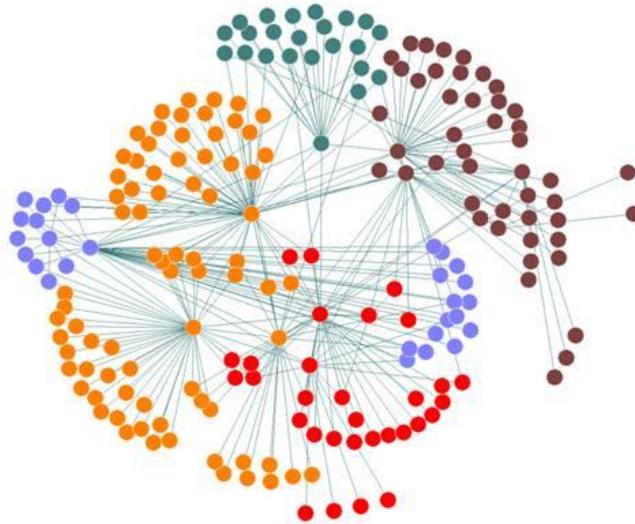




# Algorithms and Applications in Social Networks



2025/2026, Semester A

Slava Novgorodov

# Lesson #6

- Influence maximization:
  - Background, motivation and examples
- Linear Threshold Model
- Independent Cascade Model
- Theoretical properties

# Influence Maximization

# Motivation

- Advertisement – find most influential users and ask them to post an ad of your product
- Opinion making – find most influential users to spread the opinion
- Vaccination research – find people to vaccinate first

# Examples



bibars66 • Follow

La Zenia

bibars66 Laces are overrated. can't wait to play in these. #ACE16 #BeTheDifference



prognozistavki1 ++)

guymoyal\_ ביברס יאחחח

e\_dmitrichenko И ты в носках играть будешь?!

gal\_senderey6 בוא לכאר שבע חחחחחח

babkin\_offical Взаимная подписка и

babkin\_offical Я

tal\_ben\_naim מה הסיכוי שאתה חוזר... @bibras  
להפועל מתישהו? מתגעגעים אלייך

yungninjafresh В галошах будет)))

ri.cw +

mikaelrahkola @juhovoittola kato miten se on kirjottanu itte nimensä ✓problem solved



2,574 likes

JANUARY 28, 2016

# Examples



bibars66 • Follow

ממליץ בחום לכל ילד שחולם להיות שחקן.  
שחקן.  
#איתן עזריה  
👍👍👍👍👍👍

View all 35 comments

mpsmirnov Это что? Книга? Я что то не в курсе.

gal\_ginzburg פירלו הישראלי

spilberg\_ppp Нихуя не понимаю

m\_a\_goncharov @cska1909 🤔🤔 ara

sagialon8 !ביברס יתותח!

kondrratova ❤️❤️

jiblik Перевод : Советую каждому ребёнку, который хочет стать футболистом, прочесть эту книгу.

omerfridman127 פירלו של ישראל

Iguschina\_ @jiblik как называется книга??



2,384 likes

JULY 23, 2015

# Examples



**bibars66** • Follow  
La Quinta, Marbella

לא מוותר על זה בשום מקום  
בעולם.. #נס\_קפה\_של\_עלית 🙌

View all 33 comments

yoaveliaz כלפ אוהבים אותך

indiedi @rask0lnik0v אה סליחה שטעיתי

maor23235 EZ7DESIGNZ

✓FOLLOW@ez7designz✓

✓FOLLOW@ez7designz✓

✓FOLLOW@ez7designz✓

maor23235 EZ7DESIGNZ

✓FOLLOW@ez7designz✓

✓FOLLOW@ez7designz✓

✓FOLLOW@ez7designz✓!

maor23235 EZ7DESIGNZ

✓FOLLOW@ez7designz✓

✓FOLLOW@ez7designz✓

✓FOLLOW@ez7designz✓.

1.m.i.designz 1.M.I.DESIGNZ



1,524 likes

FEBRUARY 17, 2015

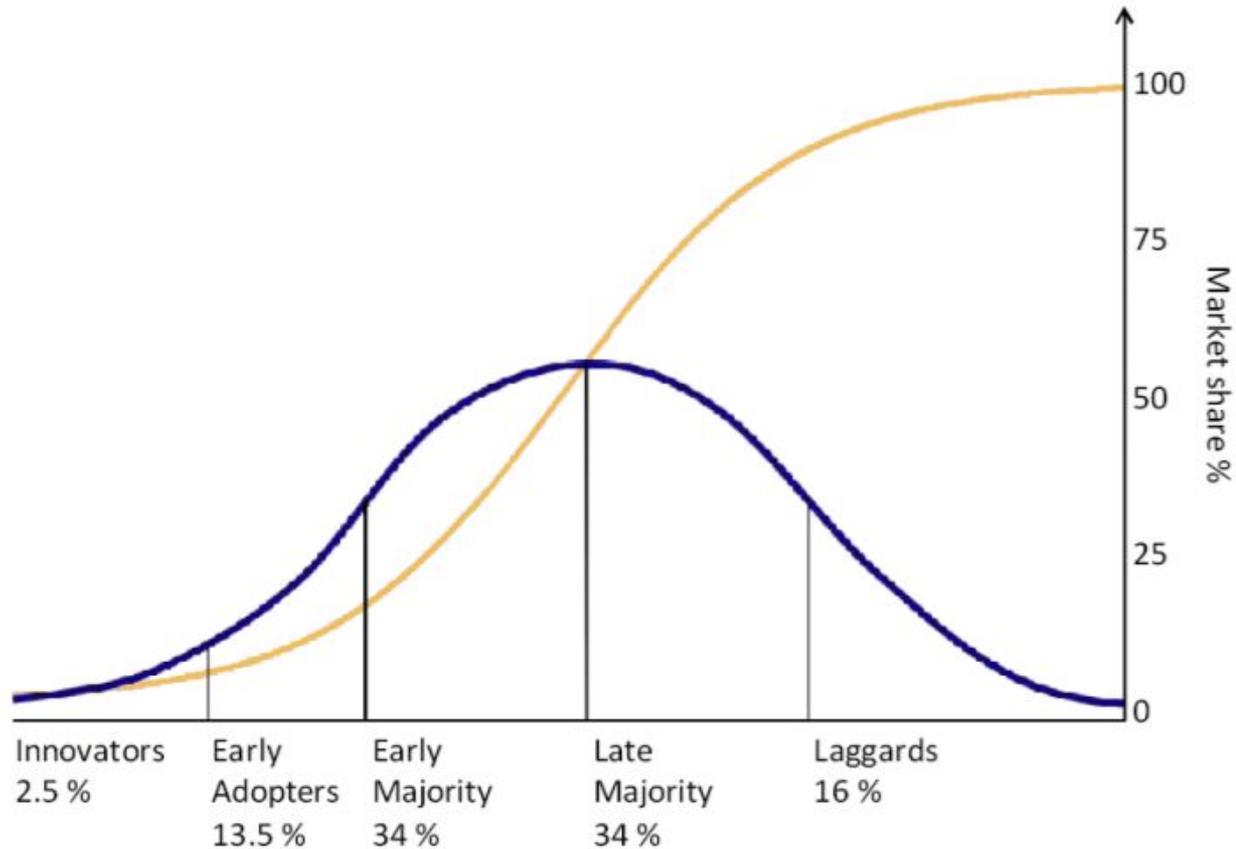
# Kate Middleton effect

- The **Kate Middleton effect** is the trend effect that she is reported to have had on others, for example in sales of particular products.
- According to NewsWeek:  
*"The Kate Effect may be worth £1 billion to the UK fashion industry"*

[https://en.wikipedia.org/wiki/Kate\\_Middleton\\_effect](https://en.wikipedia.org/wiki/Kate_Middleton_effect)



# Diffusion of innovation



# Marketing example: Hotmail

Jul 1996:	Hotmail.com started
Aug 1996:	20K subscribers
Dec 1996:	100K
Jan 1997:	1 million
Jul 1998:	12 million

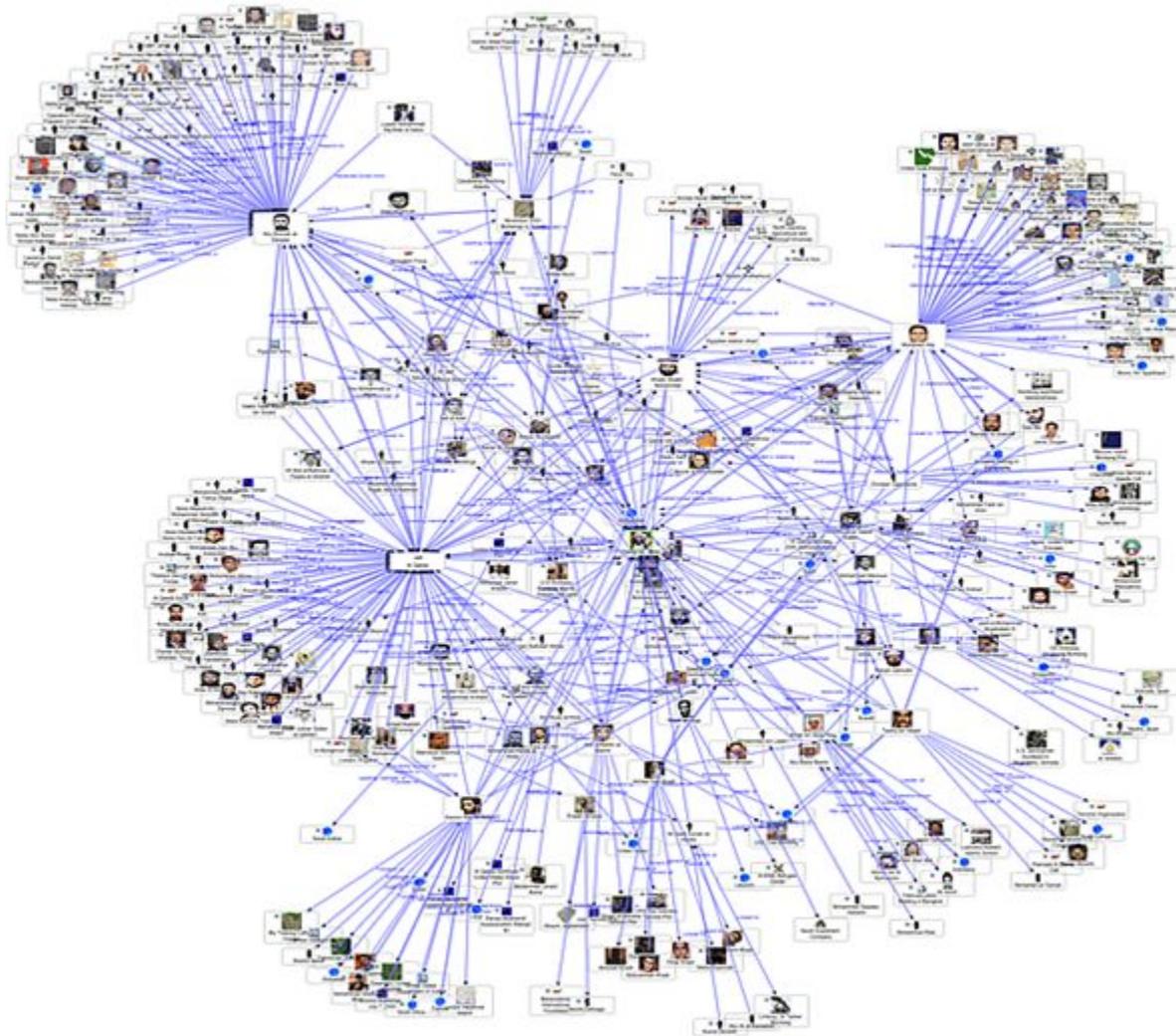


Bought by Microsoft for \$400 million

Marketing: At the end of each email sent there was a message to subscribe to Hotmail.com:

