Information Cascades and Diffusion

Dr. Natarajan Meghanathan

Associate Professor of Computer Science

E-mail: natarajan.meghanathan@jsums.edu

Following the Crowd

- When people are connected by a network, it becomes possible for them to influence each other's behavior and decisions.

- There are many settings in which it may appear to be rational for an individual to imitate the choices of others even if the individual's own information suggests an alternative choice.

- Example: You want to choose a restaurant in an unfamiliar town. You read the reviews and choose restaurant A. You go there and you do not find anyone eating; however, there are a lot of people eating in a nearby restaurant B.
  - It would appear to make more sense to join the people in restaurant B instead of on the basis of your private info.
  - This is called <u>herding or information cascade</u>.

Following the Crowd

- Information cascade has the potential to occur when people make decisions sequentially, with later people watching the actions of earlier people, and from these actions inferring something about what earlier people know.
  - A cascade then develops when people abandon their own information in favor of inferences based on earlier people's actions.
  - Example: Each potential diner (who are yet to choose a restaurant) do not get to see the reviews read by the earlier diners; but, get to see the choice of the restaurant they made.
  - Information cascade contributes to imitation in social settings (fashions, voting for the popular candidate, success of books placed highly on best-seller lists, etc).

A Simple Herding Experiment (1)

- Before delving into the mathematical models for information cascades, we start with a simple herding experiment (created by Anderson and Holt) to illustrate how these models work.

- The experiment is designed to capture situations that involve the following basic ingredients of the models:
  - There is a decision to be made (like to adopt a new technology, wear a new style of clothing, etc)
  - People make the decision sequentially, and each person can observe the choices made by those earlier.
  - Each person has some private information that helps guide their decision.
  - A person cannot directly observe the private information that other people know; but, he or she can make inferences about this private information from what they do.

A Simple Herding Experiment (2)

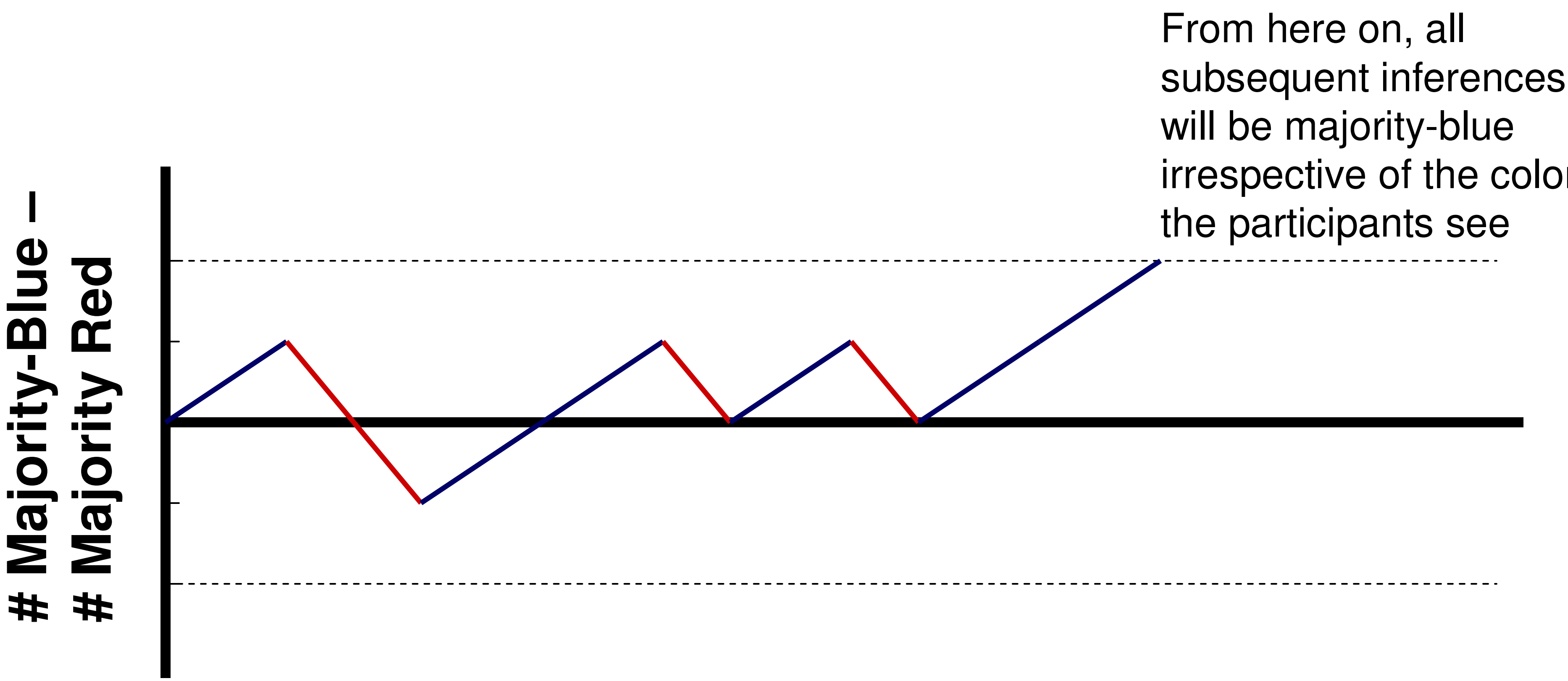
- The experiment takes place in a classroom with a large group of student participants.
- The experimenter puts an urn in front of the room, with three balls in the urn, which could be either red or blue
  - “Majority-red” urn: if two of the three balls are red
  - “Majority-blue” urn: if two of the three balls are blue
- Each student draws a ball from the urn; sees the color to himself (does not publicly display or tell it); guesses whether it would be a “majority-red” or “majority-blue” urn and then publicly announces his/her guess.
- We will assume that all students reason correctly about what to do when it is their turn to guess, using everything they have heard so far.

A Simple Herding Experiment (3)

- First Student: The first student decides based on what s/he sees.
  - If the student draws a "blue" marble, he could think the majority of them could be blue and publicly announce "majority-blue" and vice-versa
  - We will justify it mathematically later
- Second Student: For the second student, the first student's announcement is like already a draw made and appears a perfect information of what s/he saw.
  - So, if the second student also draws a ball of the same color announced by the first student, then, the second student also declares the urn to be of majority the same color.
  - If the second student draws a color that is different from that of the first student, the second student prefers to stick on to his/her drawn color and announces the urn to be of majority of that color
  - Either way, the second student's announcement looks like a perfect information of what s/he saw.
- Things do not get interesting if the absolute difference between the "number of majority-blues" and the "number of majority-reds" is less than 2. The participants tend to announce their inferences based on just the colors drawn.
  - If the first and second participants announce different colors, for the third participant, it is like again starting from fresh (no inferences could be drawn from previous draws).

A Simple Herding Experiment (4)

- Information Cascade starts to happen, only if the absolute difference between the number of “majority-blues” announced and the number of “majority-reds” announced is 2 or more.



A Simple Herding Experiment (5)

- Assume the first two students announce the same color (lets say, “majority-blue”). Remember, we have decided that the first two announcements convey perfect info.
- Third Student: Even if the third student draws “red,” the student thinks it is the only ball in red color and the other two balls (as announced by the previous two people) are in blue color. Hence, the third student announces “majority-blue,” ignoring his/her private info.
  – The rest of the class only gets to hear the announcement of the third student and not see the actual color s/he draws.
  – An information cascade has started.
- Fourth Student and Onward: Given that the first two students are considered to have told what they see (in this case, “blue” color balls), the announcement of the third student is irrelevant (conveys no information) for the fourth student, because in this case, – the third student simply decided based on what the first two students said (irrespective of what s/he saw).
  – Hence, the fourth student simply also decides like the third student and announces “majority-blue” irrespective of what s/he saw.
  – This is followed suit by the other students in the class as everyone tends to rely only on the limited genuine info available (the first two students).
  – If the first two students announce “majority-blue,” everyone else will also guess “blue”

Lessons Learnt from the Experiment

- Given the right structural conditions (like the first two students announcing "majority-blue"), each of a large group of students make exactly the same guess, even when all the decision-makers are completely rational.

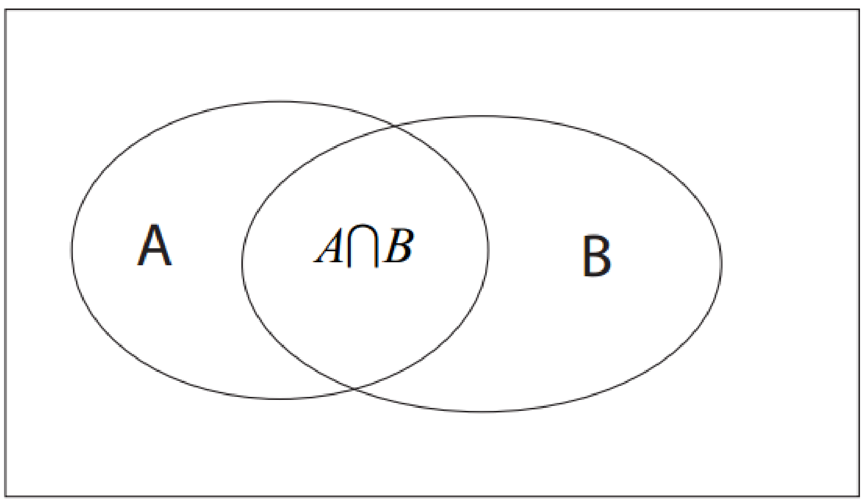
- Information cascade can sometimes lead to non-optimal outcomes.
  - Example: Consider the urn is "majority-red". There is a 1/3 chance that the first ball drawn is blue and there is another 1/3 chance that the second ball drawn is also blue (the color of the balls drawn are independent events). There is only a (1/3*1/3) 1/9 chance that the first two guess will be blue; but, if that happens, all the subsequent guesses will be blue – and they all will be wrong, since the urn is "majority-red".
  - There is a 1/9 chance of a population-wide error being not ameliorated by having many people participate.

Lessons Learnt from the Experiment

- Though information cascades has the potential to produce long runs of conformity, it can be fundamentally fragile.
  - Example: Suppose the first two students **told** “blue” and every subsequent student (3, 4, …) **guesses** blue. Lets say, students 50 and 51 turn notorious and show to everyone what they got. Lets say, they both **show** “red” color, student 52 now has four pieces of genuine information (two blue and two red colors). So, student 52 can essentially break the tie by announcing based on what s/he sees. We are simply back to the beginning.
  - This is the fragility of information cascades. Even after they have persisted for a long time, they can be overturned with little effort.

Bayes' Rule: A Model of Decision-Making under Uncertainty

- In order to build a mathematical model for how information cascades occur, we need a way to determine probabilities of events, given information that is observed.
- Given any event A, we will denote its probability of occurring by Pr[A].
- Whether an event occurs or not is the result of certain random outcomes. Imagine a large sample space consisting of a particular realization for each of these random outcomes (like in the rectangle shown below) whose area is 1.



