Processi di diffusione parte 1

Capitolo 19 del testo

Indice

1 Azioni e relazioni

La presenza di una rete influenza il comportamento degli individui che la popolano talvolta, modificando il loro comportamento e questo lo avevamo già visto con l'esperimento di Granovetter perché gli individui, proprio in virtù della presenza di una rete, interagiscono fra loro magari, soltanto osservando il comportamento gli uni degli altri.

Molte delle nostre interazioni avvengono a livello locale piuttosto che globale: ci interessa il comportamento degli individui con i quali siamo in relazione, piuttosto che il comportamento dell'intera popolazione.

Perciò, gli individui che possono influenzare il nostro comportamento sono quelli con i quali siamo in relazione.

Cerchiamo, ora di entrare nel merito della natura di questa "influenza di rete".

2 Omofilia

Abbiamo già incontrato il fenomeno della chiusura triadica, ossia, la tendenza a che si stabiliscano relazioni fra gli individui che hanno una relazione forte con uno stesso individuo.

Questo fenomeno è strettamente connesso al concetto di omofilia che si esplica in due direzioni:

- in un senso, è la tendenza a connetterci con gli individui che ci assomigliano
  - ad esempio, un individuo tende a non stabilire relazioni con chi ha idee politiche opposte alle sue
  - ma anche: se amo il mare, diventerò amico con la gente che incontro al mare (e che, dunque, come me ama il mare)

E questo per numerous ragioni:

- selezione: tendo ad essere amico a chi mi somiglia, perché sto meglio con chi mi somiglia che non con chi è diverso da me
- opportunità: se amo il mare, difficilmente, nelle mie vacanze, incontrerò gli amanti della montagna

Il fenomeno della chiusura triadica è strettamente connesso al concetto di omofilia che si esplica in due direzioni:

- Nell'altro senso, è la tendenza diventare simili agli individui con i quali siamo in relazione
  - ad assumere i loro stessi gusti
  - ad adeguarci ai loro comportamenti
  - e a diventare amici dei loro amici!

La motivazione soggiacente a questa tendenza è l'esigenza di ridurre la tensione sociale: non si può stare sempre a litigare se andare a vedere un thriller o una commedia... prima o poi, uno dei due amici si farà piacere il genere che piace all'altro.

Ma c'è anche una motivazione assolutamente razionale: se tutti i miei amici acquistano un nuovo sistema operativo, io, che li stimo (sono miei amici!), assumo che sia una buona idea e lo acquisto a mia volta e, inoltre, se mi tenessi il mio vecchio sistema operativo, magari non potrei più scambiare file con loro...

3 Processi di diffusione

Se tutti i miei amici acquistano un nuovo sistema operativo, io lo acquisto a mia volta quando valuto che questo acquisto migliorerà la mia vita. Poi, magari, altri amici lo acquisteranno e il nuovo sistema operativo si diffonderà nella rete.

Il processo di diffusione delle innovazioni è stato studiato in sociologia già a partire dalla metà del secolo scorso:

- Ryan e Gross (1943) hanno osservato i processo di adozione di nuovi semi ibridi di mais da parte di un gruppo di agricoltori in Iowa
- sebbene la maggior parte degli agricoltori veniva a conoscenza dei nuovi semi tramite le comunicazioni dei venditori,
- a parte un numero limitato di agricoltori,
- la maggior parte degli agricoltori iniziava a usare i nuovi semi solo dopo aver osservato che un certo numero di vicini / conoscenti / amici li stava utilizzando

3.1 Modellare il processo di diffusione

Vogliamo modellare il processo di diffusione in una rete e, per farlo dobbiamo stabilire le regole in base alle quali un nodo decide di cambiare comportamento, opinione, prodotto.

Intanto, definiamo un modello di decisioni individuali, ossia, non vi è coalizzazione di gruppi di nodi per prendere collettivamente la stessa decisione, nel quale le scelte dei nodi sono guidate da motivazioni di puro interesse personale: la spinta a cambiare è tanto maggiore quanto maggiore è il vantaggio che si prevede che deriverà dal cambiamento.

Assumiamo che nella rete sia stabilizzato un certo stato delle cose B (ad esempio, tutti gli individui nella rete utilizzano lo stesso sistema operativo B) e che, ad un certo istante, alcuni individui cambino il loro stato in A (ad esempio, viene lanciato sul mercato un nuovo sistema operativo A, e, a fini promozionali, viene dato gratuitamente ad un certo insieme di individui).

In quali casi un individuo decide di cambiare il proprio stato da B ad A? Assumendo che chi è nello stato A non torni mai in B.

3.2 Network Coordination Game

Modelliamo il processo di diffusione

- individuale e basato sul vantaggio personale

mediante un Network Coordination Game.