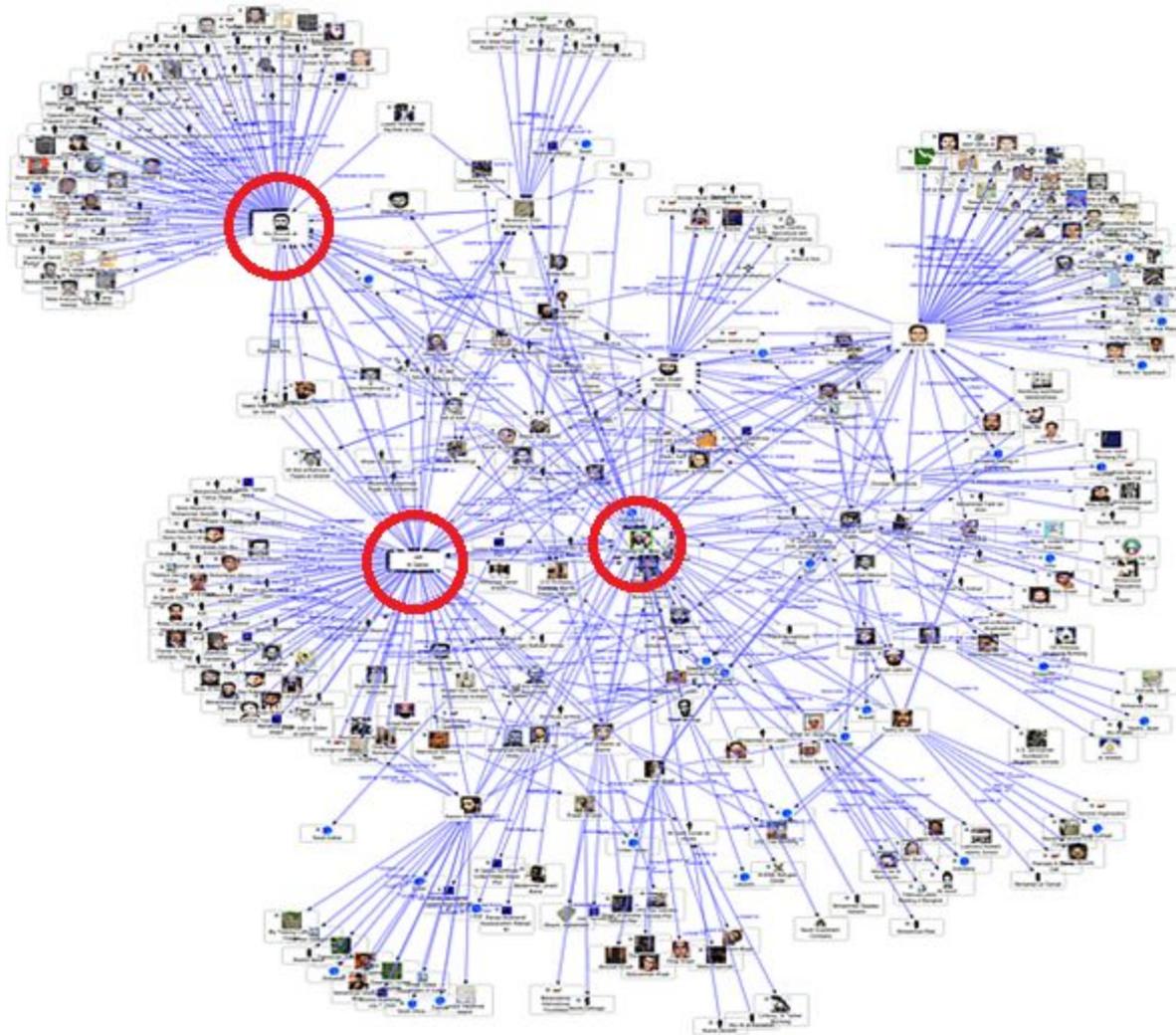
**"Get your free email at Hotmail"**

# Influence Maximization



Given a graph, find  $k$  people to maximize the number influenced of people

# Influence Maximization



Given a graph, find  $k$  people to maximize the number influenced of people

# Whom to take?



Instagram

Log In

Sign Up



b.netanyahu

Follow

Benjamin Netanyahu ראש ממשלת ישראל ויו"ר הליכוד  
[t.me/bnetanyahu](https://t.me/bnetanyahu)

1,264 posts

173k followers

3 following



Instagram

Log In

Sign Up



oren.hazan

Follow

ח"כ אורן חזן  
[www.facebook.com/orenhazanlikud](https://www.facebook.com/orenhazanlikud)

303 posts

6,620 followers

63 following



Instagram

Log In

Sign Up



tamarzandberg

Follow

Tamar Zandberg >> הצטרפו אליי, התפקדו היום למרצ  
[bit.ly/Zandberg](https://bit.ly/Zandberg)  
[tamarzandberg.co.il](https://tamarzandberg.co.il)

1,335 posts

2,844 followers

344 following

# Models of influence

- Two basic models:
  - Linear Threshold Model
  - Independent Cascade Model
- Setup:
  - A social network is represented as a directed weighted graph, with each person as a node
  - Nodes start either active or inactive
  - An active node may trigger activation of neighboring nodes
  - Monotonicity assumption: active nodes never deactivate

# Linear Threshold Model

# Linear Threshold Model

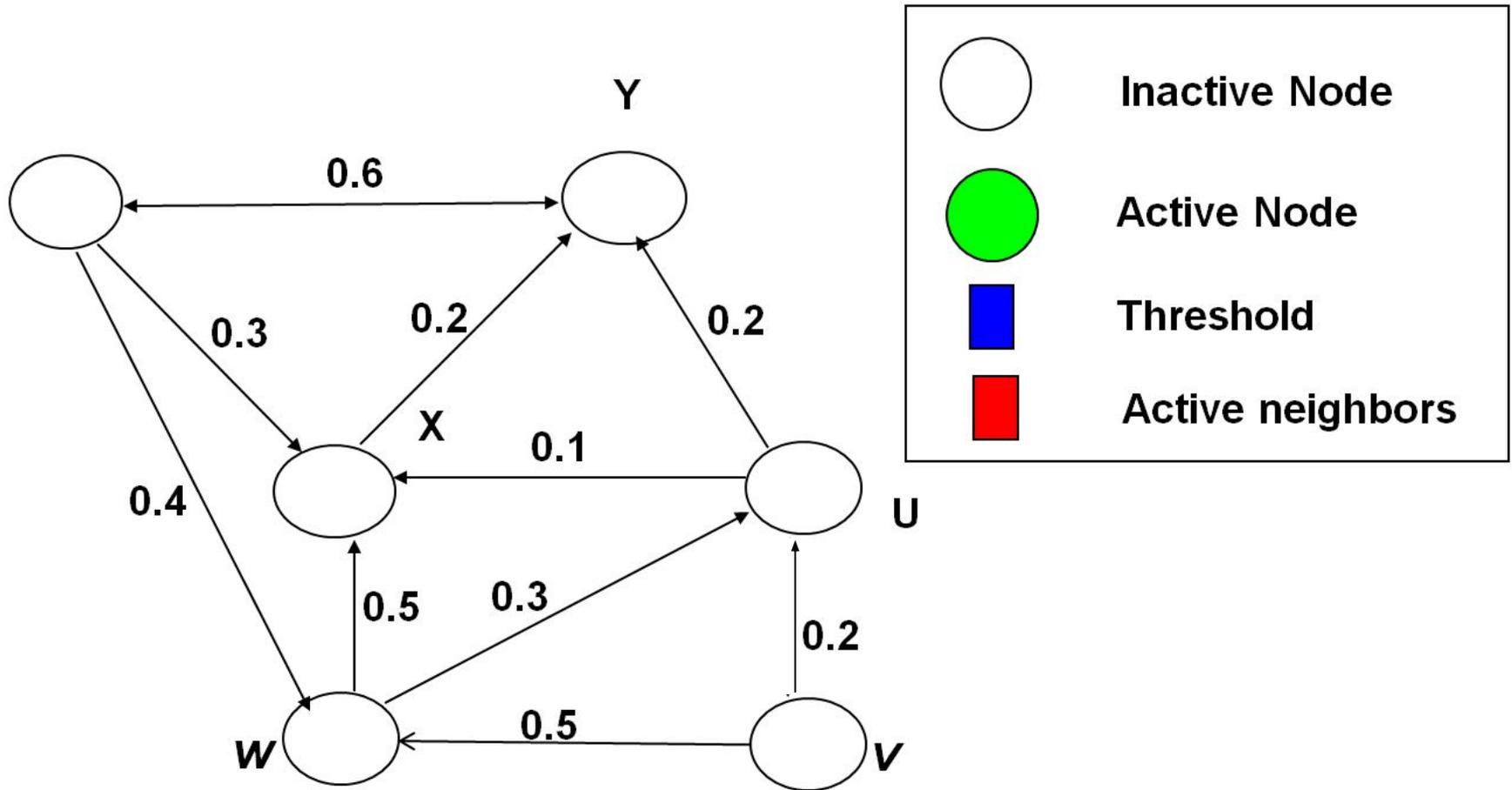
- A node  $v$  has random threshold  $\theta_v \sim U[0,1]$
- A node  $v$  is influenced by each neighbor  $w$  according to a *weight*  $b_{vw}$  such that

$$\sum_{w \text{ neighbor of } v} b_{v,w} \leq 1$$

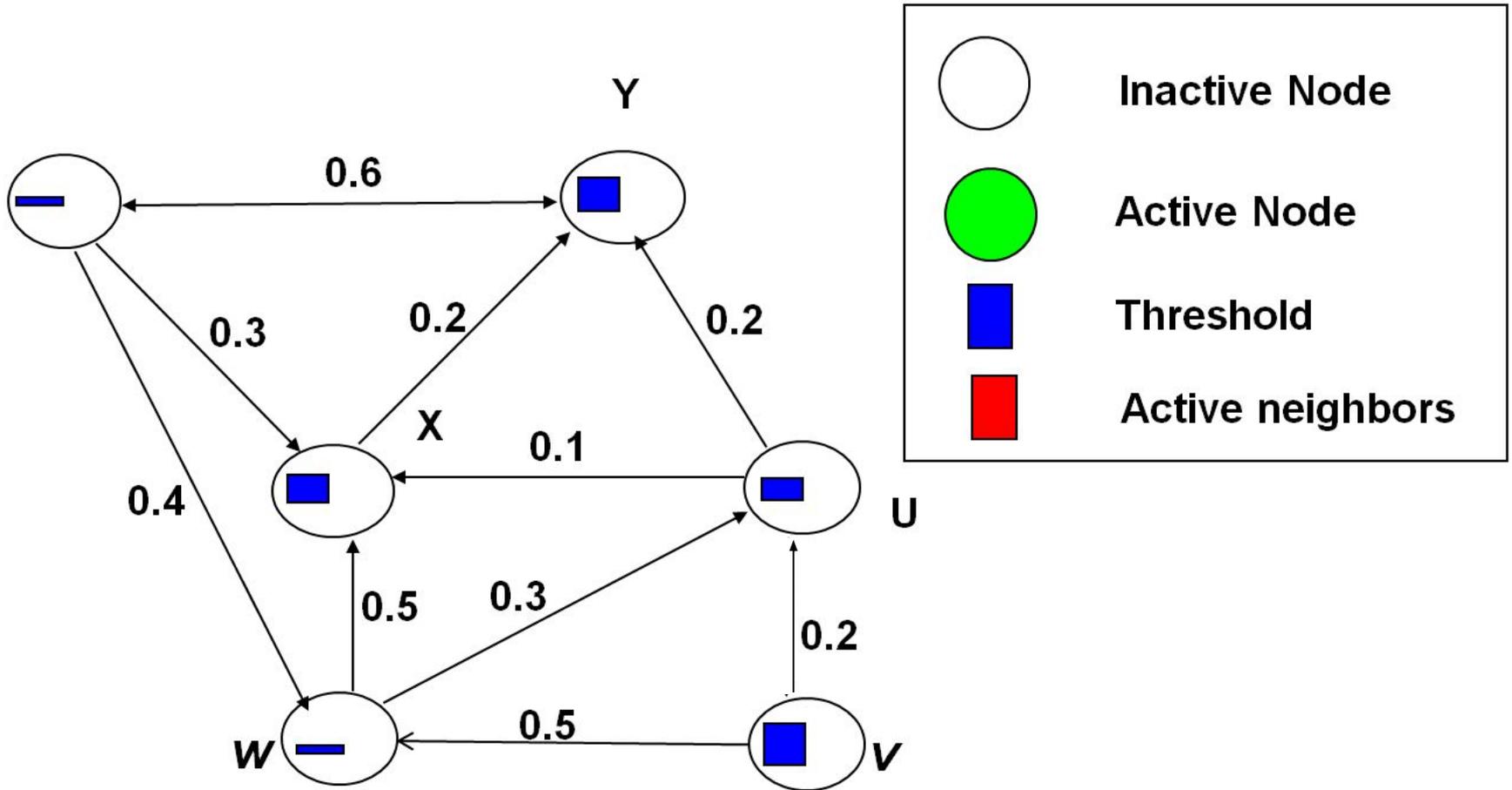
- A node  $v$  becomes active when at least (weighted)  $\theta_v$  fraction of its neighbors are active