Probability of an event corresponds to the area of the region represented with the event

Pr [A | B] should be referred to as the Probability of event A happening given that event B happened.
So, Pr [A | B] is the fraction of the area of event B and occupied by both events A and B.

$$\boxed{\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]}}$$

Bayes' Rule

$$\Pr[A \mid B] = \frac{\Pr[A \cap B]}{\Pr[B]} \Rightarrow \Pr[A \cap B] = \Pr[A \mid B] * \Pr[B]$$

$$\Pr[B \mid A] = \frac{\Pr[A \cap B]}{\Pr[A]} \Rightarrow \Pr[A \cap B] = \Pr[B \mid A] * \Pr[A]$$

$$\Pr[A \cap B] = \Pr[A \mid B] * \Pr[B] = \Pr[B \mid A] * \Pr[A]$$

$$\boxed{\Pr[A \mid B] = \frac{\Pr[A] * \Pr[B \mid A]}{\Pr[B]}}$$

Pr[A] is simply the probability of Event A happening irrespective of whether event B happened or not. Pr[A] is hence the prior probability of A.

Pr[A | B] is the probability that event A happened given that event B happened. Hence, Pr[A | B] is the posterior probability of A (given B happened).

Bayes' rule captures change from the prior probability of A to the posterior probability

Example (1): Bayes' Rule

- Lets say a city has 80% of taxi cabs in black color and the remaining 20% in yellow.
- A hit-and-run accident has happened involving a taxi and a witness states that he saw a yellow taxi (report).
- Assume that if a taxi is yellow, then a witness will claim it to be yellow after the fact 80% of the time; and if it is black, they will claim it is black 80% of the time.
- We are interested in finding out the probability that the taxi he saw is indeed yellow (actual).

Example (1): Bayes' Rule (Solution)

$$\Pr[actual = Y \mid report = Y] = \frac{\Pr[actual = Y] * \Pr[report = Y \mid actual = Y]}{\Pr[report = Y]}$$

$$\begin{aligned}\Pr[report = Y] = {} & \Pr[actual = Y] * \Pr[report = Y \mid actual = Y] + \\ & \Pr[actual = B] * \Pr[report = Y \mid actual = B]\end{aligned}$$

Pr[report = Y] = 0.2*0.8 + 0.8*0.2 = 0.32

$$\Pr[actual = Y \mid report = Y] = \frac{0.2 * 0.8}{0.32} = 0.5$$

Hence, even though there are only 20% taxi cabs in Yellow, given that the witness reported that the cab involved in the accident is yellow, the probability that the cab involved in the accident is indeed yellow is 0.5 (50%).

Example (2): Bayes' Rule

- Bayes' rule plays a crucial conceptual ingredient in the design of e-mail spam filters to automatically filter unwanted e-mail out of a user's incoming e-mail stream.

- Suppose you receive an e-mail that contains a subject "check this out". We are interested in finding out the probability that the e-mail would be a spam.

- Assume 40% of your e-mails are spam; 1% of all spam e-mails contain the word "check this out"; and 0.4% of all non-spam e-mails contain the word "check this out"

Example (2): Bayes' Rule: Solution

$$\Pr[spam \mid \text{"}checkthisout\text{"}] = \frac{\Pr[spam] * \Pr[\text{"}checkthisout\text{"} \mid spam]}{\Pr[\text{"}checkthisout\text{"}]}$$

$$\Pr[\text{"checkthisout"}] = \Pr[\text{spam}] * \Pr[\text{"checkthisout"} \mid \text{spam}] + \Pr[\text{not spam}] * \Pr[\text{"checkthisout"} \mid \text{not spam}]$$

$$\Pr[\text{"checkthisout"}] = 0.4 * 0.01 + 0.6 * 0.004 = 0.0064$$

$$\Pr[spam \mid \text{"}checkthisout\text{"}] = \frac{0.4 * 0.01}{0.0064} = 0.625$$

We could use several such signals (like “checkthisout”) to find the probability that an Email is spam (given that it has the signal).
The overall probability that an email (containing one or more of this signals is spam) is the weighted average of all the individual probabilities. The spam filter program could simply generate a random number and if it is less than the weighted average, then the e-mail is classified as “spam”.

Bayes' Rule: Herding Experiment (1)

- The urn is equally likely to be either "majority-blue" or "majority-red"
  - Pr[majority-blue] = Pr[majority-red] = ½
- Based on the compositions of the two urns, if the urn is a "majority-blue" one, the probability that a ball picked is blue is 2/3; likewise, if the urn is a "majority-red" one, the probability that a ball picked is red is 2/3.
  - Pr[blue | majority-blue] = 2/3
  - Pr[red | majority-red] = 2/3
- We are interested in finding the probability that the urn is indeed majority-blue if what we hear and see is "blue" (from the first two students)

Bayes' Rule: Herding Experiment (2)

$$\Pr["majority-blue" \mid blue] = \frac{\Pr[majority-blue] * \Pr[blue \mid "majority-blue"]}{\Pr[blue]}$$

$$\Pr[\text{blue}] = \Pr[\text{majority-blue}]*\Pr[\text{blue}|\text{"majority-blue"}] + \Pr[\text{majority-red}]*\Pr[\text{blue}|\text{"majority-red"}]$$

$$= 1/2 * 2/3 + 1/2 * 1/3 = 1/2$$

$$\Pr["majority-blue" \mid blue] = \frac{1/2 * 2/3}{1/2} = 2/3$$

This gives the basis for each of the first two students to tell the urn is "majority-blue" if they indeed see blue color.

Bayes' Rule: Herding Experiment (3)

- Lets compute the probability that the urn will be still considered to be "majority blue" by the third student, who has heard from the first two students (for sure that the ball they saw is blue) and the third student sees a "red".
- Thus, we are interested in finding
  - Pr ["majority-blue" | heard blue, blue, sees red]

$$\Pr[\text{"}majority-blue\text{"} \mid blue, blue, red] = \frac{\Pr[majority-blue] * \Pr[blue, blue, red \mid \text{"}majority-blue\text{"}]}{\Pr[blue, blue, red]}$$

$$\begin{aligned}
\Pr[\text{blue, blue, red}] &= \Pr[\text{majority-blue}] * \Pr[\text{blue, blue, red} \mid \text{"majority-blue"}] + \\
&\quad \Pr[\text{majority-red}] * \Pr[\text{blue, blue, red} \mid \text{"majority-red"}] \\
&= [\ 1/2 * (2/3)*(2/3)*(1/3)\ ] + [\ 1/2 * (1/3)*(1/3)*(2/3)\ ] \\
&= 2/27 + 1/27 = 3/27 = 1/9
\end{aligned}$$

$$\Pr[\text{"}majority-blue\text{"} \mid blue, blue, red] = \frac{1/2 * (2/3) * (2/3) * (1/3)}{1/9} = \frac{2/27}{1/9} = 2/3$$

Bayes' Rule: Herding Experiment (4)

- Thus, even if the third student sees a red color, given that the first two students indeed saw blue, gives a 2/3 chance (greater than 0.5) that the urn is indeed "majority-blue". Thus, the third student makes a guess that the urn would be majority-blue and announces the same.

- The above analysis holds good for all the subsequent students. They hear the actual colors seen by the first two students and irrespective of whether they see red or blue, the probability that the urn would be majority-blue is > ½

- Hence, it makes sense for every subsequent student (3, 4, …) to guess the urn would be majority-blue and move on.

Information Cascades:-Multiple Signals

- Consider a modified scenario of the urn-majority color problem, where every user has the same hypothesis (i.e., say urn is majority-blue).
- Each user goes to the table and picks a ball. If the user sees the ball to be of the same color as his hypothesis, he presses a “high” signal button; on the other hand, if the user sees the ball to be of a different color as his hypothesis, he presses a “low” signal button.
- Let there be N users. The Nth user gets to see the sequence of *high* and *low* signals pressed by the users 1…N-1.
  - Let there be ‘a’ high signals and ‘b’ low signals seen by the Nth user.
  - The actual ordering of the high and low signals does not matter. It is the count of the high signals and low signals that matters for the Nth user.

Information Cascades:-Multiple Signals (2)

- Let q be the probability that a user picks a ball of the <u>**same color**</u> as that of the "actual" majority color of the urn:
  - A user picking "blue" color ball, while his hypothesis majority-blue is indeed actually true (i.e., the urn is actually majority-blue)
  - A user picking "red" color ball, while his hypothesis majority-blue is indeed actually false (i.e., the urn is actually majority-red)
- Then, (1-q) is the probability that a user picks a ball of a <u>**different color**</u> (compared to his hypothesis being actually true or false):
  - A user picking "red" color ball, while his hypothesis majority-blue is indeed true (i.e., the urn is actually majority-blue)
  - A user picking "blue" color ball, while his hypothesis majority-blue is indeed false (i.e., the urn is actually majority-red)
- Remember, each of the 1, …, N-1 users (have the hypothesis: the urn is majority-blue) and press a "high" signal if they the color of the ball they pick is "blue" and press a "low" signal if they pick a "red" color ball.

Information Cascades:-Multiple Signals (3)

- Remember, **like the Nth user, each of the 1, …, N-1 users** (**have the hypothesis: the urn is majority-blue**) and press a "high" signal if they the color of the ball they pick is "blue" and press a "low" signal if they pick a "red" color ball.
- High signals
  - P[see blue | hypothesis is actually true (urn is majority-blue)] = q
  - P[see blue | hypothesis is actually false (urn is majority-red)] = 1-q
- Low signals
  - P[see red | hypothesis is actually true (urn is majority-blue)] = 1-q
  - P[see red | hypothesis is actually false (urn is majority-red)] = q
- We know that if an urn is indeed actually majority-blue, the probability of seeing a blue color ball is > ½.
- Likewise, if an urn is indeed actually majority-red, the probability of seeing a red color ball is > ½.
- Hence, for both high and low signal scenarios, q > ½. Hence, (1-q) < ½.

Information Cascades:-Multiple Signals (4)

- The Nth user has initially a probability of ‘p’ for his hypothesis (urn is majority-blue) to be true.
- Now that he has seen a sequence of S signals (‘a’ high and ‘b’ low signals) by the 1…N-1 users, we are interested in finding out :

  P[urn is majority-blue | S]
- Remember, the probability of a signal pressed by each user is independent of that pressed by another user.
- From Bayes’ theorem:

P[urn is majority-blue | S] = P[urn is majority-blue] x
P[S | urn is actually majority-blue] / P[S]

P[S] = P[urn is majority-blue] * P[S | urn is actually majority-blue] +
P[urn is majority-red] * P[S | urn is actually majority-red]

P[S | urn is actually majority-blue] = P[high signal | urn is actually majority-blue]^a x
P[low signal | urn is actually majority-blue]^b.

P[S | urn is actually majority-blue] = $(q)^a\,(1\text{-}q)^b$.

