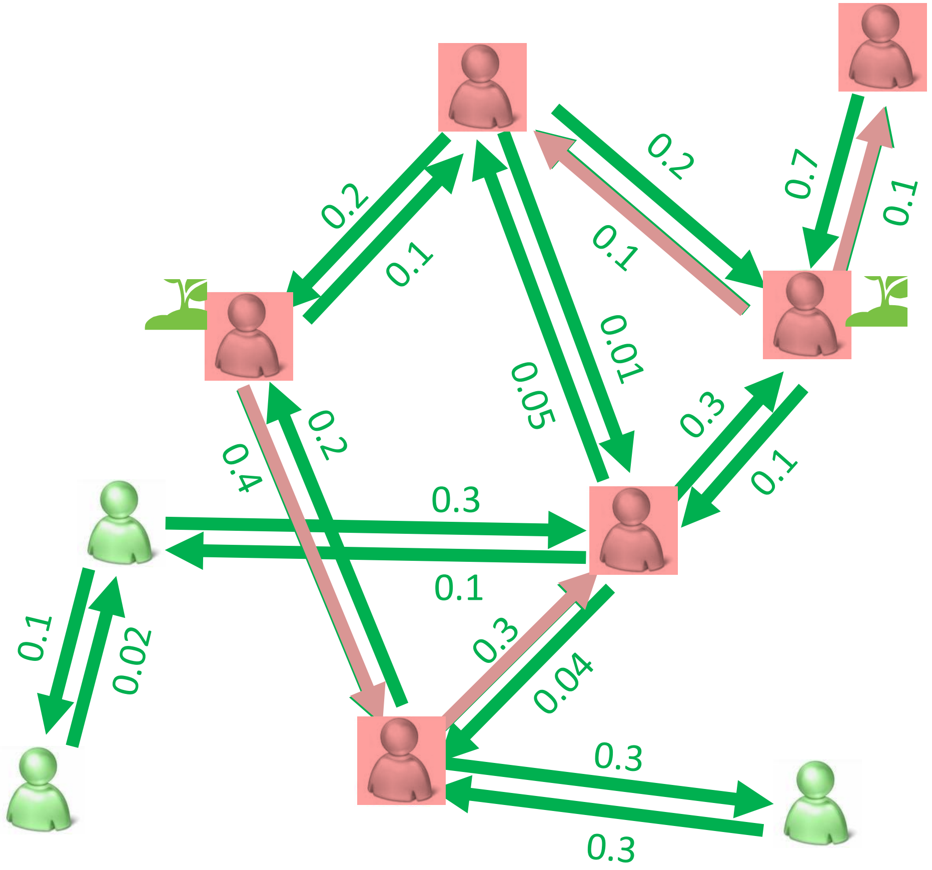
Our Approach

- Comparative study on two centrality measures:
  - Single-Node-Influence (SNI) centrality: since node's influence used as centrality
  - Shapley centrality: based on cooperative game theory, allocate total influence as credits/shares to nodes
- Axiomatic study: axiomatic characterization of both centralities
  - Provide the precise difference of the two centralities
- Algorithmic study: efficient algorithms for both centralities