$$\sum_{w \text{ active neighbor of } v} b_{v,w} \geq \theta_v$$

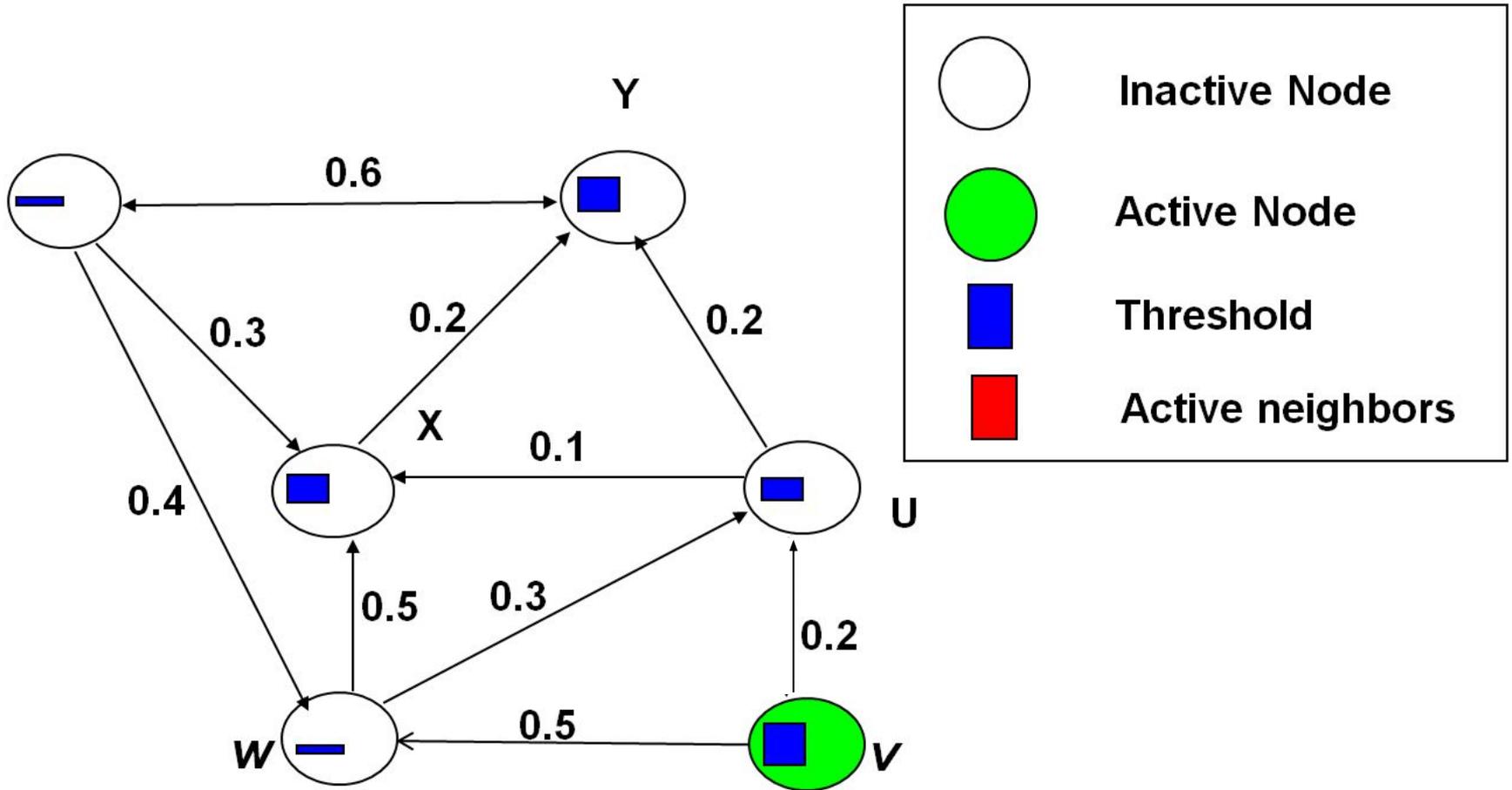
# Example



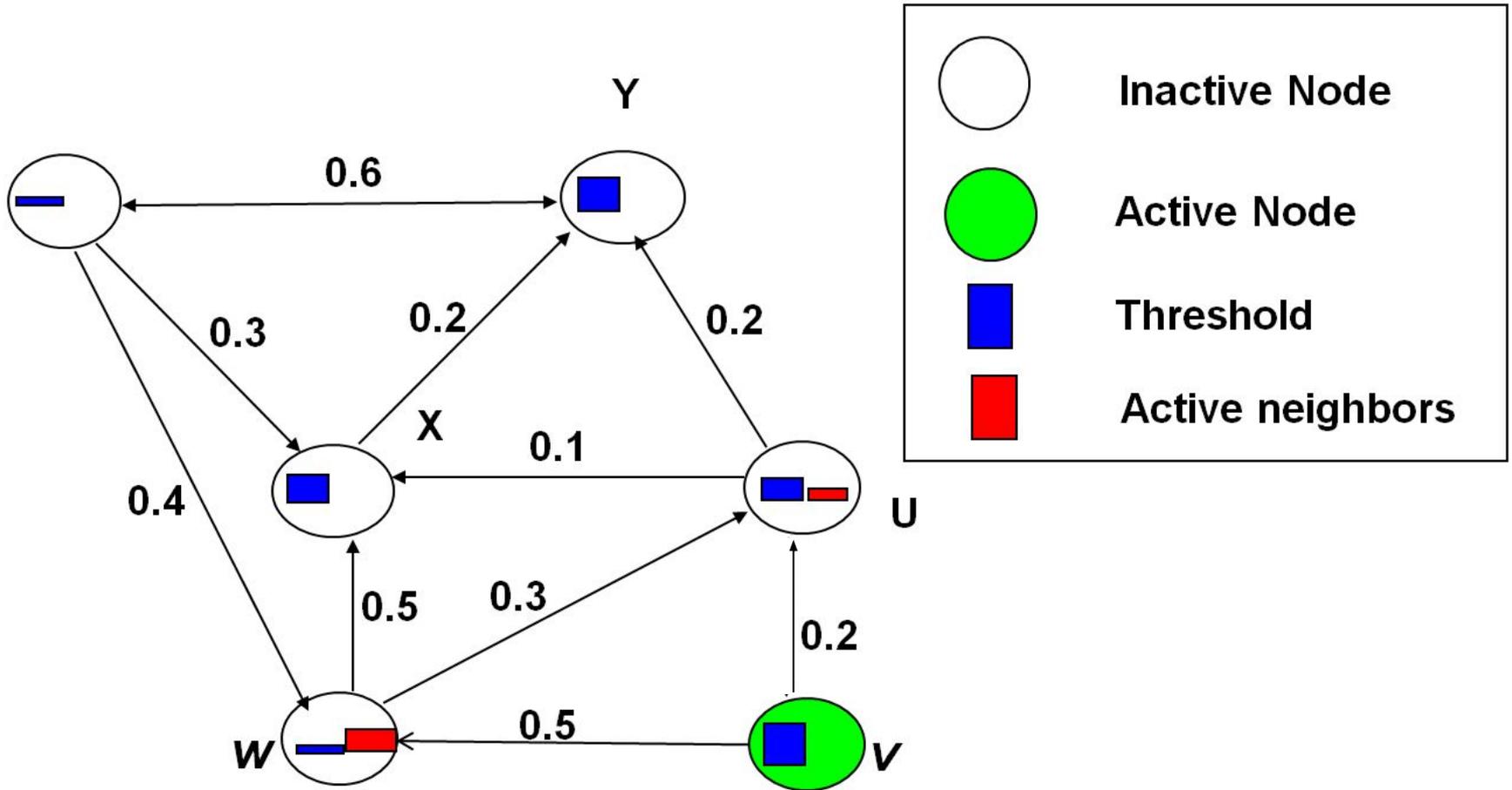
# Example



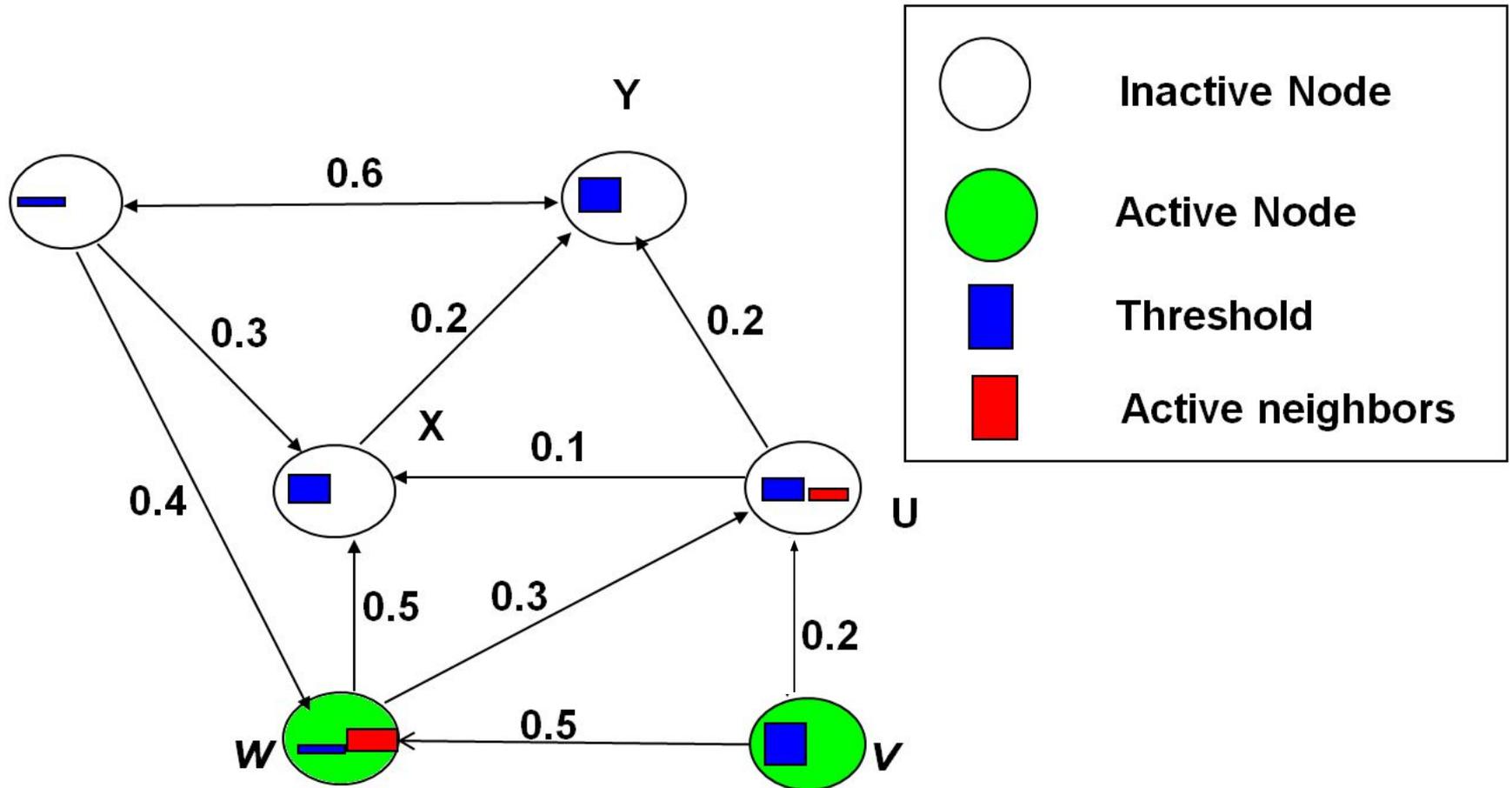
# Example



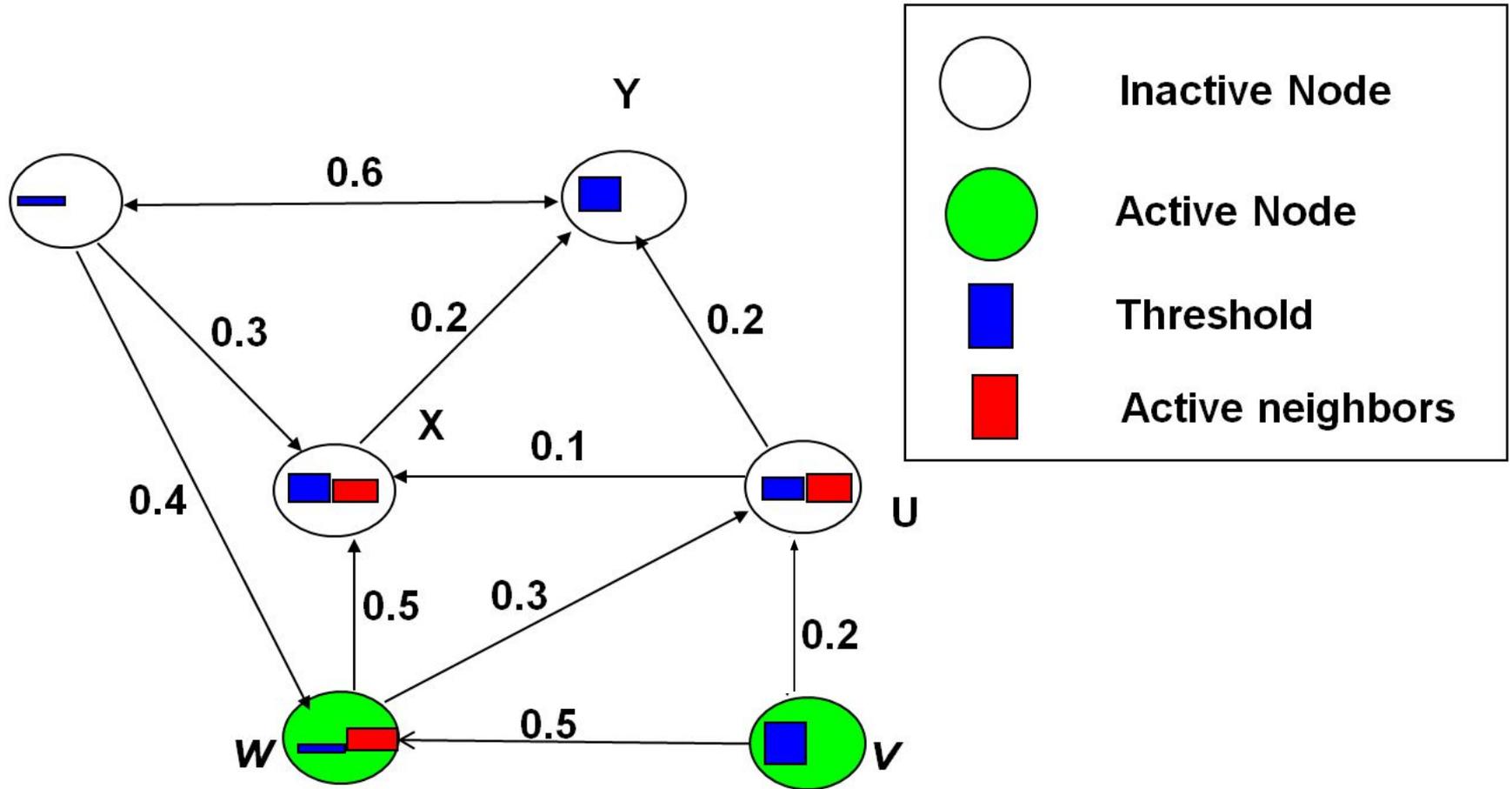
# Example



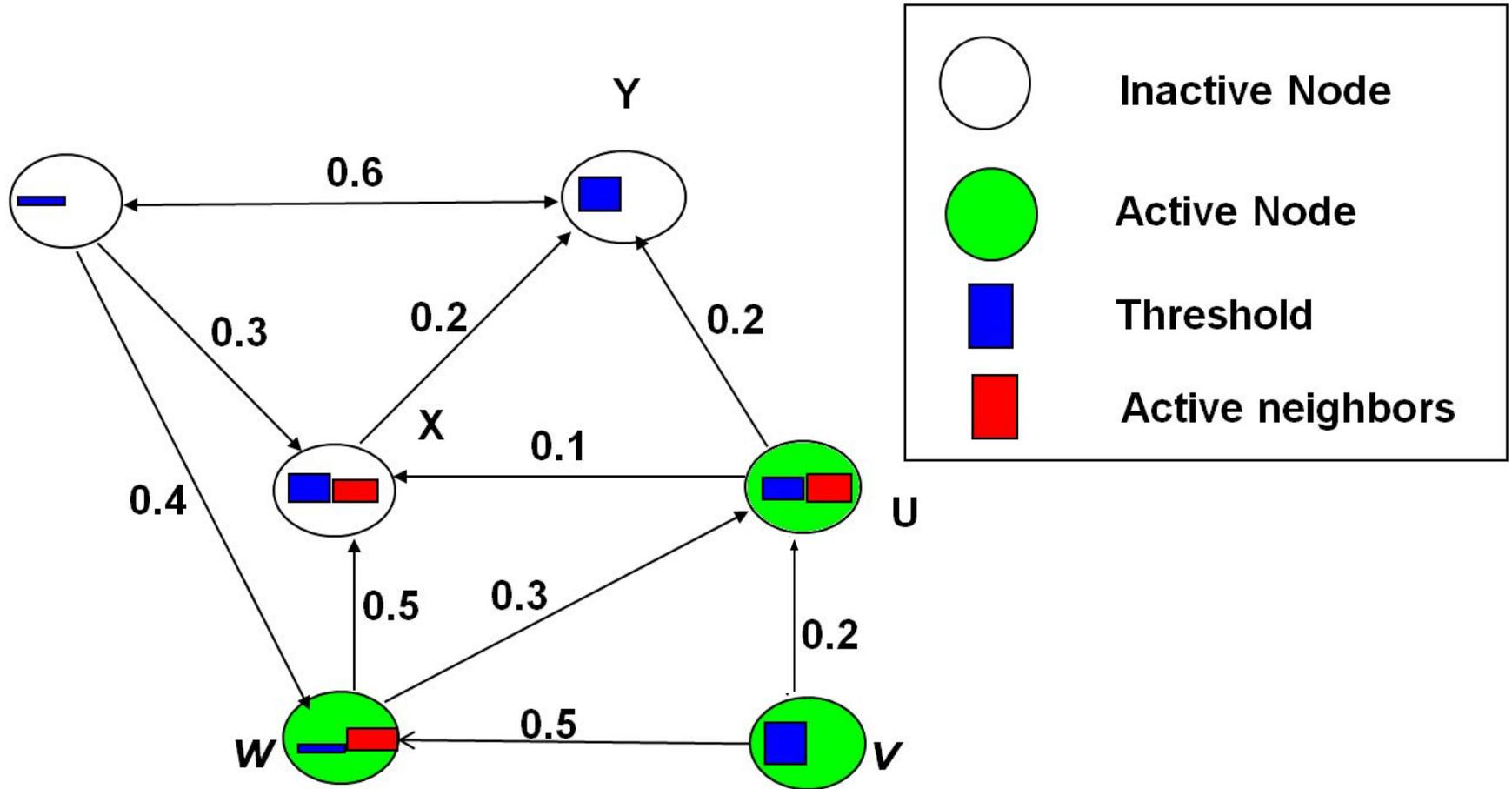