P[S | urn is actually majority-red] = P[high signal | urn is actually majority-red]^a x
P[low signal | urn is actually majority-red]^b.

P[S | urn is actually majority-red] = $(1\text{-}q)^a\,(q)^b$.

Information Cascades:-Multiple Signals (5)

P[S] = P[urn is majority-blue] * P[S | urn is actually majority-blue] + P[urn is majority-red] * P[S | urn is actually majority-red]

P[S] = p* $(q)^a$ $(1\text{-}q)^b$ + (1-p)* $(1\text{-}q)^a$ $(q)^b$.

P[urn is majority-blue | S]

= P[urn is majority-blue] * P[S | urn is actually majority-blue] / P[S]

That is,

$$\Pr[H = true \mid S] = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a}$$

We want to find for what values of 'a' and 'b' would Pr [H = true | S] > p, where p is the initial (prior) probability that the hypothesis is true.

Information Cascades:-Multiple Signals (6)

- **In Summary,**
- Let there be a sequence S of ‘a’ high signals and ‘b’ low signals (in some order) for a hypothesis H.

$$\Pr[H = true \mid S] = \frac{P[H = true] * P[S \mid H = true]}{P[S]}$$

$$P[S] = P[H = true]*P[S|H=true] + P[H = false]*P[S|H=false]$$
$$= [\ p * q^a * (1-q)^b\ ] + [\ (1-p) * q^b * (1-q)^a\ ]$$

$$\Pr[H = true \mid S] = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a}$$

We want to find for what values of ‘a’ and ‘b’ would Pr [H = true | S] > p, where p is the initial (prior) probability that the hypothesis is true.

Information Cascades:-Multiple Signals (7)

$$\Pr[H = true \mid S] = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a}$$

- Let us see what happens if we replace the second term in the denominator with $(1\text{-}p)q^a(1\text{-}q)^b$.

<u>Case 1: If 'a' > 'b':</u>

The denominator will become: $pq^a(1\text{-}q)^b + (1\text{-}p)q^a(1\text{-}q)^b = q^a(1\text{-}q)^b$.

As a result, the whole ratio will become p.

Since $q > ½$, $1\text{-}q < ½$.

$(1\text{-}p)q^b(1\text{-}q)^a < (1\text{-}p)q^a(1\text{-}q)^b$. Hence, the replacement will increase the value of the denominator. That is,

$$\frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a} > \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^a(1-q)^b}$$

$$\Pr[H = \text{true} \mid S] > p$$

Thus, if the sequence of signals is such that the number of high signals is greater than the number of low signals, the probability for the Hypothesis H to be true is greater than the prior probability **(ACCEPT THE HYPOTHESIS)**

Information Cascades:-Multiple Signals (8)

$$\Pr[H = true \mid S] = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a}$$

- Let us see what happens if we replace the second term in the denominator with $(1\text{-}p)q^a(1\text{-}q)^b$.

<u>Case 2: If 'a' < 'b':</u>

The denominator will become: $pq^a(1\text{-}q)^b + (1\text{-}p)q^a(1\text{-}q)^b = q^a(1\text{-}q)^b$.

As a result, the whole ratio will become p.

Since $q > ½$, $1\text{-}q < ½$.

$(1\text{-}p)q^b(1\text{-}q)^a > (1\text{-}p)q^a(1\text{-}q)^b$. Hence, the replacement will decrease the value of the denominator. That is,

$$\frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a} < \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^a(1-q)^b}$$

$$\Pr[H = true \mid S] < p$$

Thus, if the sequence of signals is such that the number of high signals is lower than the number of low signals, the probability for the Hypothesis H to be true is lower than the prior probability **(REJECT THE HYPOTHESIS).**

Information Cascades:-Multiple Signals (9)

$$\Pr[H = true \mid S] = \frac{pq^a(1-q)^b}{pq^a(1-q)^b + (1-p)q^b(1-q)^a}$$

- Let us see what happens if we replace the second term in the denominator with $(1\text{-}p)q^a(1\text{-}q)^b$.

<u>Case 3: If 'a' = 'b':</u>

The denominator will become: $pq^a(1\text{-}q)^b + (1\text{-}p)q^a(1\text{-}q)^b = q^a(1\text{-}q)^a$, canceling with the numerator.

$$\Pr [\, H = \text{true} \mid S] = p$$

Thus, if the sequence of signals is such that the number of high signals is equal to the number of low signals, the probability for the Hypothesis H to be true is equal to the prior probability (an indication that more signals need to be observed).

Sequential Decision Making and Cascades

- Consider a person N.
- If the number of acceptances and rejections of the hypothesis (say, "majority-blue") are equal, then person N uses his own signal as the tie breaker and decides accordingly.

- If the number of acceptances differs (absolute difference) from the number of rejections by 1, then the person N still decides based on his private (own) signal.
  - If the person's private signal reinforces the current state of the hypothesis being true or false (prior to N), then accordingly the absolute difference between the number of acceptances and number of rejections becomes 2. An information cascade will begun from the next person.
  - If the person's private signal contradicts the current state of the hypothesis being true or false (prior to N), then the number of acceptances will be equal to the number of rejections.

Sequential Decision Making and Cascades (2)

- If the absolute difference between the number of acceptances and rejections is greater than 2,
  - we observe that the posterior probability of the hypothesis being true (or false) is greater than the prior probability of the hypothesis being true (or false).
- Hence, private signals are ignored.

- The person publicly accepts (or rejects) the hypothesis, depending on the majority signal → accordingly, the number of acceptances (or rejection) gets incremented, and will continue to be incremented.
  - An information cascade has begun.

Concluding Observations

- The general conclusion is:
  - When people can see what others do but not what they know, there is an initial period when people rely on their own private information, but as time goes on, the population can tip into a situation where people – still behaving fully rationally – begin ignoring their own information and follow the crowd.
- Marketing Strategy:
  - Make an initial set of customers buy your product (but not reveal their experiences with the product!)
  - The general public is likely to be swayed by the observation that several people had already bought the product.
- Decision making in meetings:
  - If all people have access to the same information, then an information cascade could happen, if the participants are asked to convey the decision in a sequence.
  - To avoid information cascade in such public settings, the participants could be asked to privately vote and a majority decision can be then taken.

Concluding Observations (2)

- Cascades can be wrong:
  - If the hypothesis is supposed to be false; but, if the first two people happen to get high signals, a cascade of acceptances will start immediately, even though it is the wrong choice for the population.

- Cascades can be based on very little information:
  - Since people ignore their private info once a cascade starts, only the pre-cascade info influences the behavior of the population.
  - If the cascade starts relatively quickly in a large population, the private signals of the individuals are ignored.

- Cascades are fragile:
  - Cascades are easy to stop too.
  - People receiving the private signals of few of those who earlier took part in the experiment can overturn a long-lived cascade.

Diffusion in Networks

- So far, we focused on mechanism where an individual's choices is based on the knowledge that the individual has about everyone else.
- Many of our interactions with the rest of the world happen at a local, rather than a global level
- We often do not care as much about the full population's decisions as about the decisions made by friends and colleagues.
  - In a work setting, we may choose technology to be compatible with the people we directly collaborate with, rather than the universally most popular technology
  - We may adopt political views that are aligned with those of our friends, even of they are nationally in minority.
- We will now focus on taking decisions based on the direct benefits that we get in the interaction with our neighbors.
  - People tend to align their behaviors with those of their network neighbors if it could give better payoffs (benefits)

The Diffusion of Innovations

- We will consider how new practices, opinions, technologies, etc, spread from person to person through a network, as people influence their neighbor nodes (friends and colleagues) to adopt new ideas.

- Research has shown that though people get to hear about an innovation much earlier, they are less likely to adopt it until they see some of their neighbors adopting the invention and they see a direct benefit of adopting the same as their neighbor.
  - Example: One tends to adopt a technology that a majority of his/her friends have adopted to communicate.

- The benefits of adopting a new behavior or innovation increases as more and more neighbors adopt it.
  - It makes sense to adopt a new behavior once sufficient proportion of neighbors have done so.

- Homophily could sometimes appear to be a barrier to diffusion of innovations as the latter tends to come from “outside” the system; while, Homophily is a measure of people aligning with those who are like themselves.

Modeling Diffusion through a Network

- Assume each node has a choice between two possible behaviors (innovations, etc), labeled A and B.
- For two nodes *v* and *w* linked by an edge:
  - If both adopt A, they each get a payoff ‘a’ > 0
  - If both adopt B, they each get a payoff ‘b’ > 0
  - If they adopt opposite behaviors, they each get a payoff of 0.
- The strategy to be adopted by a node v depends on the choice(s) of its neighbors.
- Assume a node v has fraction ‘p’ of ‘d’ neighbors who adopt A and fraction (1-p) of d neighbors who adopt B.
  - If the node adopts A, it gets a payoff of pda
  - If the node adopts B, it gets a payoff of $(1-p)db$
  - Hence, A is the better choice if $pda \geq (1-p)db$
  - Rearranging the terms, we get:

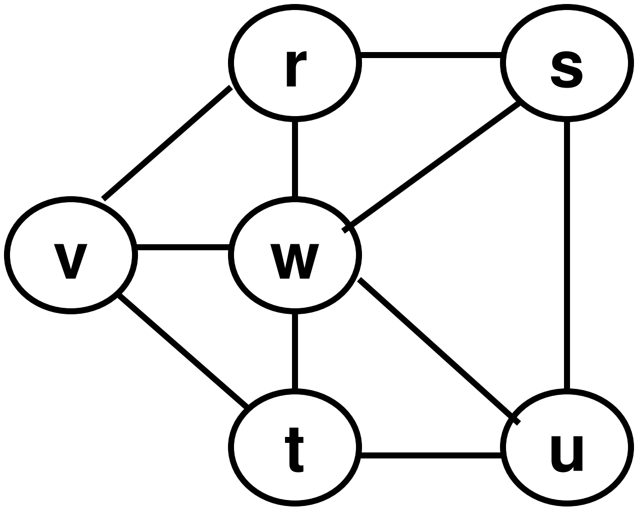
$$q = \frac{b}{a+b}$$

$$p \geq \frac{b}{a+b}$$

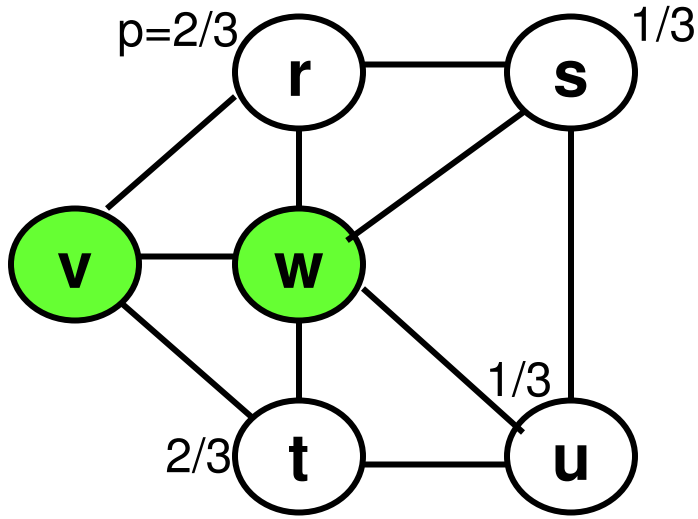
Hence, if at least b/(a+b) of the neighbors adopt A, the node should adopt A.