Sia (u, v) un arco della rete. Assumiamo che il beneficio reciproco di adottare A o B sia quello illustrato in tabella:

V \ u	**A**	**B**
A	a, a	0,0
B	0,0	b,b

Tabella 1: Beneficio reciproco

Se u e v adottano entrambi A allora entrambi hanno un beneficio pari ad a. Se u e v adottano entrambi B allora entrambi hanno un beneficio pari a b. Altrimenti nessuno dei due ha alcun beneficio (dalla reciproca relazione).

3.3 Modello lineare e Soglia di adozione

Ma un nodo nella rete ha, in generale, più di un vicino. Cosa accade quando qualcuno dei vicini di un nodo u è nello stato A e qualcun altro è nello stato B?

Semplicemente, detti n_A il numero di vicini di u nello stato A e n_B il numero di vicini di u nello stato B:

- se u rimane nello stato B ha un beneficio pari a $b \cdot n_B$, se u passa allo stato A ha un beneficio pari a $a \cdot n_A$,
- u rimane nello stato B se $b \cdot n_B > a \cdot n_A$
- u passa allo stato A se $a \cdot n_A \geq b \cdot n_B$
- (si osservi che, a parità di beneficio, u passa ad A (l'innovazione è preferibile al vecchio stato))

ossia, poiché $n_B = |N(u)| - n_A$, u passa allo stato A se: $a \cdot n_A \geq b(|N(u)| - n_A)$ ossia, se $\frac{n_A}{|N(u)|}a \geq \frac{|N(u)|-n_A}{|N(u)|}b$ ossia, detto $p_A = \frac{n_A}{|N(u)|}$ se $p_A a \geq (1 - p_A)b$.

u passa allo stato A se, detta p_A la frazione dei vicini di u che è in A, vale che:

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

chiamiamo

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

la soglia di adozione di A.

- Quando q è molto piccolo, occorrono pochi vicini nello stato A per indurre un nodo a cambiare stato
- e q è molto piccolo quando a è molto più grande di b
- ossia, quando lo stato A è molto migliore dello stato B
- Quando $a = b$ occorrono almeno la metà dei vicini nello stato A per indurre un nodo a cambiare stato
- e questo accade quando lo stato A è confrontabile con lo stato B
- Infine, quando a è molto più piccolo di b, occorrono molti vicini nello stato A per indurre un nodo a cambiare stato
- e questo accade quando lo stato A è peggiore dello stato B
- e quindi è faticoso / costoso / rischioso adottare A

4 Game e configurazioni di equilibrio

Cerchiamo, ora di capire, se e quali configurazioni di equilibrio ha il network coordination game che abbiamo appena introdotto: configurazioni, cioè, in cui nessun nodo cambia il proprio stato da B ad A (perché, ricordiamo, assumiamo che nessun nodo "torni indietro" cambiando il proprio stato da A a B).

Osserviamo che esistono almeno due configurazioni di equilibrio: quelle banali

- quando A non è stato introdotto nella rete, cosicché tutti i nodi sono nello stato B
- quando, dopo che A è stato introdotto nella rete, tutti i nodi sono passati nello stato A

Perché, una volta che A viene introdotto nella rete, esso inizia a diffondersi:

- Quando termina il processo di diffusione?
- Riesce sempre a raggiungere tutti i nodi?
- Oppure, talvolta, la diffusione si blocca prima di aver raggiunto tutti i nodi, in configurazioni di equilibrio intermedie?
- E, in questo caso, perché si blocca?

Cerchiamo di capire, innanzi tutto, con qualche esempio.

4.1 Esempi di processi di diffusione

Lo stato A, viene forzato all'inizio sui nodi v e w. In questo caso, $a = 3$ e $b = 2$. A è migliore di B. Quindi, $q = \frac{2}{3+2}$. Ossia, per adottare A, un nodo deve avere i 2/5 dei vicini nello stato A. Perciò, un nodo dopo l'altro, tutti i nodi della rete adottano A.