# Example



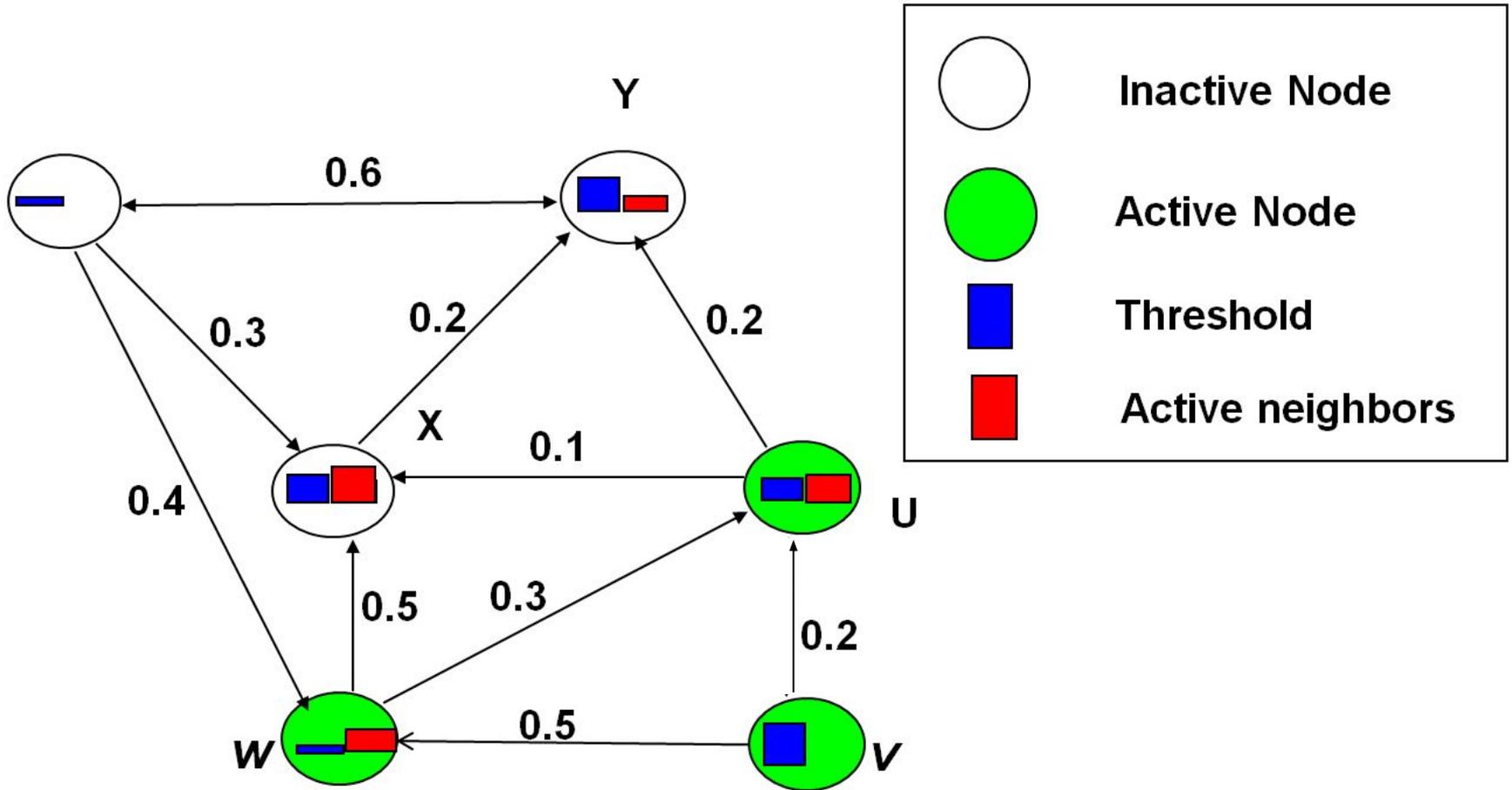
# Example



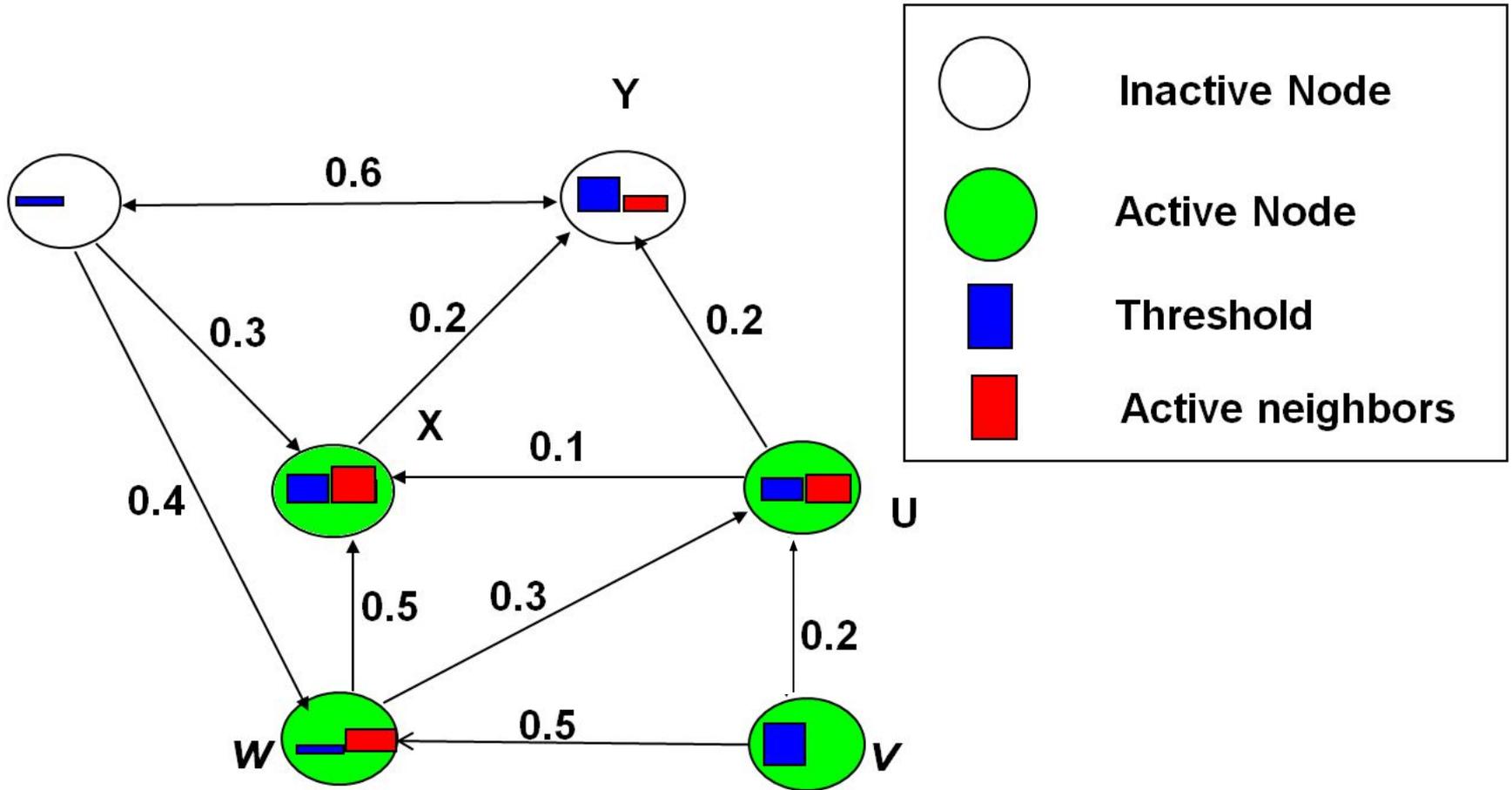
# Example



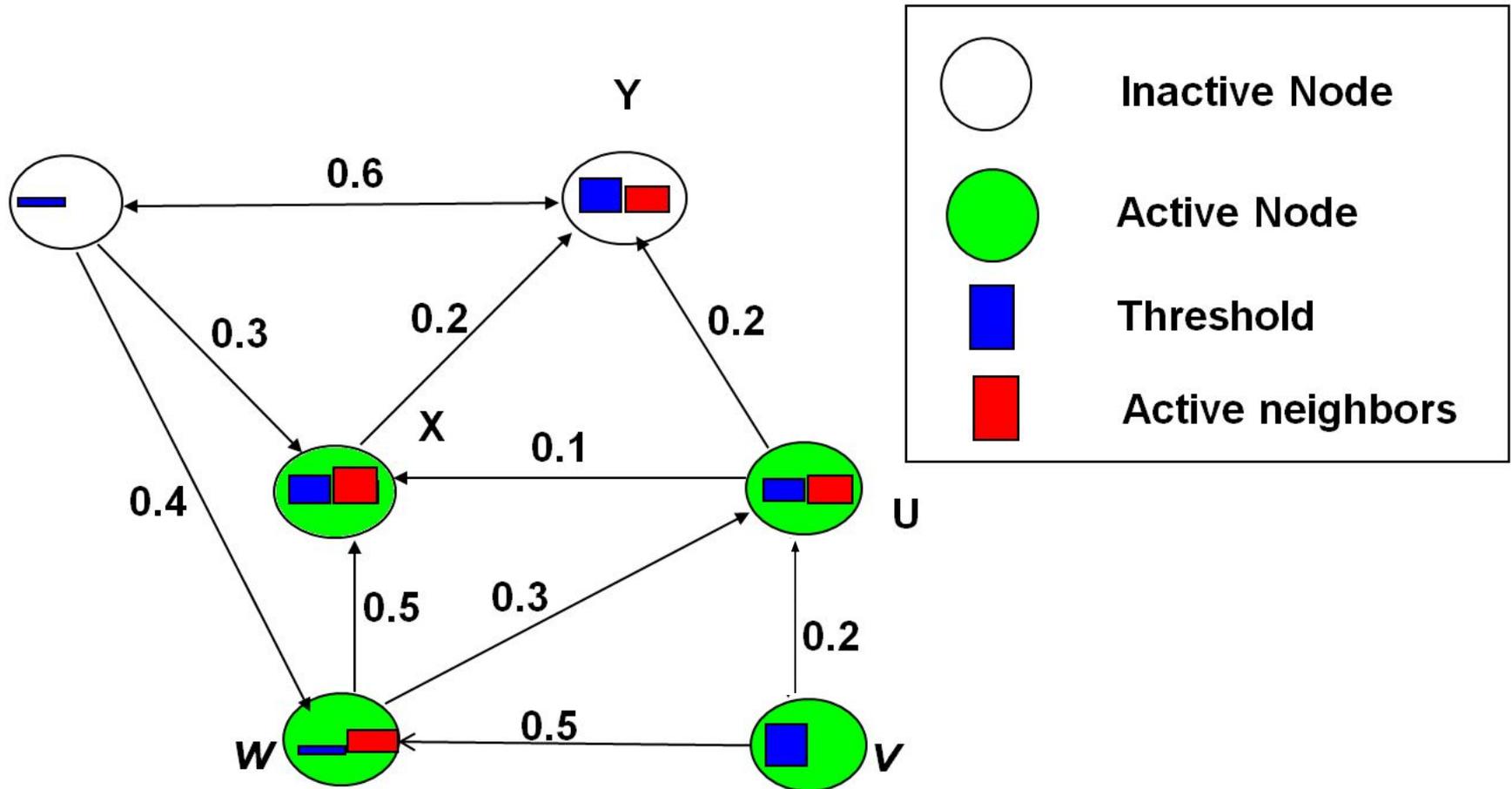
# Example



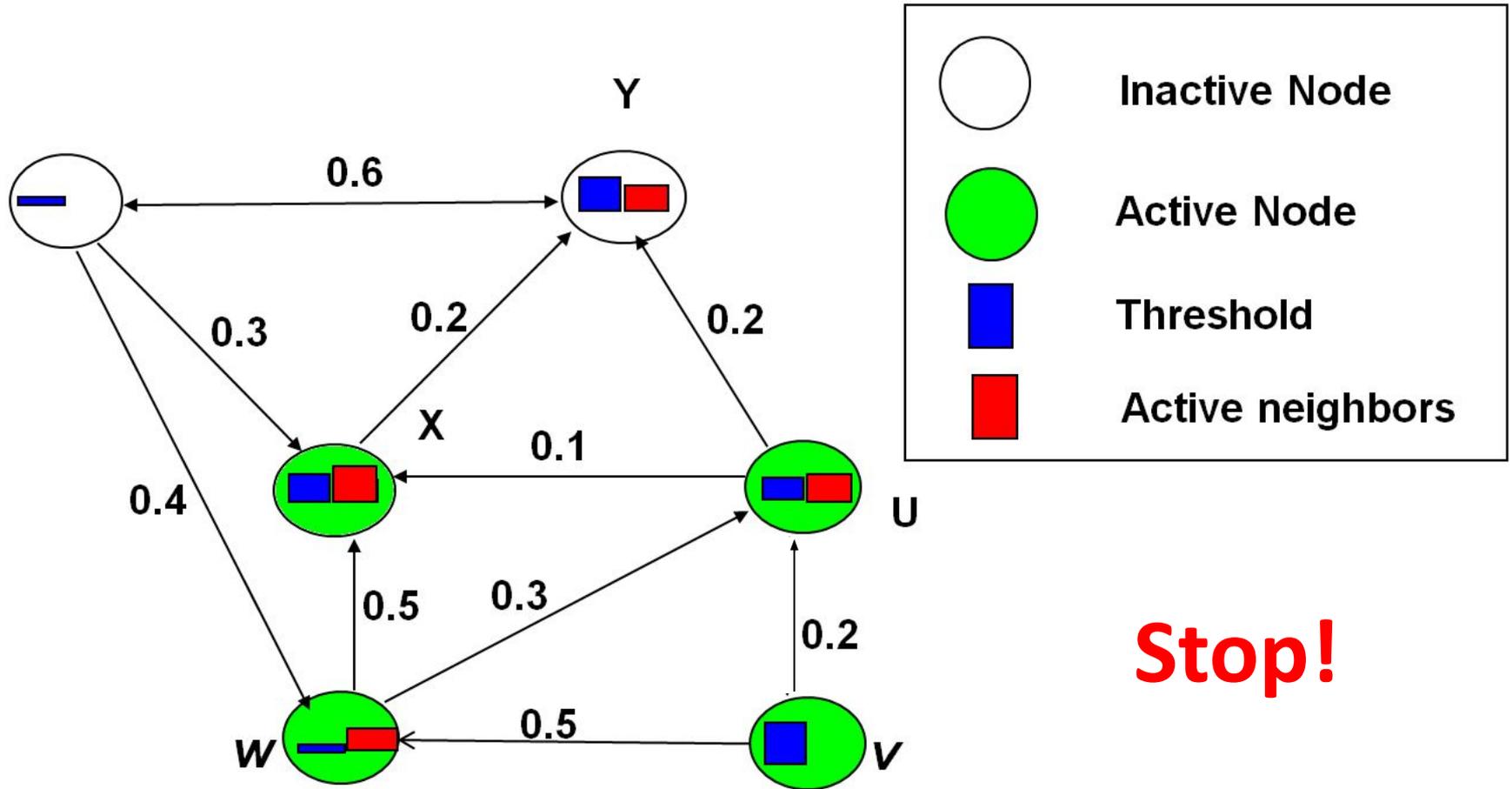
# Example



# Example



# Example



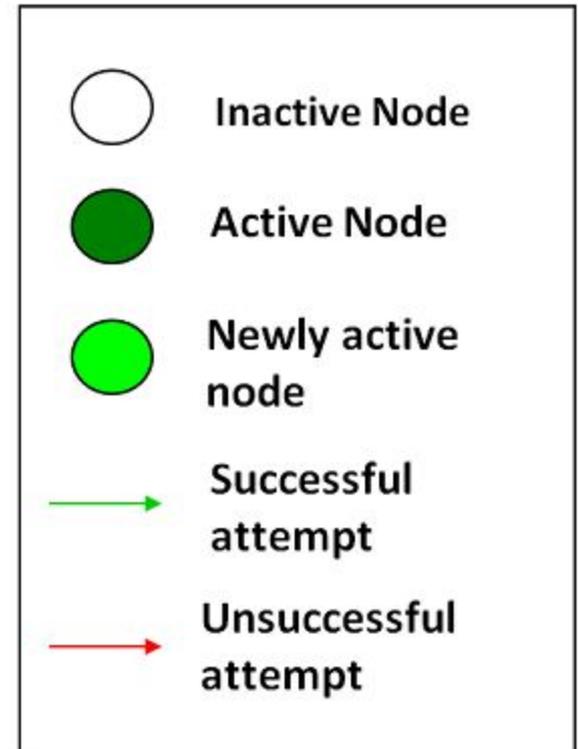
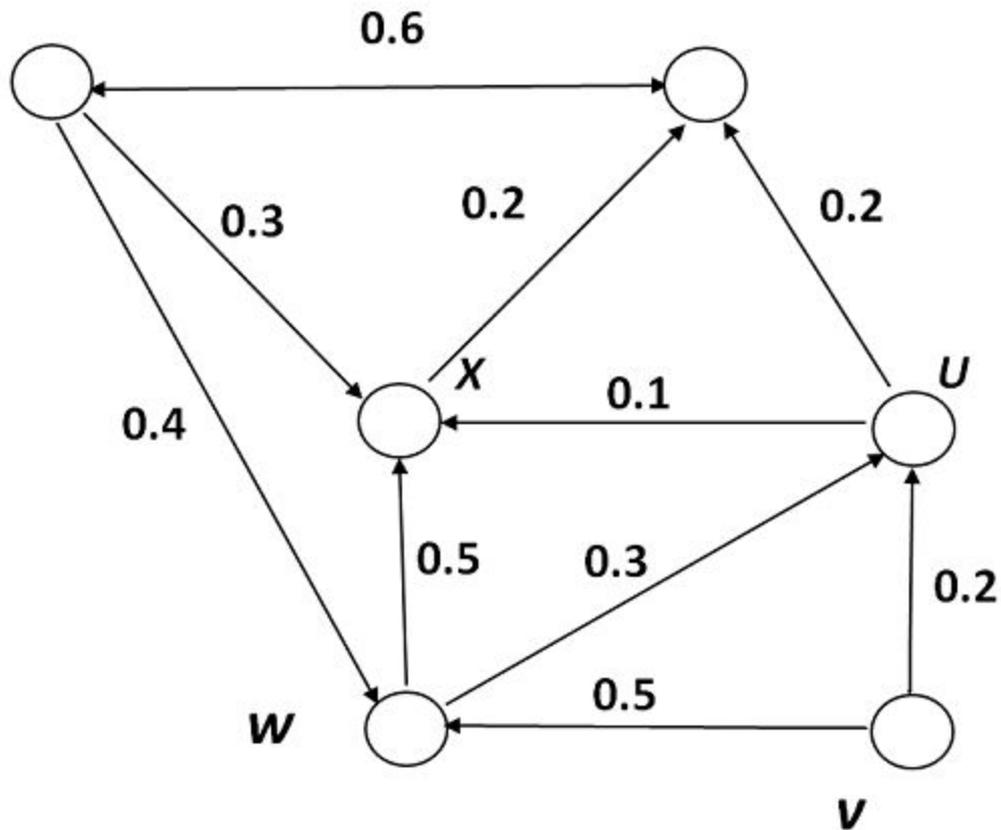