Initially, except the initial adopters, We will assume all other nodes have adopted B

Diffusion Modeling: Example 1



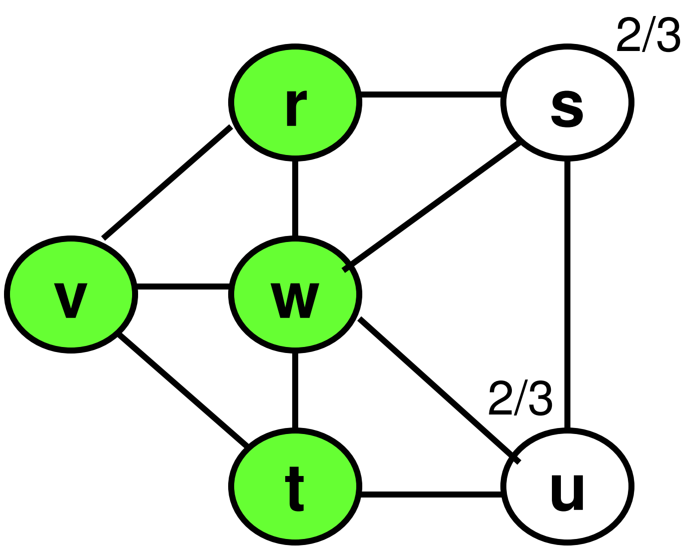
Underlying Network



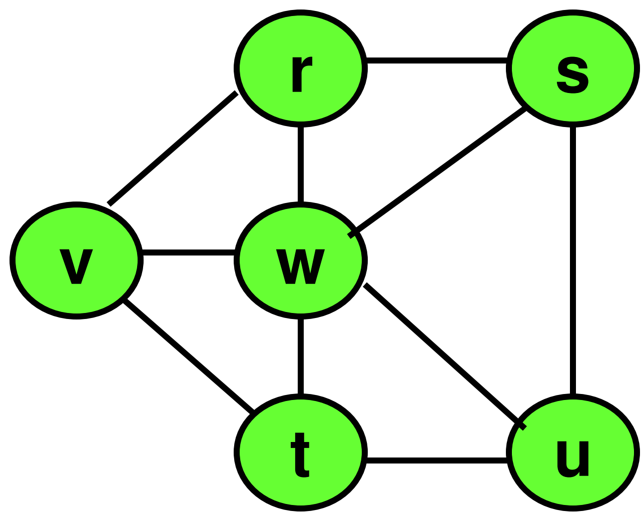
Initial Adopters

Let a = 3,
b = 2

q = 2/(2+3)
q = 2/5 = 0.4



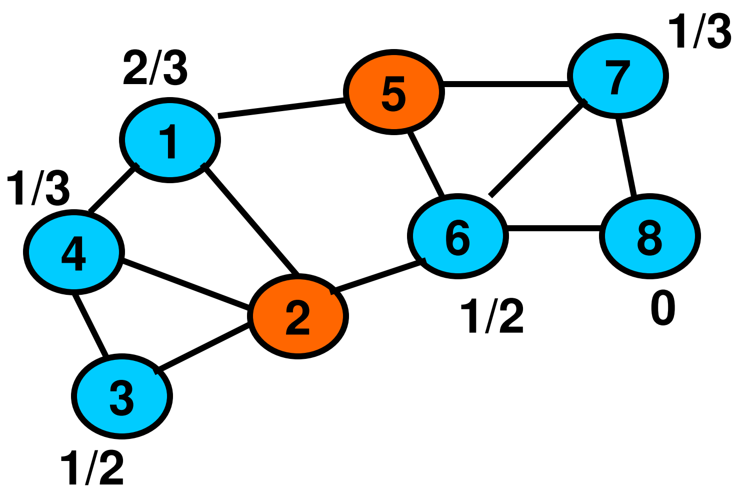
After time tick 1



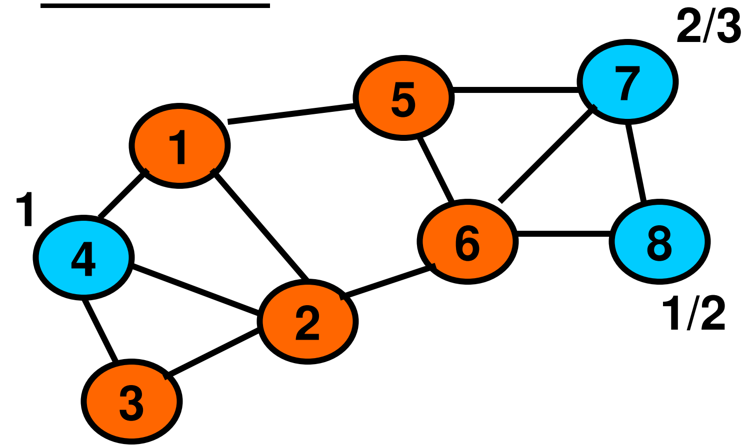
After time tick 2

Example 2: Opinion Diffusion

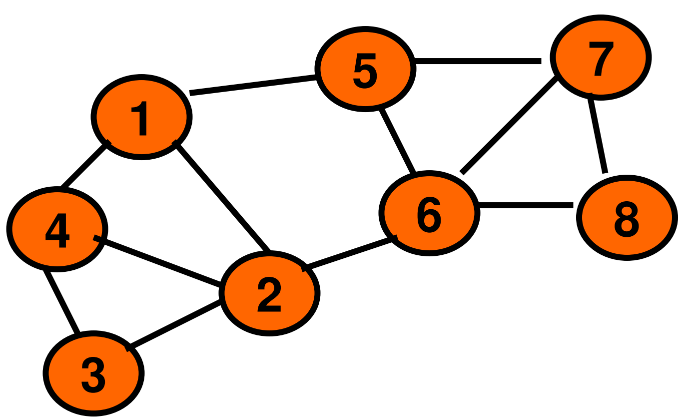
a (Red) = 1; b (Blue) = 1; q = 1/2



Iteration 1



Iteration 2



Exercises:

(1) Try this example, with only node 5 being initially Red and others Blue.
You could observe that diffusion couldn't happen at all.

(2) Try this example, with only node 2 being initially Red and others Blue.
Does diffusion happen completely or stops after awhile?

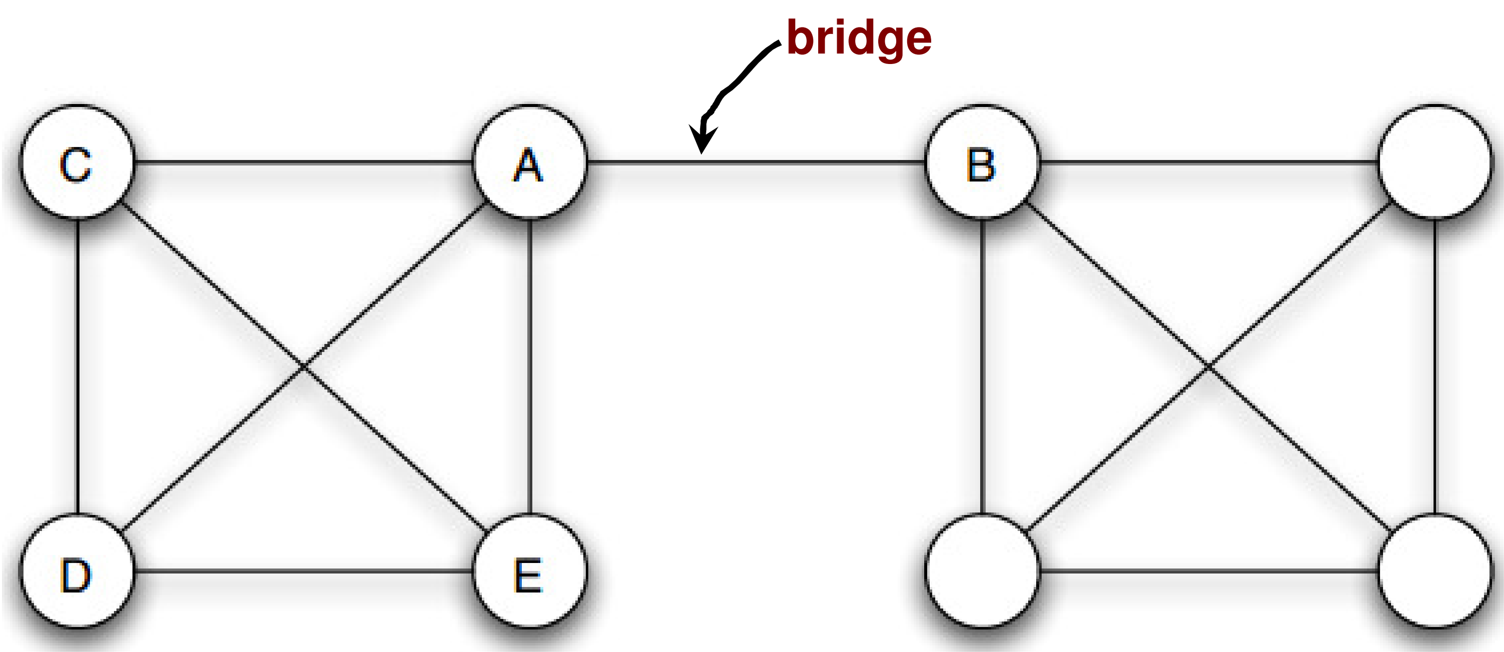
Modeling Diffusion: Nutshell, Issues

- In a nutshell:
  - An initial set of nodes adopt A while everyone else adopt B
  - Time runs forward in unit steps; in each step, each node uses the threshold rule to decide whether to switch from B to A.
  - The process stops either when every node has switched to A, or when we reached a step when no one wants to switch, at which point we say that the network has stabilized for A and B to coexist.
- Note that
  - The number of neighbors of a node that adopt A only increases with time. Hence, no node switches back to B after adopting A.

- Issues:
  - When does every node in the network eventually switches from B to A? If so, we say that a "complete cascade" has happened.
  - If this does not happen, what causes the spread of A to stop?
    - The answer to the above depends on the network structure, the value of the threshold q, and the choice of the initial adopters.