Definitions of SNI and Shapley Centralities

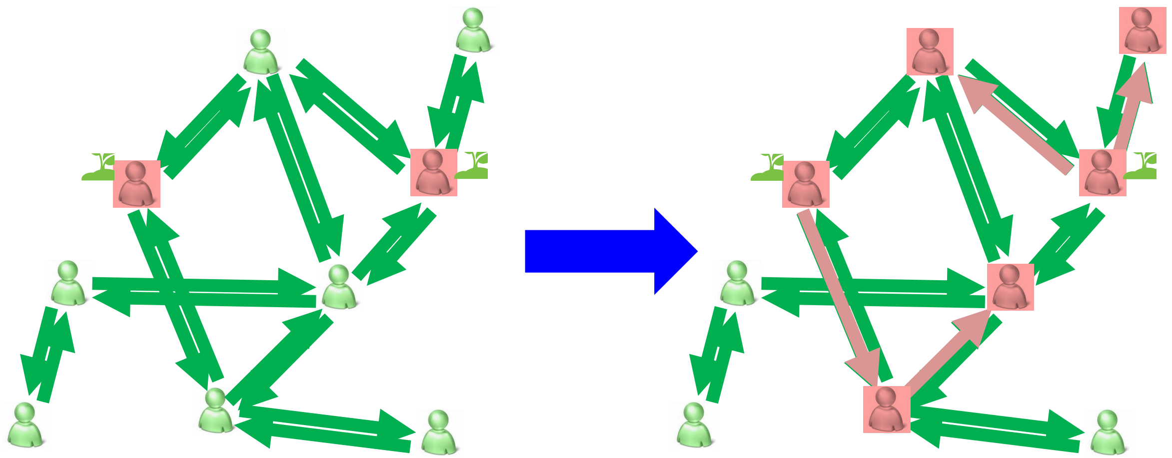
Stochastic Influence Propagation Models

- Model how influence stochastically propagate in a network, starting from a seed set
- Classical models: Independent Cascade (IC) Model, triggering model [Kempe, Kleinberg, Tardos '03]
  - No need to understand the mechanism for this talk
- Influence spread $\sigma(S)$: expected number of nodes activated
  - Measure the power of set S



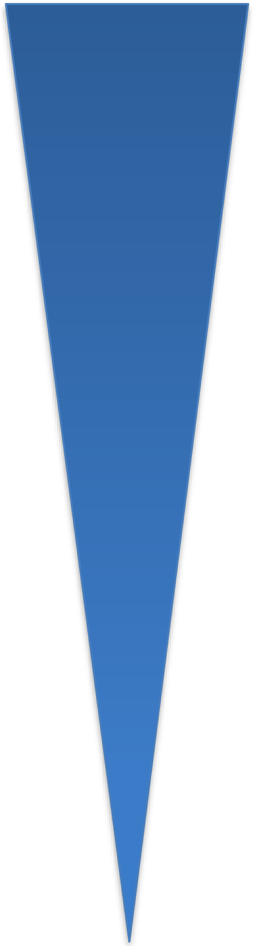
General Influence Instance

- Influence instance $I = (V, E, P_I)$
  - $P_I: 2^V \times 2^V \to [0,1]$
  - $P_I(S,T)$: probability that seed set S activates exact target set T
  - $S \subseteq T$
- Influence spread:
  - $\sigma(S) = \sum_{T \subseteq V} P_I(S,T) \cdot |T|$



Influence-based Centrality Measure

- Influence-based centrality measure ψ
  - $\psi: \{\mathcal{I}\} \to \mathbb{R}^n$
- Centrality measure as dimension reduction



Single-Node-Influence (SNI) Centrality

- Node v's SNI centrality is v's influence spread

$$\psi_v^{SNI}(\mathcal{I}) = \sigma_{\mathcal{I}}(\{v\})$$

- Natural and intuitive
- Measure node's power in isolation

Cooperative Game Theory and Shapley Value

- Measure individual power in group settings
- Cooperative game over $V = [n]$, with characteristic function $\tau: 2^V \to \mathbb{R}$
  - $\tau(S)$: cooperative utility of set S
- Shapley value $\phi: \{\tau\} \to \mathbb{R}^n$: $\phi_v(\tau) = \frac{1}{n!}\sum_{\pi\in\Pi}(\tau(S_{\pi,v} \cup \{v\}) - \tau(S_{\pi,v}))$
  - Π: set of permutations of V
  - $S_{\pi,v}$: subset of V ordered before v in permutation π
  - Average marginal utility on a random order
- Enjoy a unique axiomatic characterization

marginal utility

Shapley Centrality

- Node v's Shapley Centrality is the Shapley value of the influence spread function

$$\psi_v^{Shapley}(\mathcal{I}) = \phi_v(\sigma_{\mathcal{I}})$$

  – Treat influence spread function as a cooperative utility function
- Measure node's power in groups
- More precisely, node's marginal influence in a random order

Axiomatic Characterizations of Shapley and SNI Centralities

Why Axiomatization?

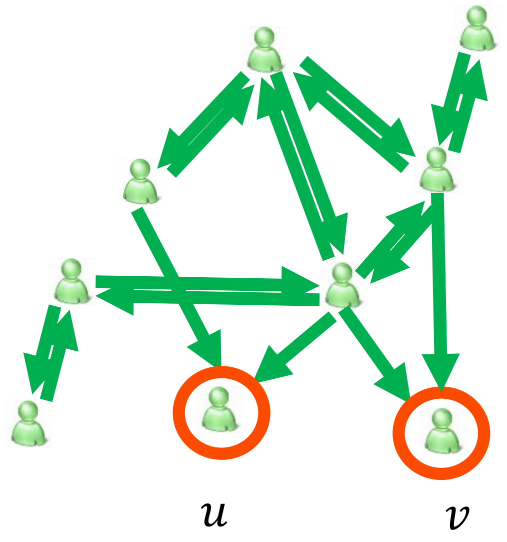
- Provide unique characterization of a centrality measure
- Know the determining factors of a centrality measure
- Axiomatic comparison among different centrality measures