**Stop!**

# Independent Cascade Model

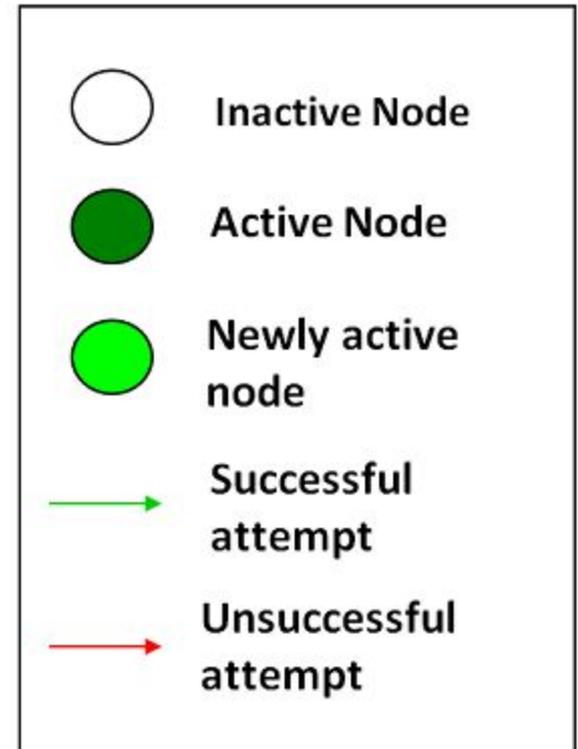
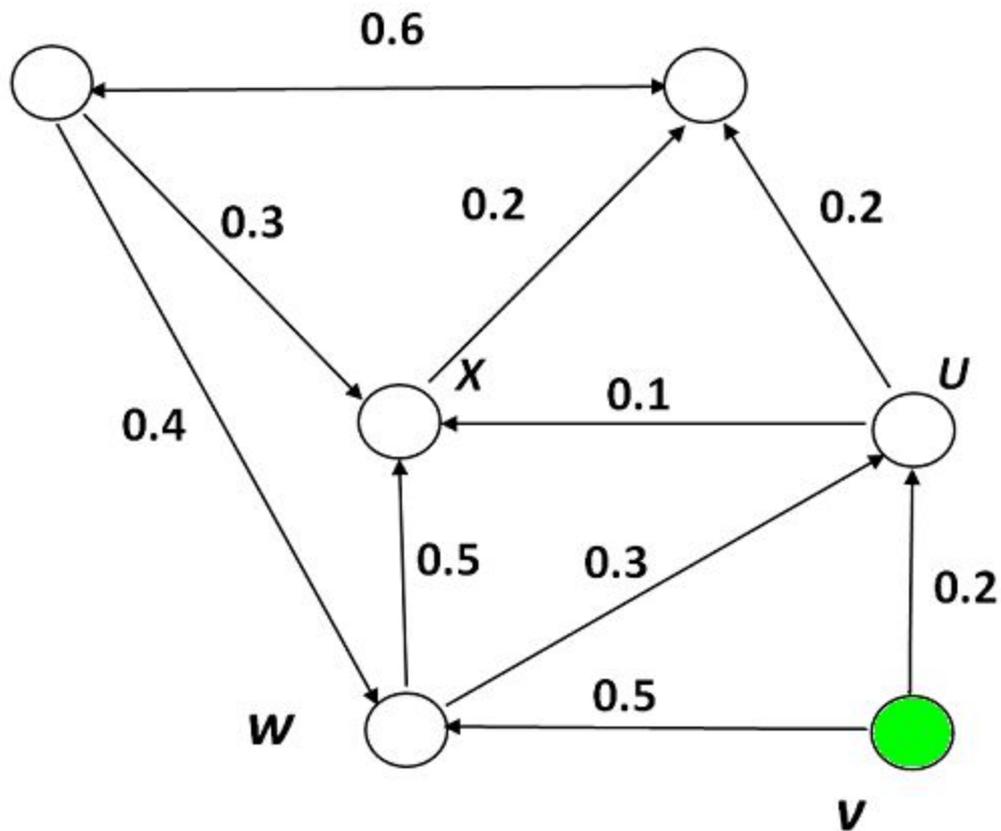
# Independent Cascade Model

- When node  $v$  becomes active, it has a **single** chance of activating each currently inactive neighbor  $w$ .
- The activation attempt succeeds with probability  $p_{vw}$ .