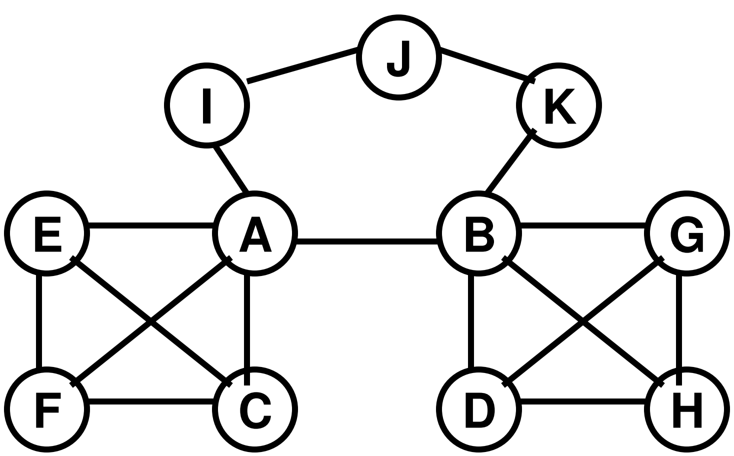
Bridges

- An edge joining two nodes A and B in a graph is a *bridge* if deleting the edge would cause A and B to lie in two different components.
  - The edge is the only route between its endpoints, the nodes A and B.
- Bridges provide nodes with access to parts of the network that are unreachable by other means.



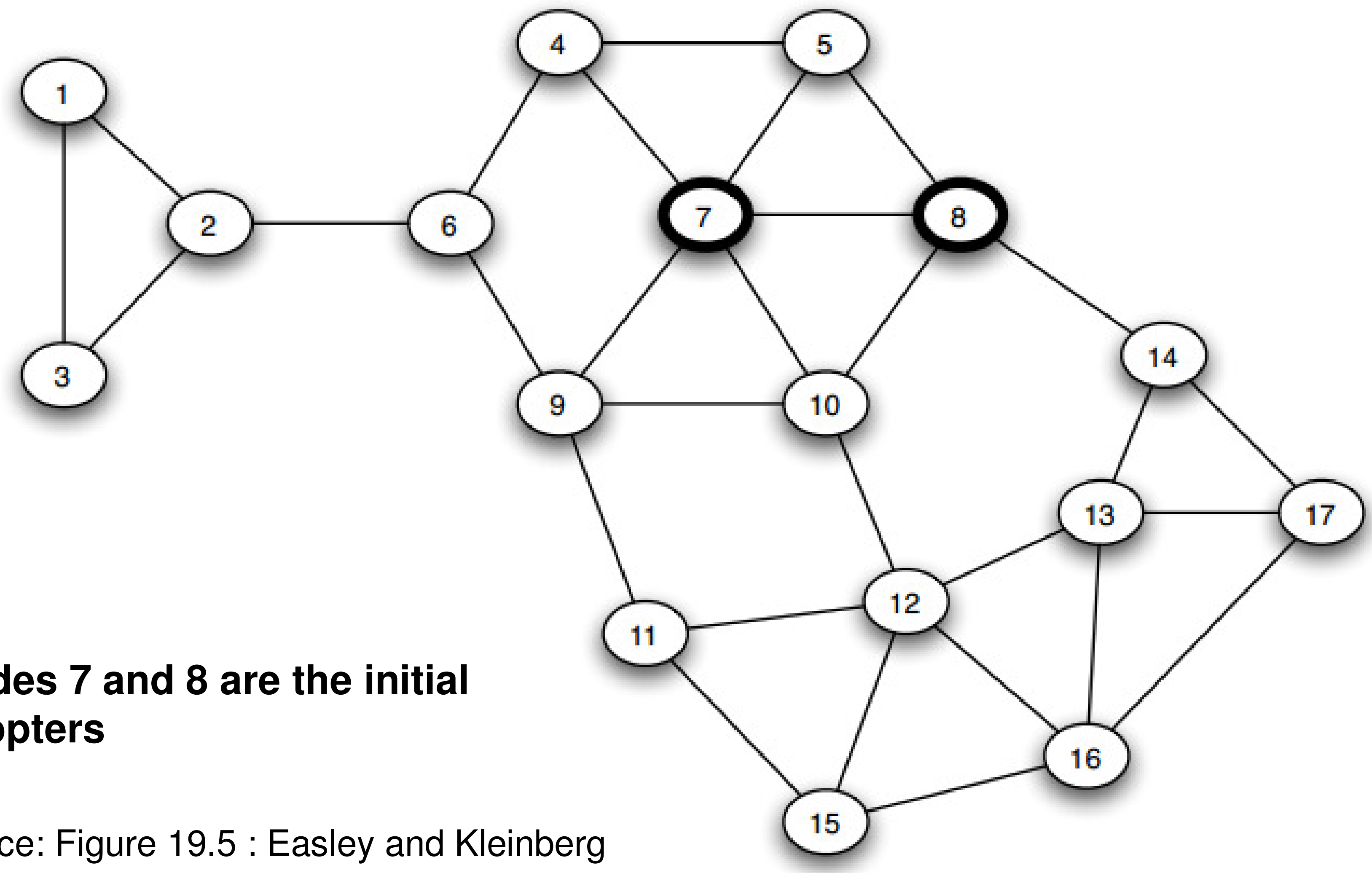
Local Bridge

- We say that an edge joining two nodes A and B in a graph is a local bridge if A and B have no common neighbors, and
  - Deleting the A – B edge would increase the distance between the two nodes to more than 2 hops.
  - However, deleting the A – B edge would not put A and B in two different components
- The <u>span of a local bridge</u> is the number of hops it takes to connect the two end nodes of the local bridge if the edge is removed from the graph.
- An <u>edge could be a local bridge only if it does not form a side of any triangle in the graph</u>.



A local bridge provides its endpoints access to parts of the network (and hence sources) of information, that they would otherwise be away from.
The closely-knit groups to which the endpoints belong to roughly know the same things that the two nodes do (and hence individually may may not be any new source of information)

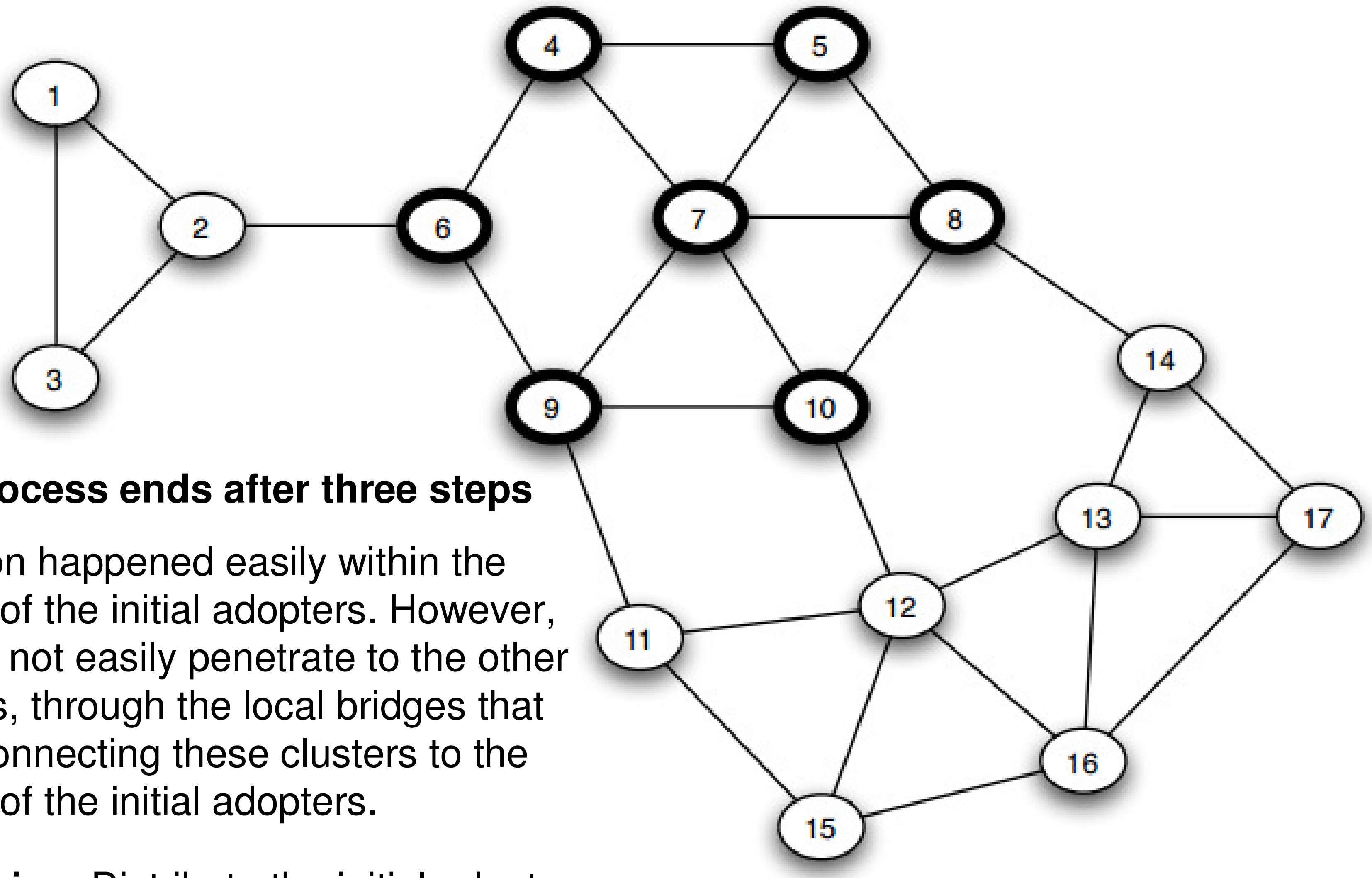
Cascade and Clusters Motivating Example



Nodes 7 and 8 are the initial adopters

Source: Figure 19.5 : Easley and Kleinberg

Cascade and Clusters: Motivating Example (2)