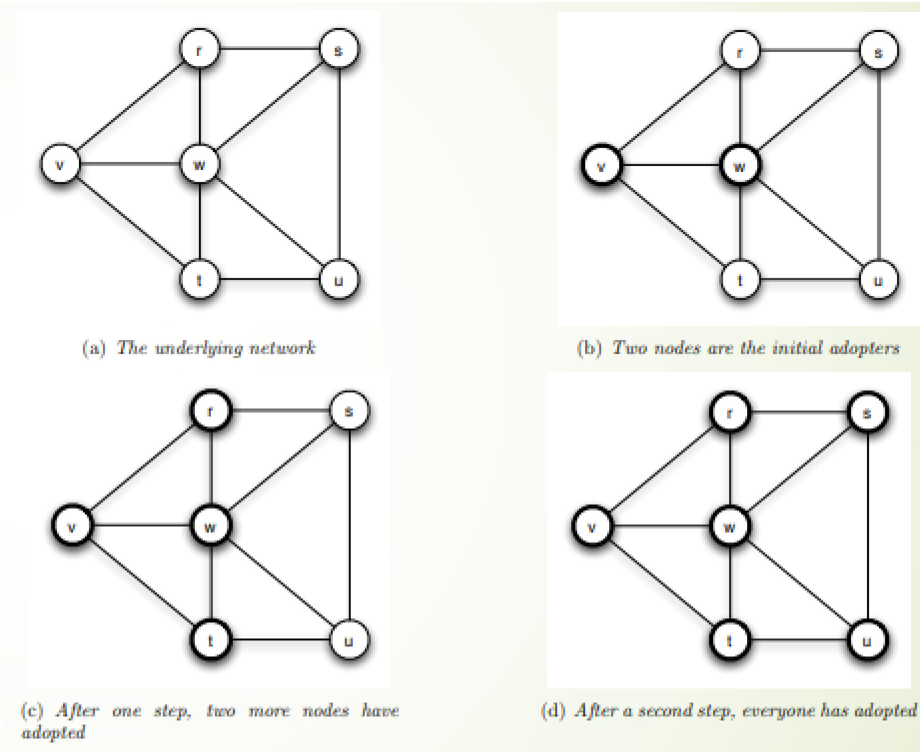
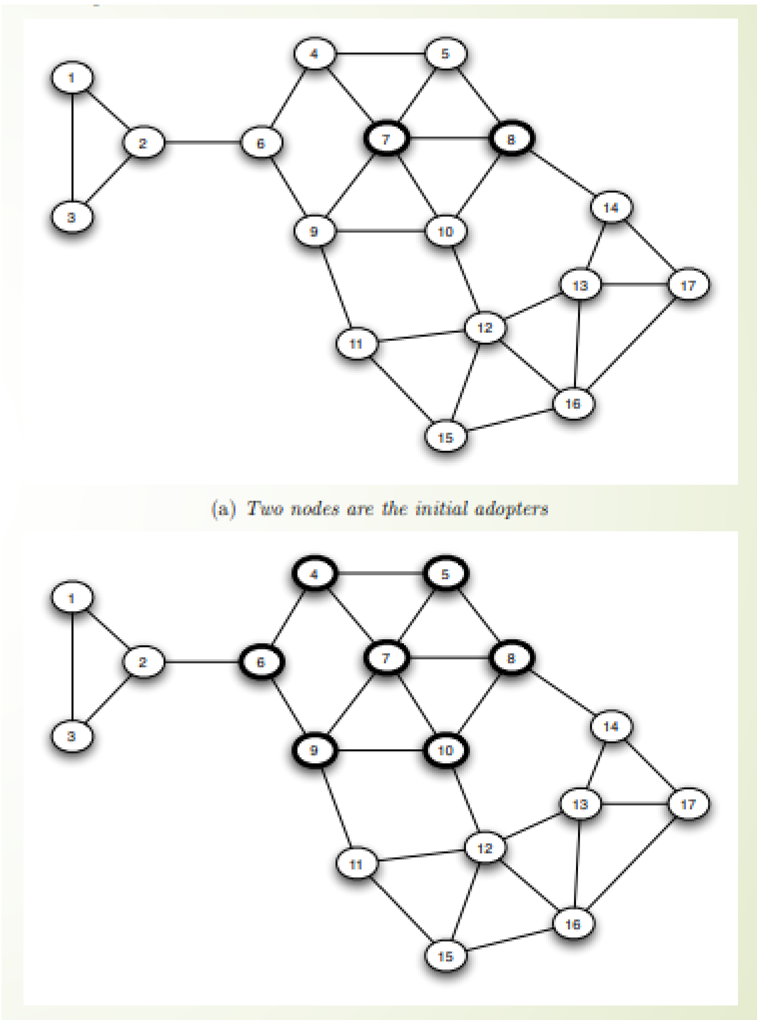


Figure 19.3: Starting with v and w as the initial adopters, and payoffs $a = 3$ and $b = 2$, the new behavior A spreads to all nodes in two steps. Nodes adopting A in a given step are drawn with dark borders; nodes adopting B are drawn with light borders. (a) The underlying network (b) Two nodes are the initial adopters (c) After one step, two more nodes have adopted (d) After a second step, everyone has adopted.

5 Diffusione e cascate complete

Altro esempio:



(a) *Two nodes are the initial adopters*

(a) Two nodes are the initial adopters (7, 8). (b) After one step (con $a = 3, b = 2$), A non riesce a raggiungere i nodi fuori dall'esagono.