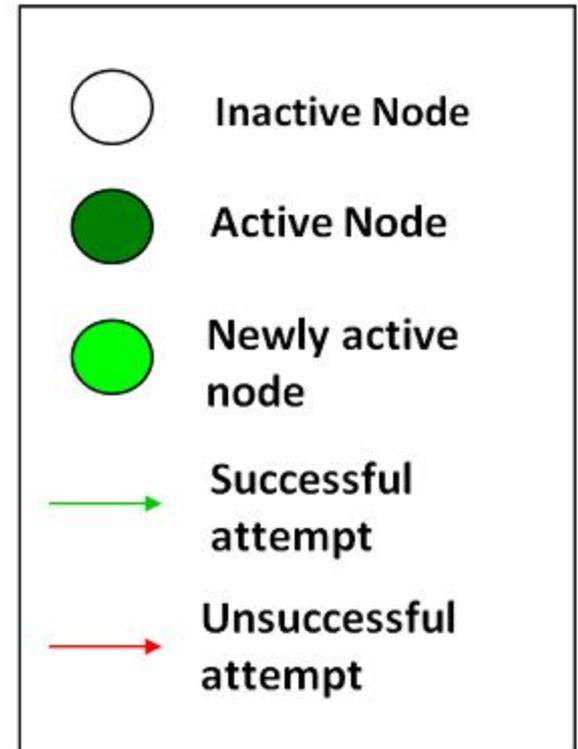
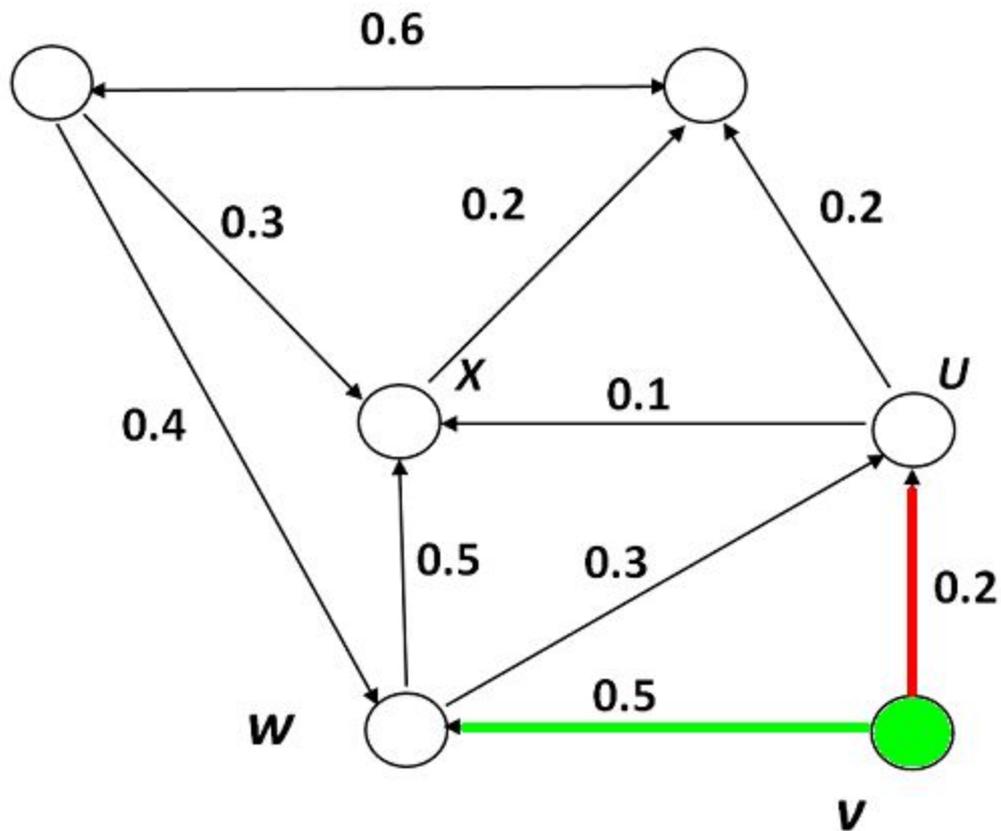
# Example



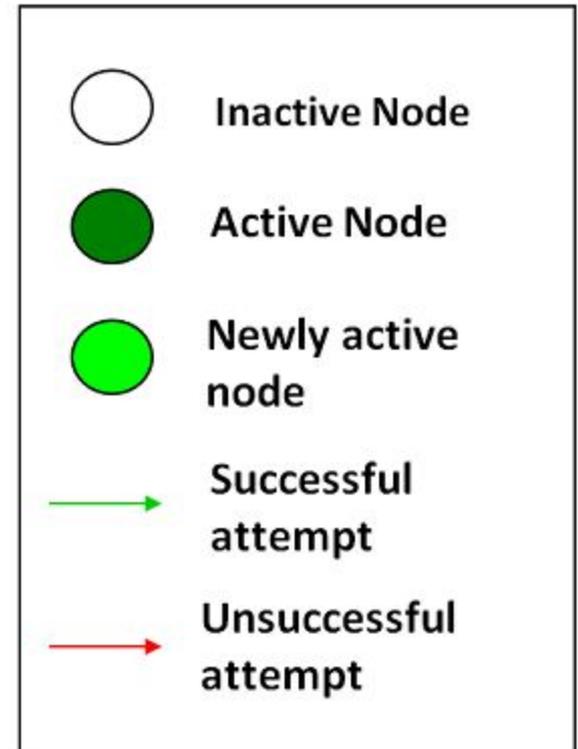
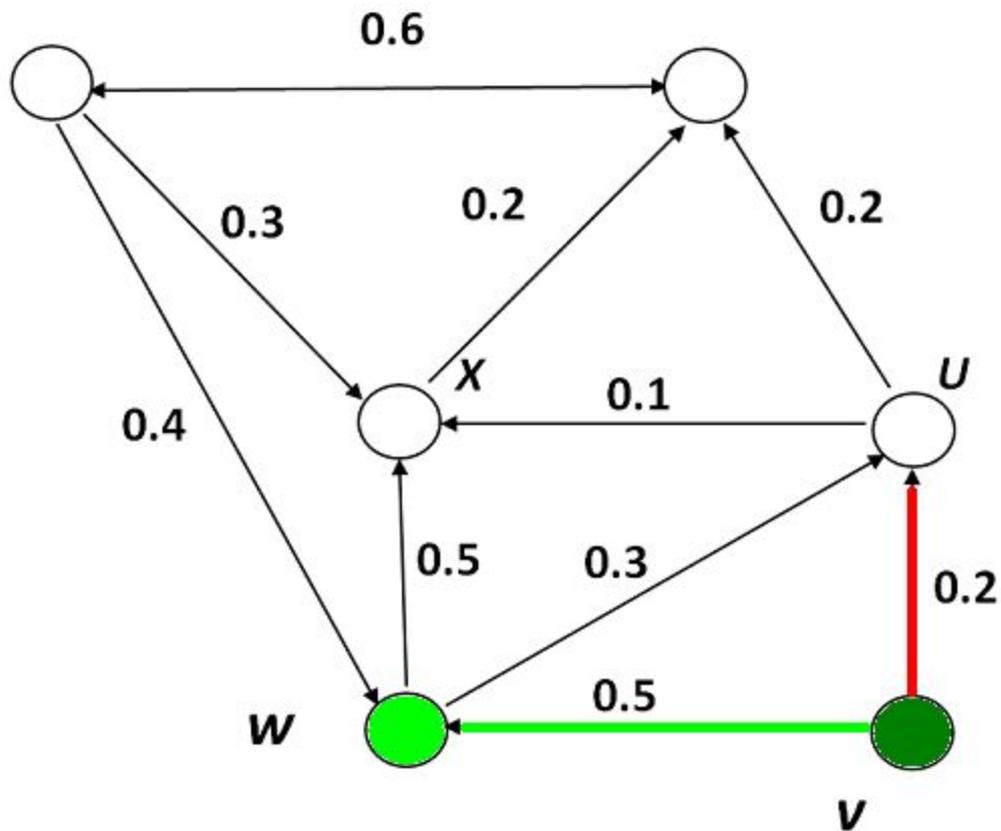
# Example



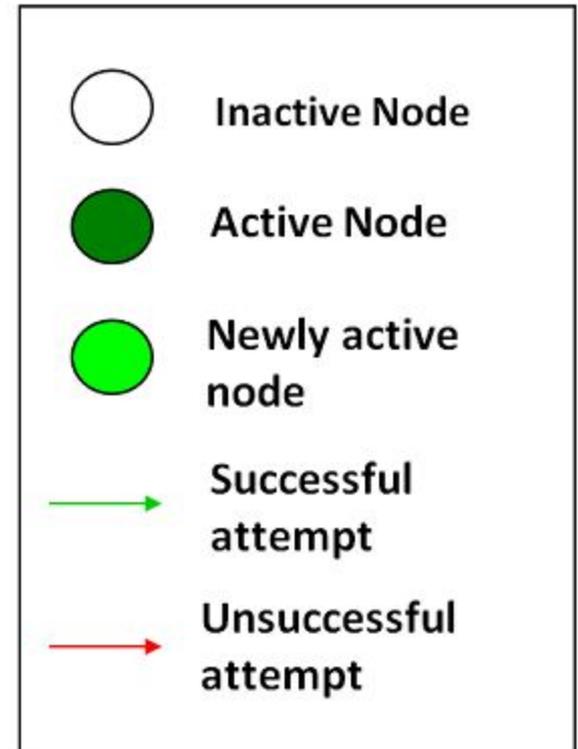
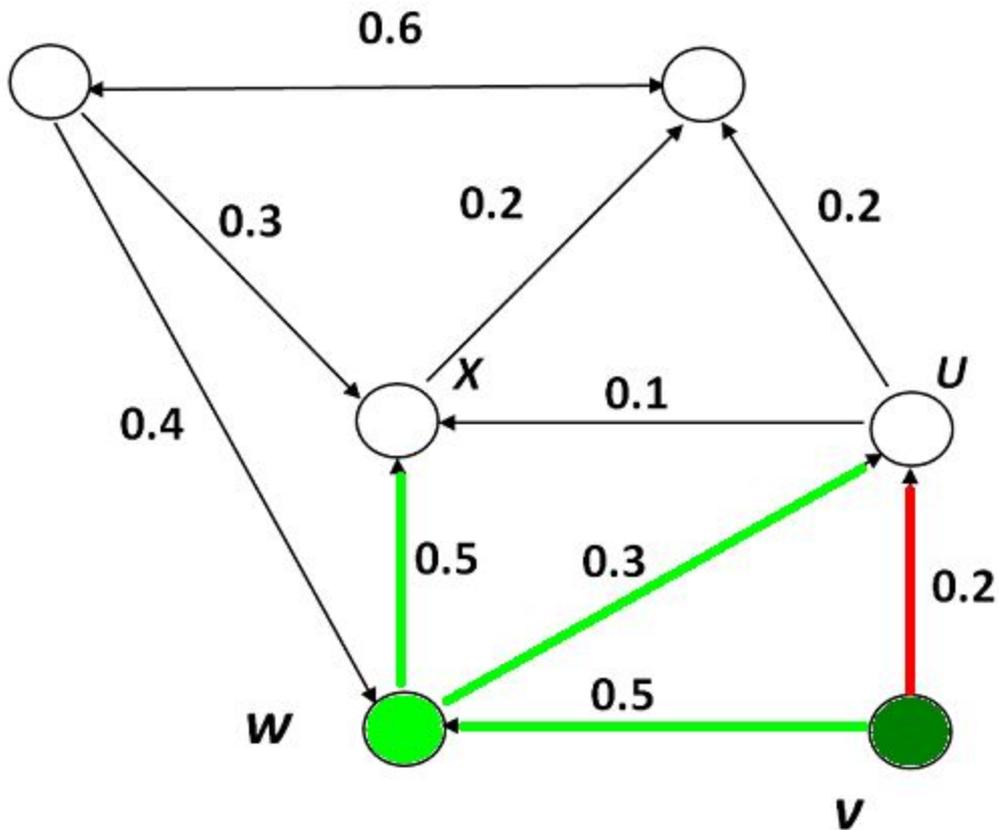
# Example