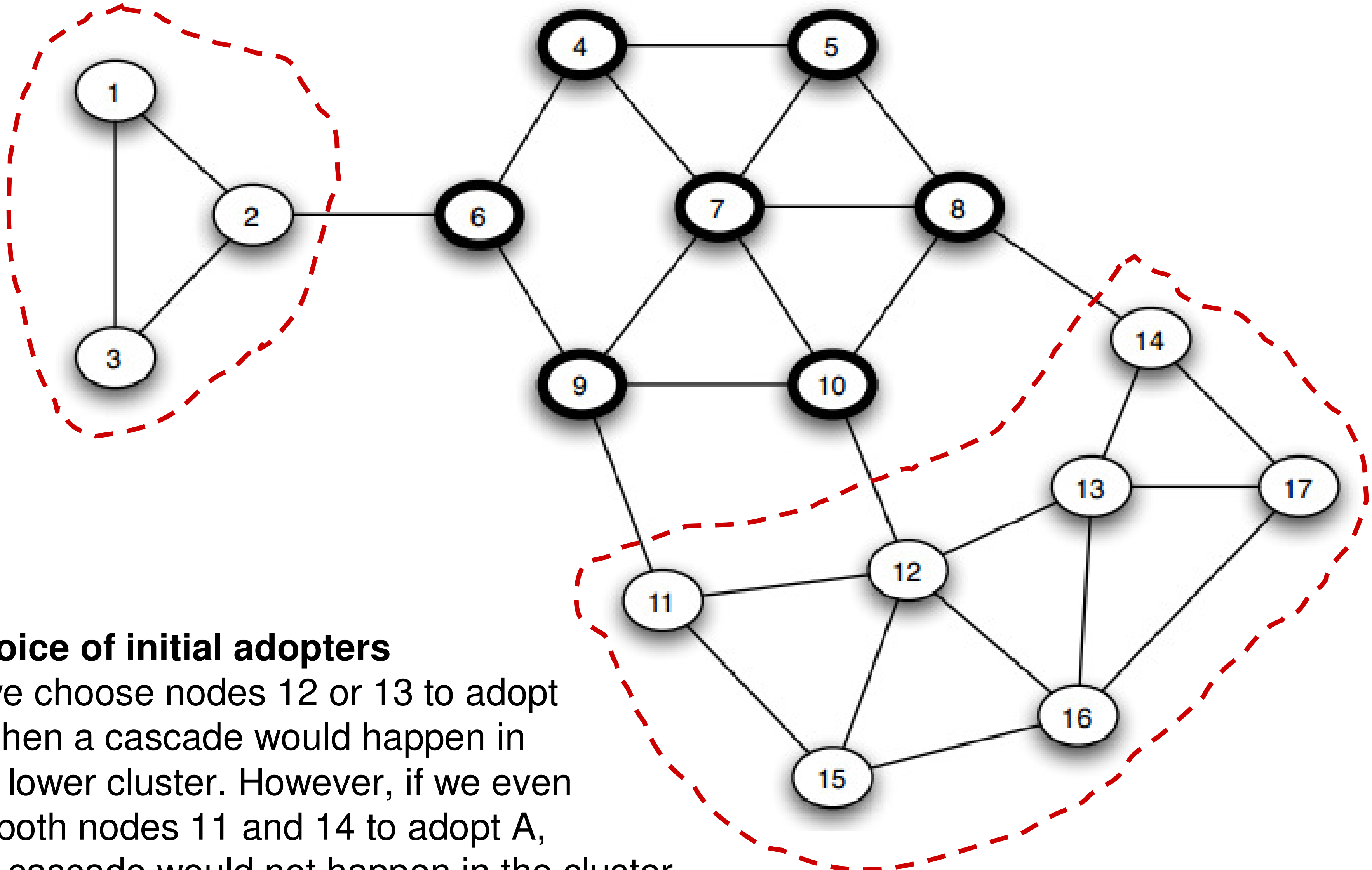
The process ends after three steps

Diffusion happened easily within the cluster of the initial adopters. However, it could not easily penetrate to the other clusters, through the local bridges that were connecting these clusters to the cluster of the initial adopters.

Strategies: Distribute the initial adopters across clusters, raise the quality of A (equivalent to lowering the threshold for adoption)

Source: Figure 19.5 : Easley and Kleinberg

Cascade and Clusters: Motivating Example (3)



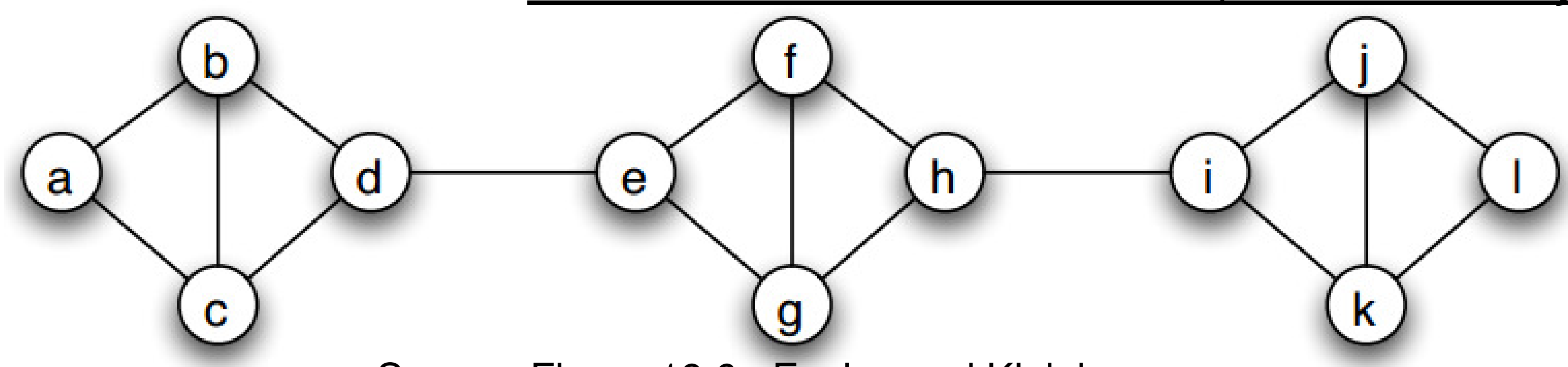
Choice of initial adopters
If we choose nodes 12 or 13 to adopt A, then a cascade would happen in the lower cluster. However, if we even let both nodes 11 and 14 to adopt A, the cascade would not happen in the cluster.

Source: Figure 19.5 : Easley and Kleinberg

What makes Diffusion to Stop?

- The answer is Homophily.
  - The spread of a new behavior could stop when it tries to break into a tightly-knit community. Thus, innovations are hard to arrive from outside <u>densely connected communities</u>.
- A “densely connected community” has the property that when you pick up a node within the community, its neighbors are also more likely to belong to that community as well.
- We say that a cluster is of density *p* if every node in the cluster has at least a *p* fraction of its neighbor nodes within the cluster.

<u>A Collection of 4-node Clusters, each of density 2/3</u>



Source: Figure 19.6 : Easley and Kleinberg

Theorems: What makes Diffusion Stop?

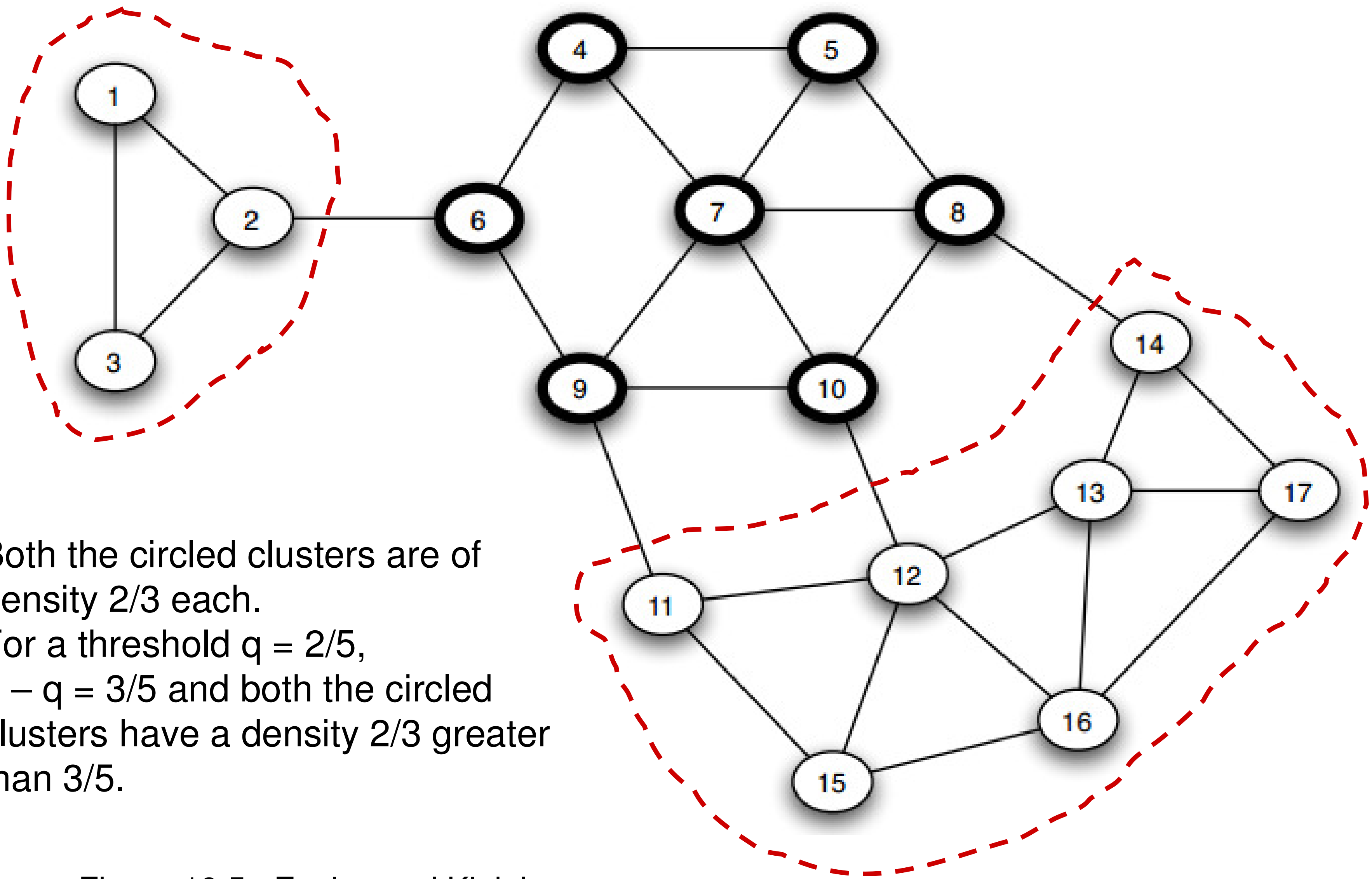
- Let S be the set of initial adopters.
- Let the remaining network refer to the portion of the network consisting of all nodes other than the initial adopters.
- Consider a set S of initial adopters of behavior (or innovation) A, with a threshold of *q* for nodes in the remaining network to adopt A.

- Theorem 1: If the remaining network contains a cluster of density greater than 1-q, then the set of initial adopters will not cause a complete cascade.

- Theorem 2: Whenever a set of initial adopters does not cause a complete cascade with threshold q, then the remaining network must contain a cluster of density greater than 1-q.

- **Conclusion: A set of initial adopters can cause a complete cascade at threshold q if and only if the remaining network contains no cluster of density greater than 1-q.**

Illustrative Example: Cluster Density



Both the circled clusters are of density 2/3 each.
For a threshold q = 2/5,
1 – q = 3/5 and both the circled clusters have a density 2/3 greater than 3/5.

Source: Figure 19.5 : Easley and Kleinberg

1: Clusters are Obstacles to Cascades

- Given that *q* is the threshold fraction of neighbors (that have already adopted the desired behavior, say A) a node should have to adopt A, if there is a cluster in the remaining network of density greater than 1-q, then A cannot penetrate into the cluster.