Shapley Centrality: An Axiomatic Characterization

- Five axioms uniquely determining Shapley centrality
- Axiom 1 (Anonymity). Invariant under node id permutation
- Axiom 2 (Normalization). Sum of centrality measure is $\boldsymbol{n}$
  - For every instance $\mathcal{I}$, $\sum_{v \in V} \psi_v(\mathcal{I}) = n$
  - Average centrality measure per node is $\mathbf{1}$
  - A share division of the total influence spread

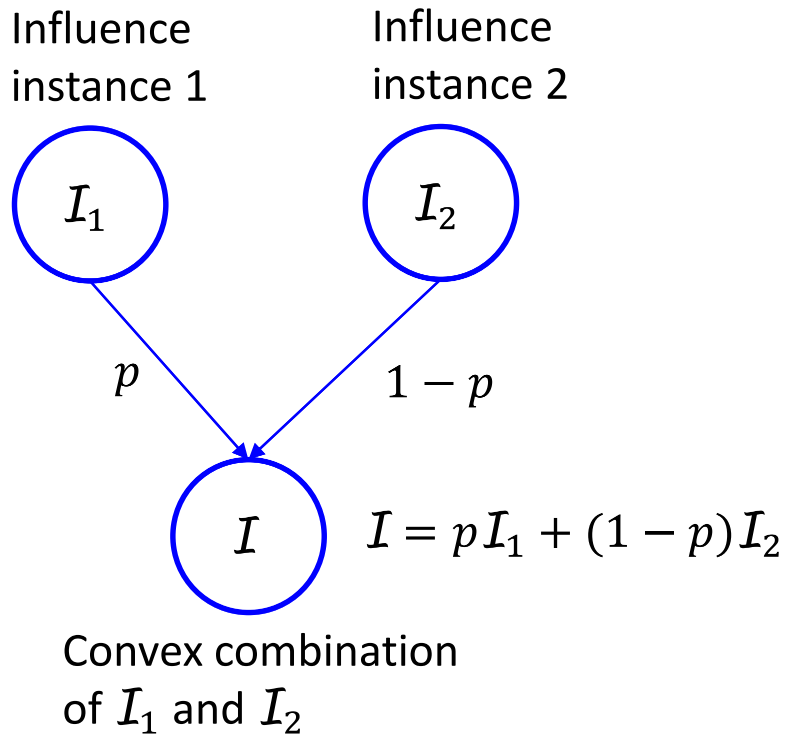
Axiom 3 (Independence of Sink Nodes)

- Axiom 3: Sink node projection preserves the centrality of other sink nodes
- Sink node
  - v is a sink node in $\mathcal{I}$, if $\forall S, T \subseteq V \setminus \{v\}$
    $$P_{\mathcal{I}}(S \cup \{v\}, T \cup \{v\}) = P_{\mathcal{I}}(S, T) + P_{\mathcal{I}}(S, T \cup \{v\})$$
  - Sink nodes have no influence to others, but others may influence sink nodes.
- Sink node projection: $\mathcal{I} \setminus \{v\} = (V \setminus \{v\}, E \setminus \{v\}, P_{\mathcal{I} \setminus \{v\}})$
  $$P_{\mathcal{I} \setminus \{v\}}(S, T) = P_{\mathcal{I}}(S, T) + P_{\mathcal{I}}(S, T \cup \{v\})$$
  - Equivalent to removing the sink node and its incident links in the triggering model



Axiom 4 (Bayesian)

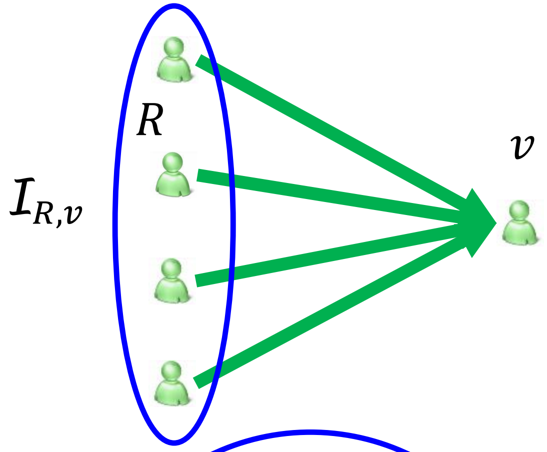
- Bayesian combination (convex combination) of influence instances gives convex combination of centrality measures.