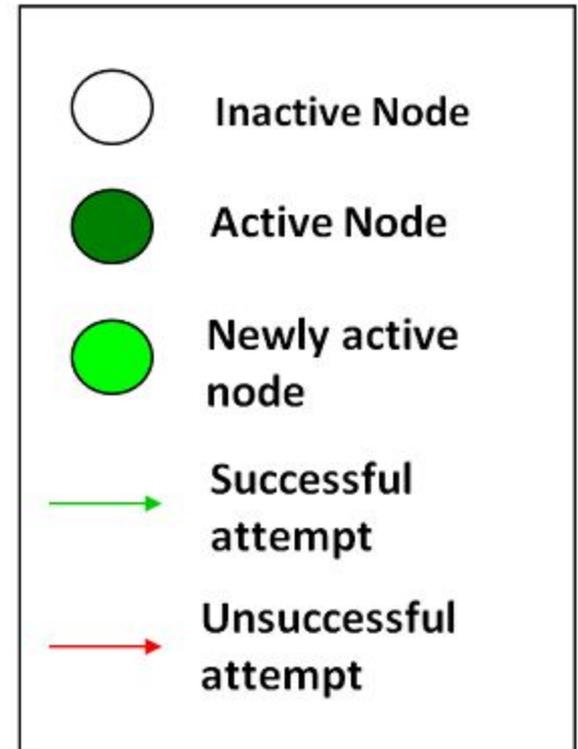
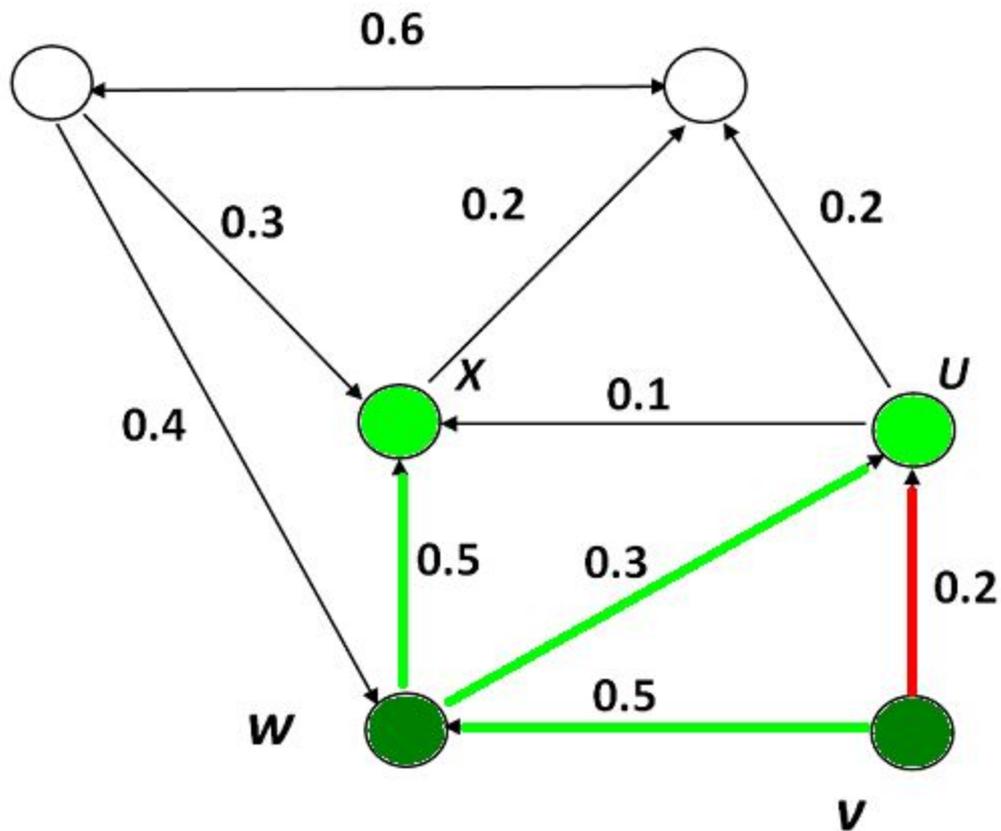
# Example



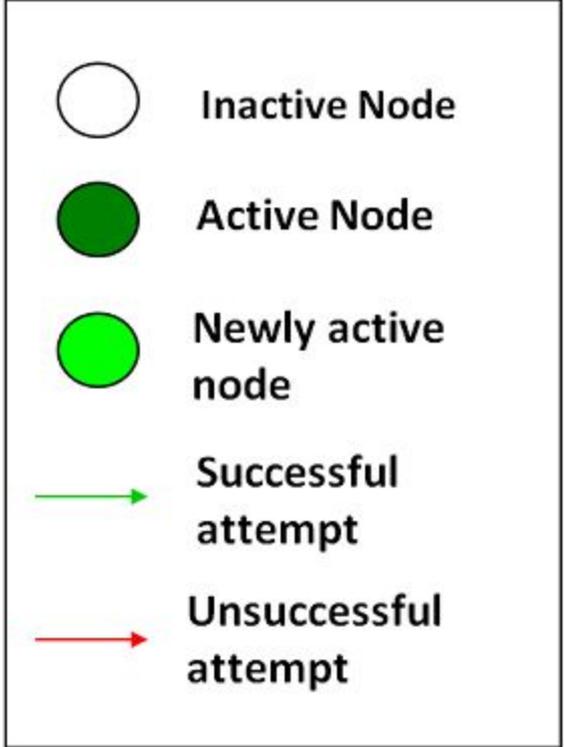
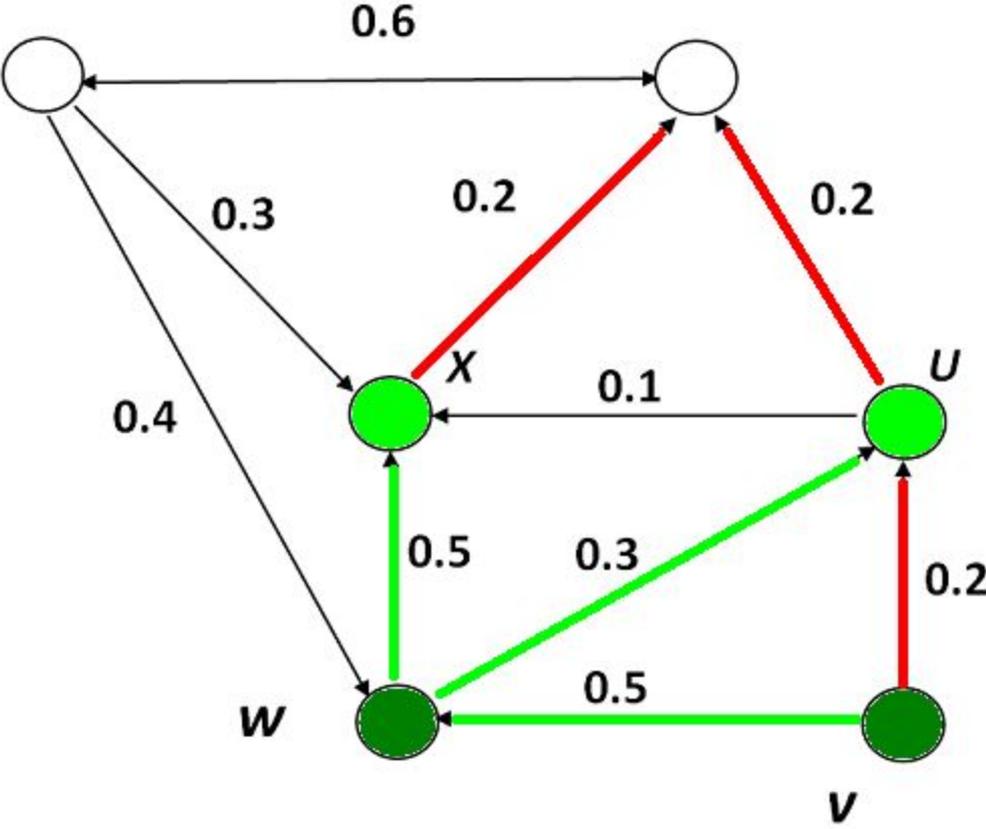
# Example



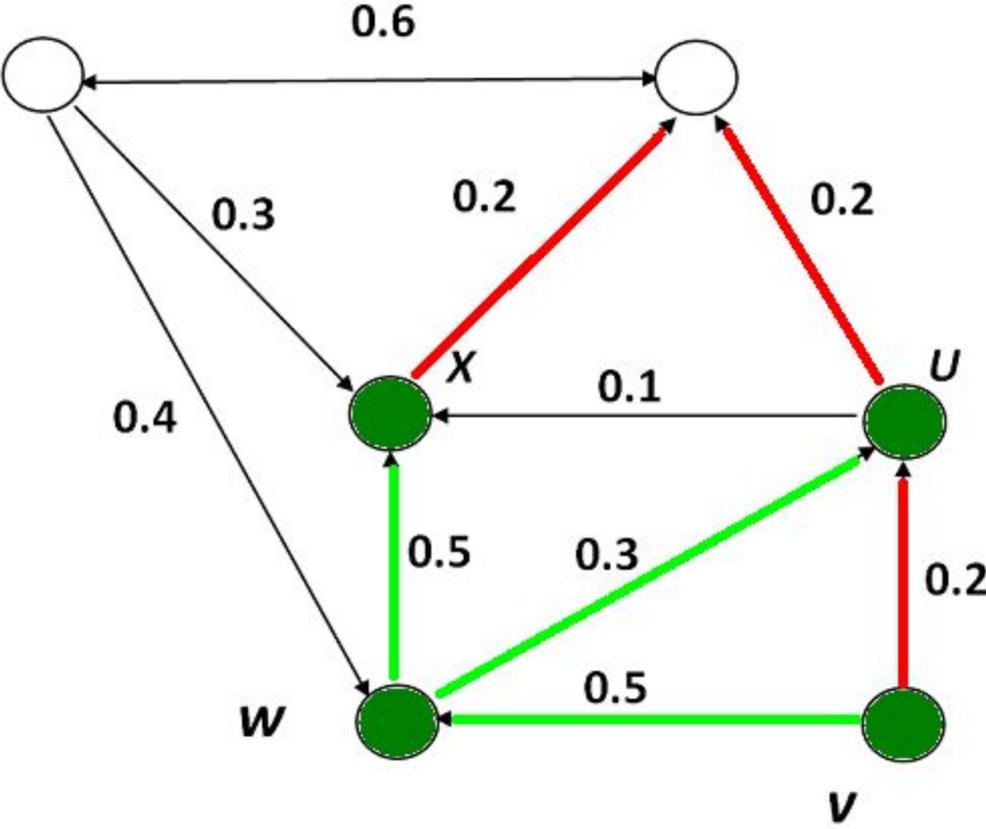
# Example



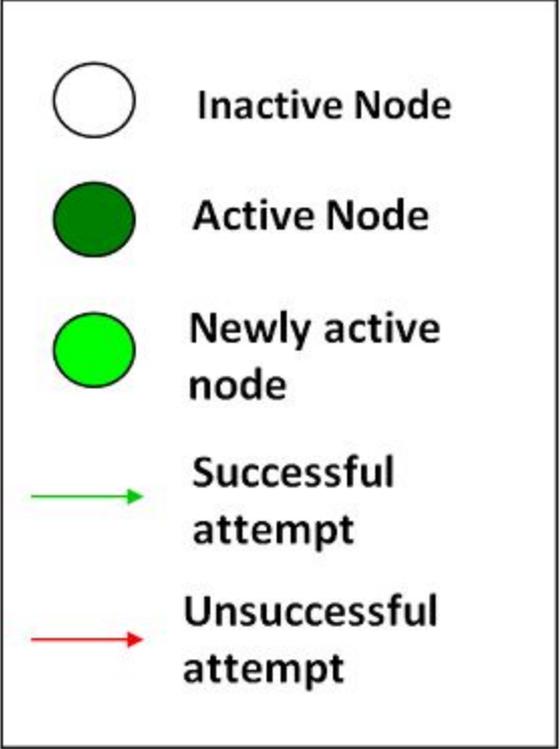
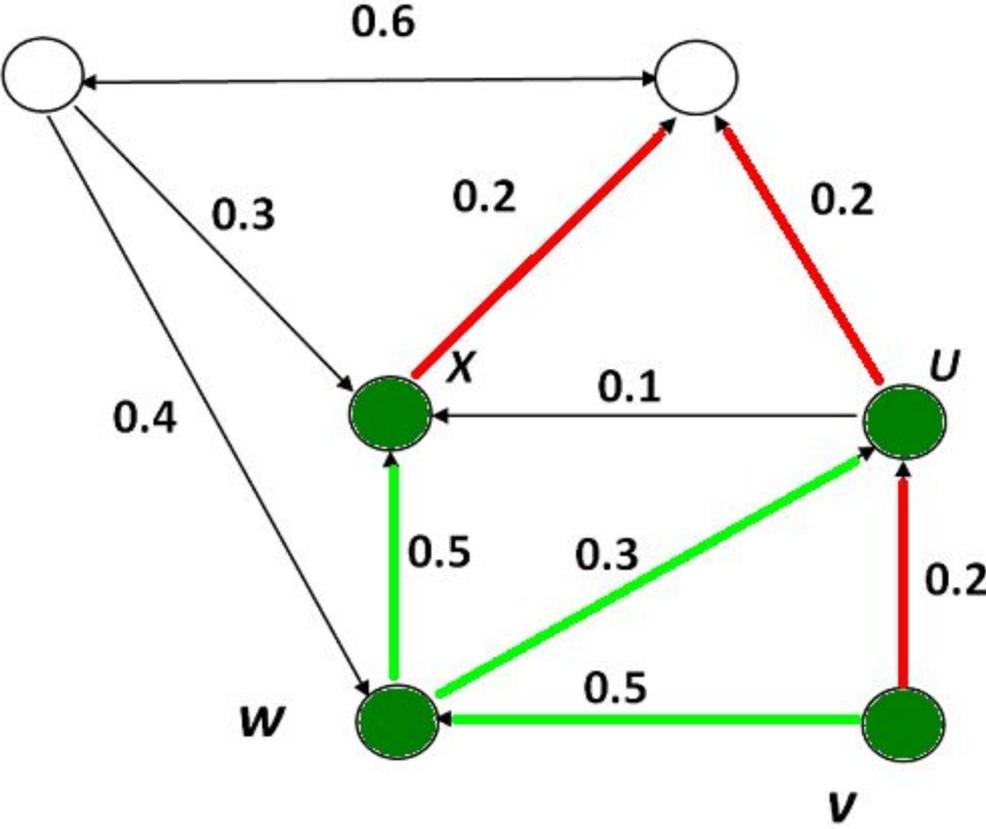
# Example