- We will prove this by contradiction.
- Consider a node u to be the first node of the cluster (of density greater than 1-*q*) to adopt A and that no other node in the cluster has yet adopted A.
- All of node u's neighbors that contributed for the threshold *q* should have come from outside the cluster. Hence, the fraction of neighbors of node u that are inside the cluster could be at most 1-*q*.
- But, since the cluster is of density greater than 1-*q*, each node in the cluster should have the fraction of neighbor nodes within the cluster to be greater than 1-*q*. This contradicts that the cluster is of density greater than 1-*q*.
- Hence, for a node u in the cluster to adopt A, it could only have a fraction of neighbors less than *q* outside the cluster, which makes it not qualified to adopt A.

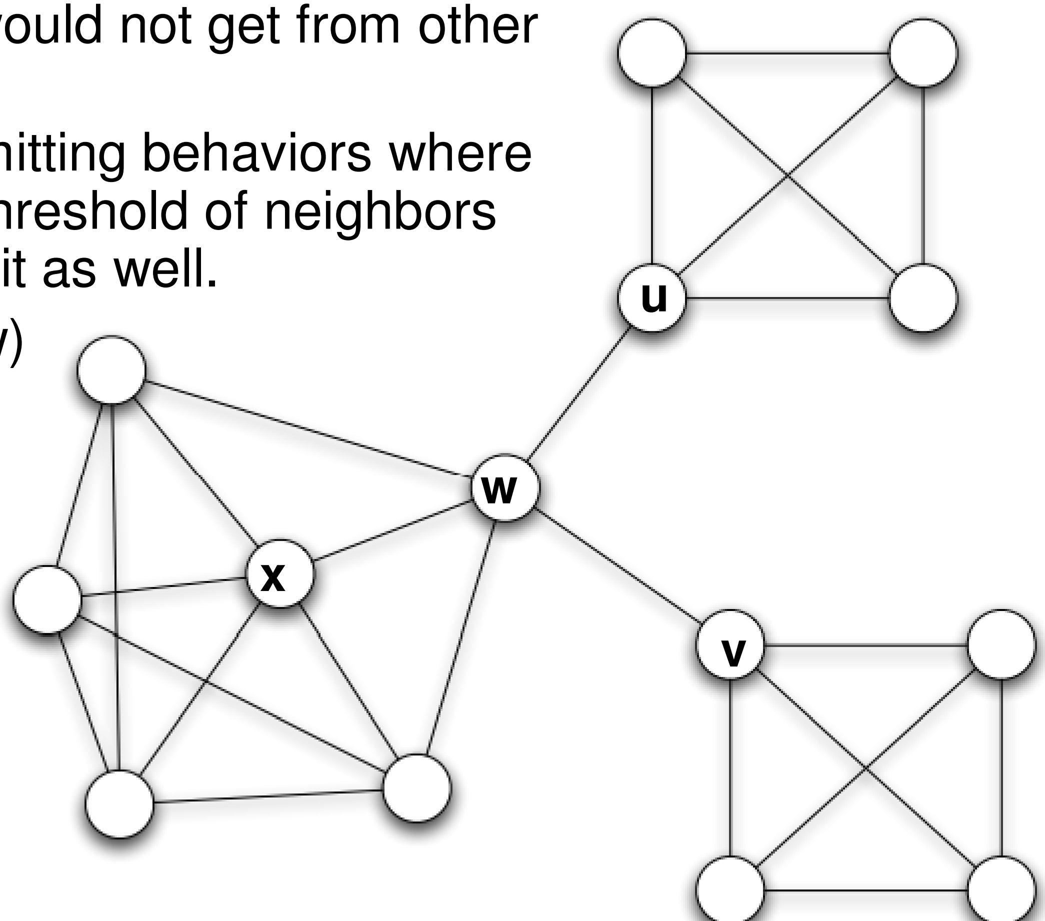
2: Clusters are the <u>only</u> Obstacles to Cascades

- Let R be the set of nodes in the remaining network that have not yet adopted the desired behavior A.
- Let us assume that diffusion has stopped without creating a complete cascade, leaving a set R of nodes that stayed with the alternate (original) behavior B itself.
- For every node u in this set R, the fraction of neighbors who adopted A is less than q (otherwise, u would have adopted A). Hence, the fraction of neighbors who stayed with B (and hence lie in the set R) is greater than 1-q.
- The above observation holds good for every node in R. Hence, R is a cluster of density larger than 1-q.

- **Conclusion: A set of initial adopters can cause a complete cascade at threshold q if and only if the remaining network contains no cluster of density greater than 1-q.**

Bridges and Local Bridges: Tradeoff

- Bridges, local bridges and structural holes are powerful ways to convey awareness of new things that the nodes in a cluster would not get from other edges.
- But, they are weak in transmitting behaviors where one needs to see a higher threshold of neighbors doing it before a node does it as well.
- Nodes u and v (learn from w) and have strong informational advantages over other members of their respective tightly-knit communities – but for behaviors with higher thresholds they will still want to align themselves with others in their own community.



Source: Figure 19.11: Easley and Kleinberg

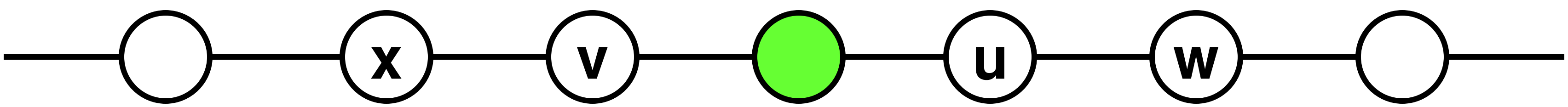
Cascades, Clusters and Local Bridges

- Cascades and clusters truly are natural opposites: If a network has clusters connected with weak ties and the initial adopters are not adequately spread across the different clusters, clusters block the spread of cascades.

- A world-spanning system of weak ties in a global network is able to spread awareness of information with remarkable speed, political mobilization moves more sluggishly, needing to gain momentum within neighborhoods and small communities.
  - Thresholds provide a possible reason: social movements tend to be inherently risky undertakings, and hence individuals tend to have higher thresholds for participating; under such conditions, local bridges that connect very different parts of the network are less useful.

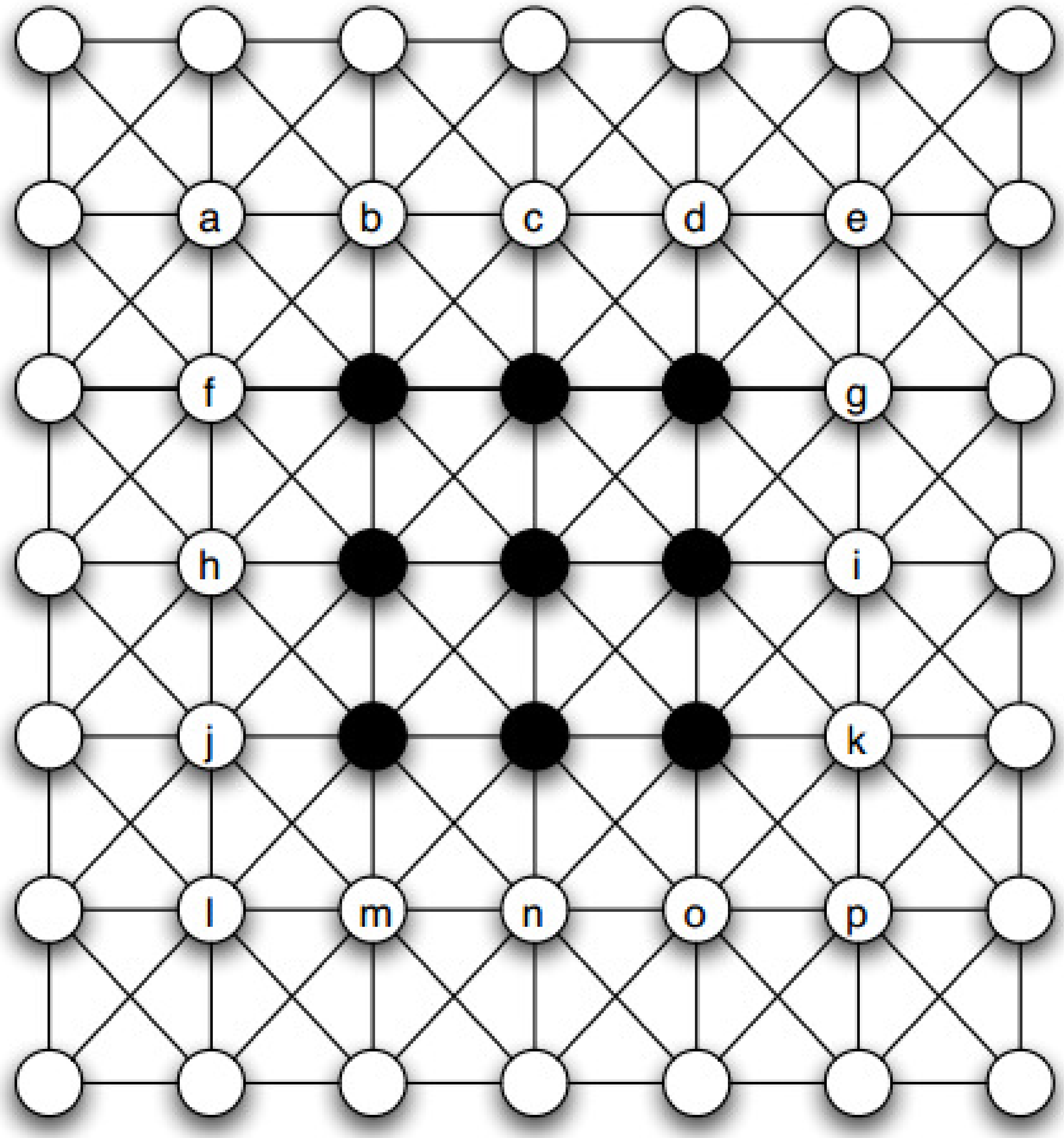
Cascade Capacity of a Network

- The cascade capacity of a network is the maximum threshold value (for every node in the network) at which any “small” finite set of initial adopters can cause a complete cascade.
- We will analyze cascade capacity using the notion of infinite networks, where there is an infinite set of nodes; but, each individual node is only connected to a finite number of other nodes.



For threshold values $q \le ½$, the above infinite network can run into a complete cascade, starting with the initial adopters (even one initial adopter would be sufficient) as shown above.
In the first round, u and v accept A; in the second round, x and w accept behavior A, and so on. For $q > ½$, the network would <u>not</u> run into a complete cascade.

Hence, the **cascade capacity** of the above network is **½.**



Example (2)

Consider an infinite grid
Where each node is
Connected to 8 of its
Nearest neighbors.
The black nodes are the
Initial adopters.

The nodes c, h, n, i adopt
A in the first round.
Nodes b, d, f, j, m, o, g, k
adopt A in round 2.
Nodes a, e, l, p adopt A in
round 3.
This spreads to the next
square ring of nodes.