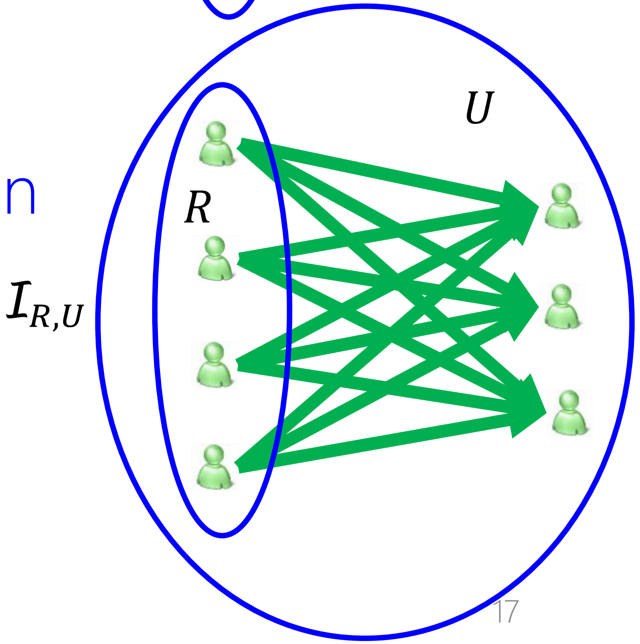


$$\psi(I) = p\psi(I_1) + (1-p)\psi(I_2)$$

Axiom 5 (Bargaining with Critical Sets)

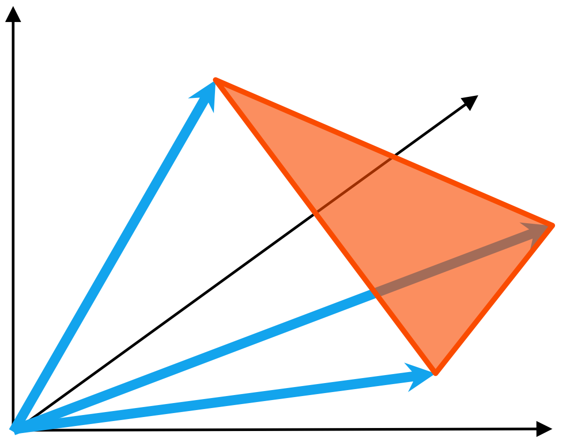
- r-vs-1 critical set instance $I_{R,v}$
  - Bipartite graph: set R vs. a sink node v; $|R| = r$
  - Set R together activates all nodes
  - Missing any one in R, generates no further influence
- The sink node in the r-vs-1 critical set instance $I_{R,v}$ has centrality $\frac{r}{r+1}$
  - Smaller than 1, because others can influence v
  - When R gets larger, getting close to 1, because coalition in R gets weaker
- Can be explained by Nash bargaining solution
- Extend to general critical set instance $I_{R,U}$





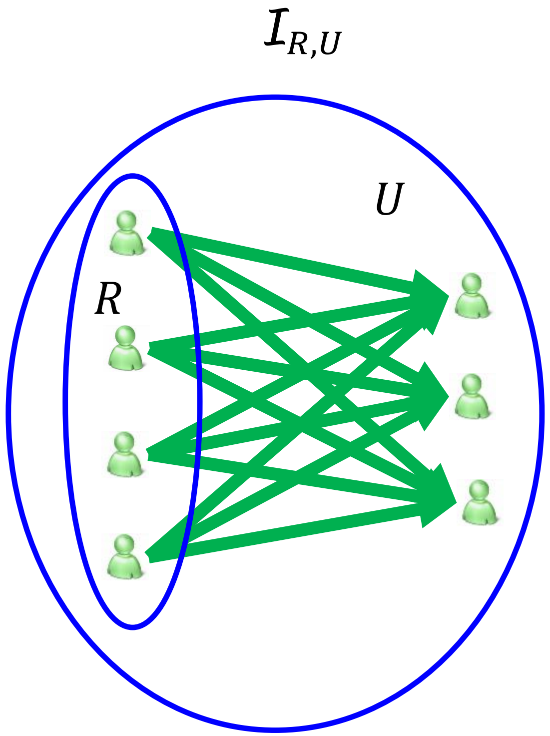
Characterization Theorem for Shapley Centrality