# Example



# Example



**Stop!**

# Theoretical properties

# Influence Maximization Problem

- Influence of node set  $S$ , denoted as  $I(S)$  (or  $f(S)$ ):  
The **expected** number of active nodes at the end, if set  $S$  is the initial active set.
- Problem:
  - Given a parameter  $k$ , find a  $k$ -node set  $S$  to maximize  $I(S)$

# Properties of $I(S)$

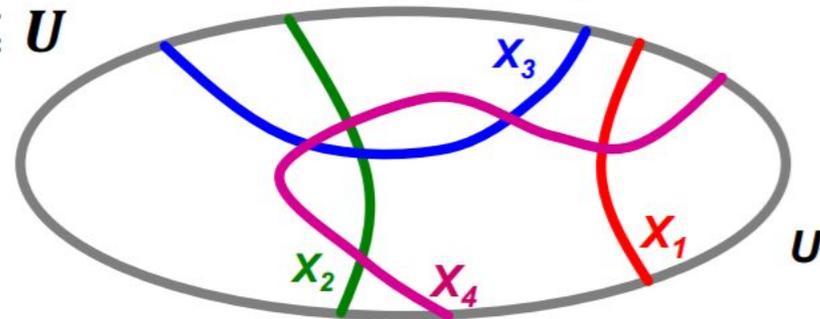
- Non-negative
- Monotone  $I(S \cup \{v\}) \geq I(S)$
- Submodular
  - Function  $I$  is submodular iff:

$$\forall S \subset T \subset N, \forall v \in N \setminus T, \\ I(S \cup \{v\}) - I(S) \geq I(T \cup \{v\}) - I(T)$$

# NP-Hardness of IM

- The problem is NP-Hard! (by reduction from the Set Cover Problem)
- Reminder - Set Cover Problem:

Given universe of elements  $U = \{u_1, \dots, u_n\}$   
and sets  $X_1, \dots, X_m \subseteq U$

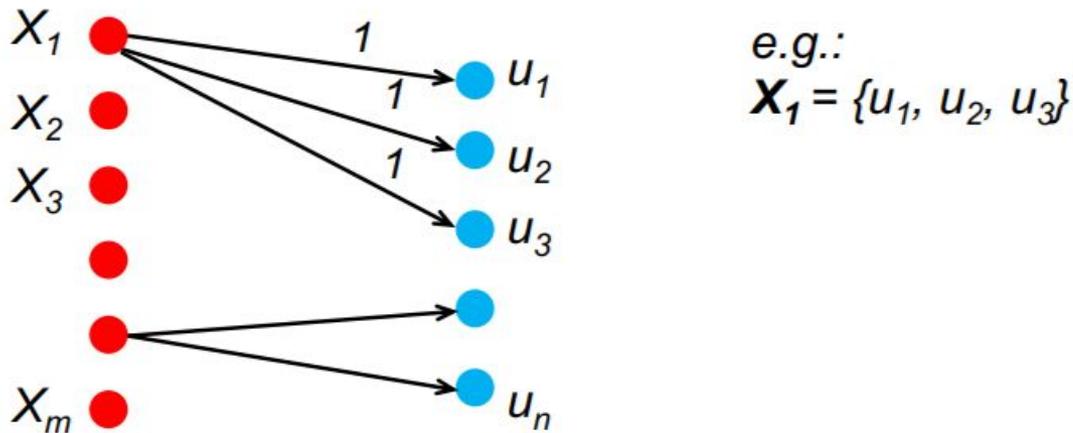


**Q: Are there  $k$  sets among  $X_1, \dots, X_m$  such that their union is  $U$ ?**

# The reduction (sketch)

Given an instance of Set Cover Problem with sets  $X_1 \dots X_m$ :

1. Build a bi-partite graph  $X$  to  $U$  by creating edge  $(X_i, u)$  for each  $u \in X_i$  (with weight 1)



2. Solution of k-IM problem will solve the k-Cover Set

# Approximation algorithm

Greedy Hill Climbing algorithm:

Start with  $S_0 = \{ \}$

For  $i = 1 \dots k$

- Activate node  $u$  that  $\max f(S_{i-1} \cup \{u\})$
- Let  $S_i = S_{i-1} \cup \{u\}$

Example:

Eval.  $f(\{a\}), \dots, f(\{e\})$ , pick argmax of them

Eval.  $f(\{d, a\}), \dots, f(\{d, e\})$ , pick argmax

Eval.  $f(\{d, b, a\}), \dots, f(\{d, b, e\})$ , pick argmax

# Approximation quality

- Hill climbing produces a solution  $S$  where:

$$f(S) \geq (1 - 1/e) * f(\text{OPT}) \quad [1 - 1/e \sim 0.63]$$

- Claim holds with 2 must properties of  $f$ :

**$f$  is monotone:** (activating more nodes doesn't hurt)

if  $S \subseteq T$  then  $f(S) \leq f(T)$  and  $f(\{\}) = 0$

**$f$  is submodular:** (activating each additional node helps less)

adding an element to a set gives less improvement

than adding it to one of its subsets:  $\forall S \subseteq T$

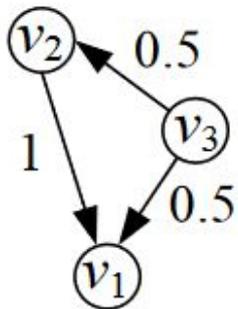
$$\underbrace{f(S \cup \{u\}) - f(S)}_{\text{Gain of adding a node to a small set}} \geq \underbrace{f(T \cup \{u\}) - f(T)}_{\text{Gain of adding a node to a large set}}$$

Gain of adding a node to a small set

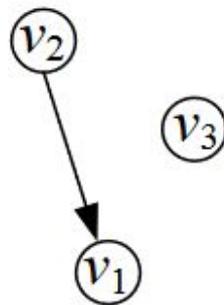
Gain of adding a node to a large set

# How to compute the $I(S)$

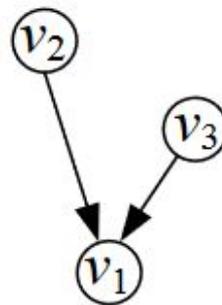
- Independent Cascade Model:
  - Take the original graph and generate an instance where the weights of edges are the probabilities
  - Repeat the process many times and compute the average (expected) number of edges reachable



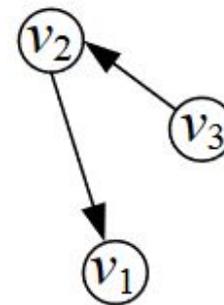
(a)  $G$



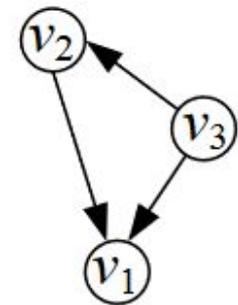
(b)  $g_1$



(c)  $g_2$



(d)  $g_3$

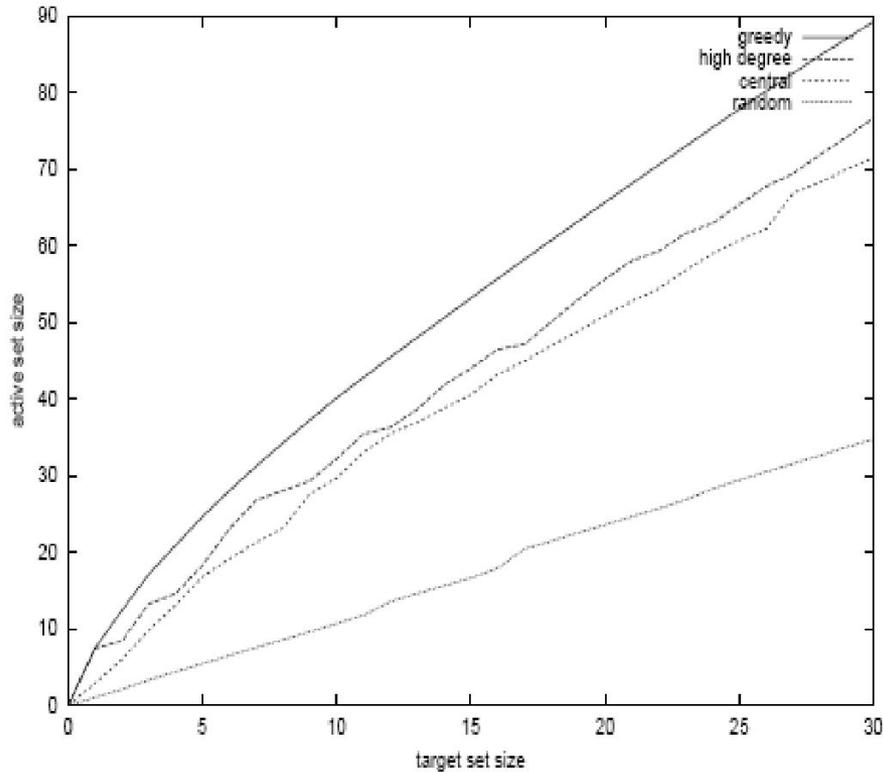


(e)  $g_4$

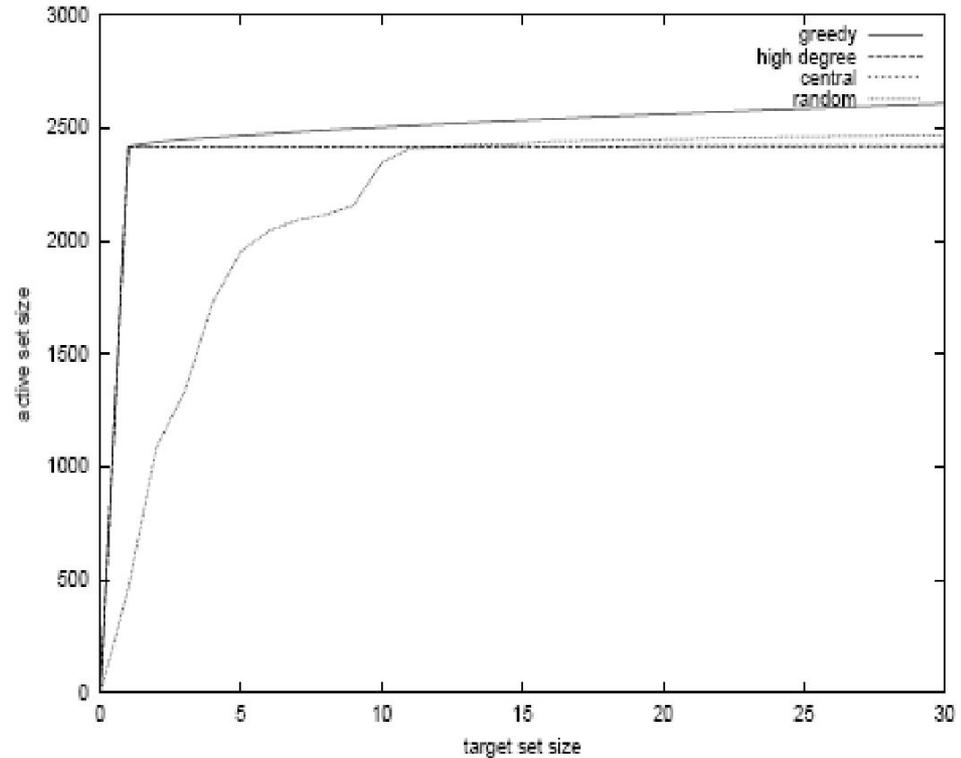
# Experimental Results

- Collaboration graph obtained from co-authorships in papers from arXiv's high-energy physics theory section
  - Claim: co-authorship networks capture many “key features”
  - Simple settings of the influence parameters
  - For each paper with 2 or more authors, edge was placed between them
- Competitors:
  - Degree centrality: Pick nodes with highest degree
  - Closeness centrality: Pick nodes in the “center” of the network
  - Random nodes: Pick a random set of nodes

# Experimental Results



probability = 1%



probability = 10%

# Discussion

- Greedy approach is very slow!
  - The complexity is  $O(k * n * m * R)$   
R – rounds, n – nodes, m – edges
- Optimization ideas:
  - Faster reachability computation
  - Heuristics like degree discount
- Open problems:
  - More realistic models
  - Negative influence



**Thank you!**  
**Questions?**