The **Cascade Capacity**
of the grid network is **3/8**.
i.e., threshold q ≤ 3/8
leads to a complete
cascade.

Extensions of the Basic Cascade Model: Heterogeneous Thresholds

- Each node *v* in the network has its own (a different) value for the threshold q_v.
  - The payoff values for a node *v* are a_v and b_v for associating with behaviors A and B respectively.
- A node *v* adopts behavior A only if the fraction of neighbor nodes in the network

Threshold for node v to adopt

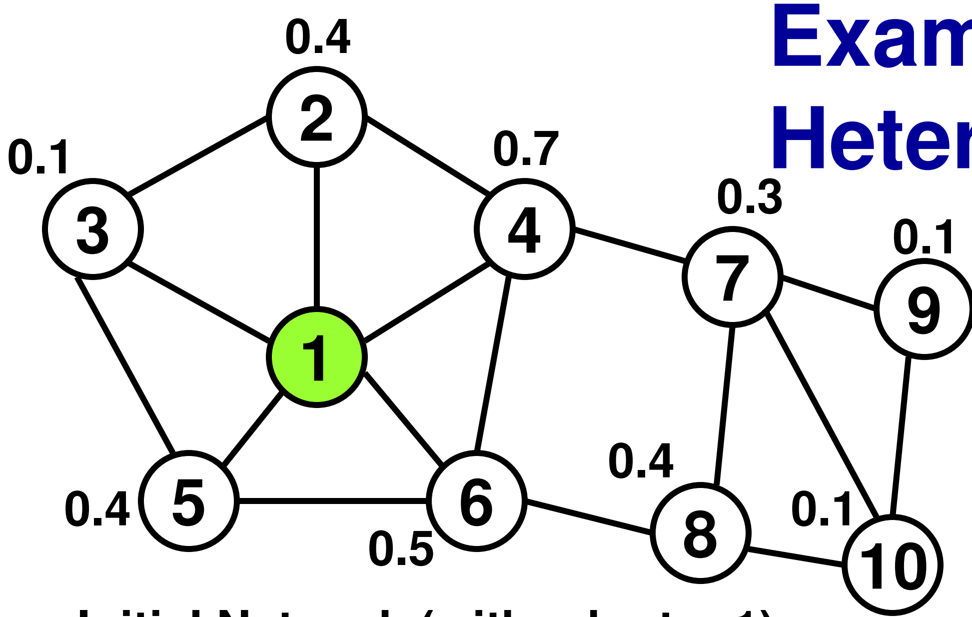
$$p \geq \frac{b_v}{a_v + b_v} \qquad q_v = \frac{b_v}{a_v + b_v}$$

To effectively realize complete cascade, it is important to not only identify influential nodes (nodes with larger degrees) as potential initial adopters, it is also imperative to identify influential nodes that have access to easily influenceable people.

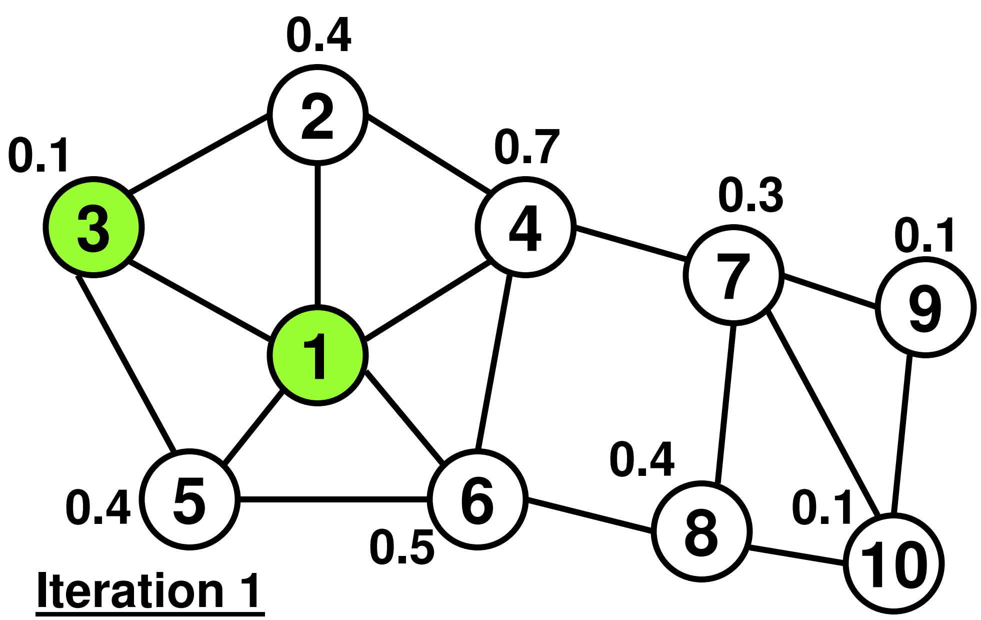
A blocking cluster in a network is a set of nodes such that for every node *v* in this Set, there exists more than 1-q_v fraction of neighbor nodes that are also in this set.

For a given set of initial adopters and node thresholds, complete cascade is possible only if the remaining network does not contain any blocking cluster.

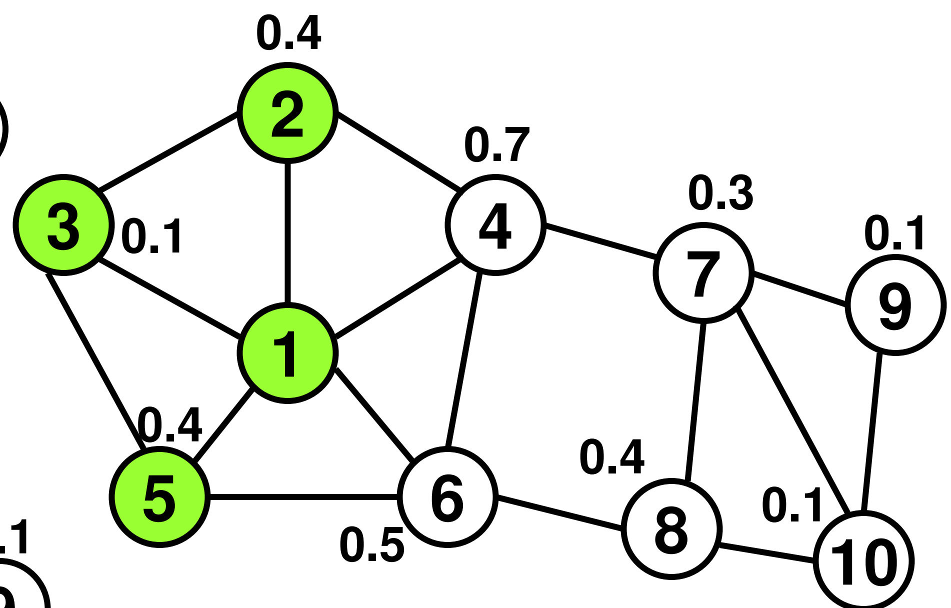
Example: Heterogeneous Thresholds



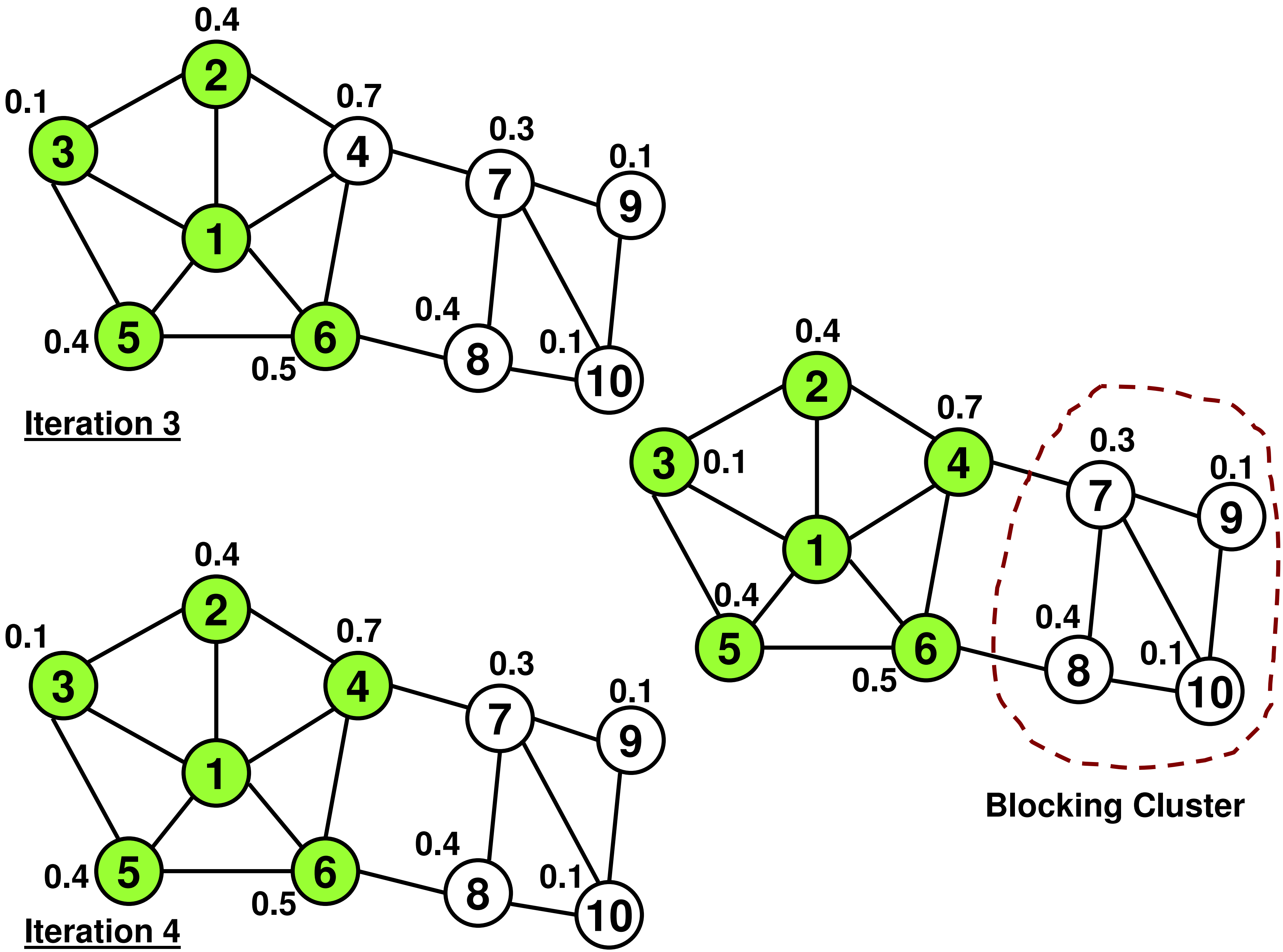
Initial Network (with adopter 1)



Iteration 1



Iteration 2


Iteration 3
Iteration 4
Blocking Cluster