- Characterization Theorem: Shapley centrality is the unique centrality measure satisfying Axioms 1-5, and these axioms are independent.
- Proof sketch:
  - Use vector representation of influence instances
  - Find a set of instances (critical instances $\{\mathcal{I}_{R,U}\}$) as a set of basis for the vector space
  - Centrality of basis instances are uniquely determined by the axioms
  - Linearity of convex combination preserves uniqueness



Axiomatic Characterization of SNI Centrality

- Axiom 4 (Bayesian). Centrality measure of Bayesian influence instance respects the linearity-of-expectation principle
- Axiom 6 (Uniform Sink Nodes). Every sink node has centrality 1.
- Axiom 7 (Critical Nodes). In any critical instance $\mathcal{I}_{R,U}$, the centrality of a node in R is $\mathbf{1}$ if $|R| > 1$, and is $|U|$ if $|R| = 1$.
- Theorem: SNI centrality is the unique one satisfying Axioms 4, 6, 7, and these axioms are independent.



Comparison of Shapley and SNI Centrality

	SNI Centrality	Shapley Centrality
From definition	Focus on single node influence	Focus on influence in groups
On normalization	NO	YES, consider share division
On sink nodes	Treat them the same, only consider outgoing influence	Not the same, consider incoming influence
On critical nodes	Always 1 when $\|R\| > 1$, not good for threshold-like influence models	Always greater than 1, decreasing when $\|R\|$ increases, consider individual power in a group setting
Summary	Node influence power in isolation	Node irreplaceable power in group setting

Scalable Algorithm

Algorithmic Challenge

- Influence spread computation is #P-hard
- Shapley value definition involves factorial

Our Approach

- Based on the reverse reachable set (RR-set) approach for influence maximization [Borges et al'14, Tang et al'14, '15]
  - RR set $\boldsymbol{R}$: randomly select a node $\boldsymbol{v}$, reserve simulate diffusion (in the triggering model), the set of nodes reversely reachable from $\boldsymbol{v}$ is $\boldsymbol{R}$
  - Key property: $\sigma(S) = n \cdot \mathbb{E}_{\boldsymbol{R}}[\mathbb{I}\{S \cap \boldsymbol{R} \neq \emptyset\}]$
- For SNI: repeatedly sample RR sets, estimate influence spread of all nodes together ---$\psi_u^{SNI} = \sigma(\{u\}) = n \cdot \mathbb{E}_{\boldsymbol{R}}[\mathbb{I}\{u \in \boldsymbol{R}\}]$
- What about Shapley?
  - Key property for Shapley: $\psi_u^{Shapley} = n \cdot \mathbb{E}_{\boldsymbol{R}}[\mathbb{I}\{u \in \boldsymbol{R}\}/|\boldsymbol{R}|]$
  - Almost the same algorithmic structure as SNI

Our Result

- SNI and Shapley centrality share the same algorithmic structure
- Can approximate SNI and Shapley centralities with $\boldsymbol{\varepsilon}$ multiplicative error, with probability $\mathbf{1 - 1/n^{\ell}}$

$$\begin{cases} |\hat{\boldsymbol{\psi}}_v - \psi_v| \leq \varepsilon \psi_v & \forall v \in V \;\; with \; \psi_v > \psi^{(k)}, \\ |\hat{\boldsymbol{\psi}}_v - \psi_v| \leq \varepsilon \psi^{(k)} & \forall v \in V \;\; with \; \psi_v \leq \psi^{(k)}. \end{cases}$$

k-th largest centrality

- Running time: $O\left(\frac{1}{\varepsilon^2} \cdot \ell(m+n) \log n \cdot \frac{\mathbb{E}[\sigma(\tilde{v})]}{\psi^{(k)}}\right)$

Near linear time

Constant in many graphs

Conclusion and Future Work

- We provide dual axiomatic and algorithmic characterization
  - Axiomatically, exact characterization of SNI and Shapley centrality
  - Algorithmically, efficient computation for both using the same algorithmic structure
- Future work
  - SNI and Shapley centrality can be viewed as two end points in a spectrum, from node based centrality to group based centrality, what about others in the middle?
  - Extending traditional degree, distance, betweenness centralities etc. to influence based centralities?
  - More efficient algorithms?

Thank you, and questions?