- **Caso 1)** se $a = 3$ e $b = 2$, ossia $q = \frac{2}{5}$ (è il caso illustrato in figura).
- ossia, per adottare A, un nodo deve avere i 2/5 dei vicini nello stato A.
- In questo caso, non tutti i nodi della rete adottano A:
- come si vede in figura, A non riesce a raggiungere i nodi fuori dell'esagono.
- **Caso 2)** se $a = 4$ e $b = 2$, ossia $q = \frac{2}{6} = \frac{1}{3}$
- dopo aver raggiunto tutti i nodi dell'esagono
- A è adottato da 2, 11 e 14
- poi da 1, 3, 12, 13, 17
- e infine da 15 e 16
- Tutti i nodi hanno adottato A!

Dai due esempi possiamo trarre una serie di conclusioni:

- intanto, non sempre tutti i nodi si adeguano all'innovazione
- poi, se questo accade, possiamo aumentare il beneficio derivante dall'adozione di A per indurre tutti i nodi ad adottarlo
  - ad esempio, possiamo abbassare il costo dell'acquisizione di A
  - aumentandone l'attrattiva!

Ma dovrebbe essere chiaro anche che l'eventualità che tutti i nodi arriveranno ad adottare A dipende dai nodi che scegliamo per forzare lo stato A all'inizio del processo:

- dal loro numero
- ma anche dalla loro posizione all'interno della rete

ad esempio, mantenendo $a = 3$ e forzando A su 4 nodi (invece che su 2):

- se forziamo A sui nodi 7, 8, 2 e 12, tutti i nodi adotteranno A
- se forziamo A sui nodi 7, 8, 2 e 14 o, peggio ancora, 7, 8, 4 e 5, non tutti i nodi adotteranno A

Come mai?

5.1 Definizioni

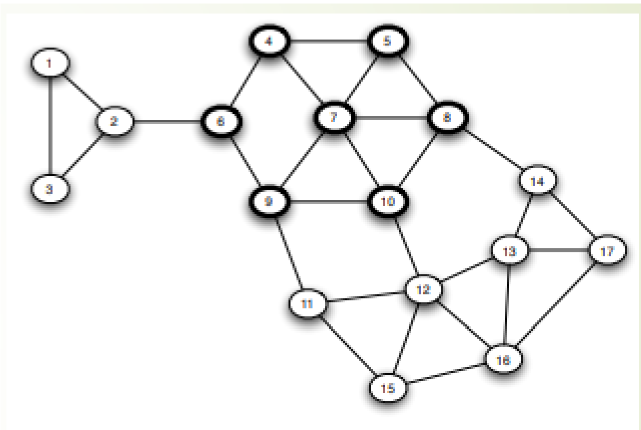
Prima di procedere, abbiamo bisogno di qualche definizione e di qualche notazione. Inizialmente, lo stato A viene forzato su un certo insieme V_0 di nodi che chiameremo iniziatori. Come abbiamo visto, una volta che A viene introdotto nella rete, esso inizia a diffondersi, ossia, ha inizio un processo di diffusione che procede in una sequenza di passi discreti:

- al passo 1, un insieme V_1 di vicini di nodi in V_0 adotta A
- al passo 2, un insieme V_2 di vicini di nodi in $V_0 \cup V_1$ adotta A
- ... al passo t, un insieme V_t di vicini di nodi in $V_0 \cup V_1 \cup \cdots \cup V_{t-1}$ adotta A ...

ossia, indichiamo con V_i l'insieme dei nodi che adottano A al passo i.

Viene generata una **cascata completa** se ad un certo passo t tutti i nodi hanno adottato A, ossia se esiste $f \geq 0$ tale che $\bigcup_{0 \leq i \leq f} V_i = V$.

6 Cascate e clusters



L'esagono centrale (nodi 4, 5, 6, 7, 8, 9, 10).

Se ripensiamo all'esempio, sembra che l'innovazione abbia difficoltà a uscire dall'"esagono centrale" - i nodi 4, 5, 6, 7, 8, 9, 10, che appare come un gruppo coeso di nodi, in qualche modo, in qualche senso, una comunità (vediamo in quale senso).

Un **cluster di densità** p è un sottoinsieme di nodi $V' \subseteq V$ tale che la frazione di vicini che ogni suo nodo ha in V' è almeno p:

$$\forall u \in V' \left[\frac{|N(u) \cap V'|}{|N(u)|} \geq p \right]$$

Il sottoinsieme $\{4, 5, 6, 7, 8, 9, 10\}$ è un cluster di densità $p = \frac{2}{3}$.

Osservazione: l'unione di due cluster di densità p è ancora un cluster di densità p. Ad esempio, poiché anche $\{1, 2, 3\}$ è un cluster di densità $\frac{2}{3}$, allora $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ è un cluster di densità $\frac{2}{3}$.

ESERCIZIO: provare a dimostrare questa proprietà.

6.1 Teorema 1

Notazione: sia $G = (V, E)$ un grafo e sia $V' \subseteq V$ indichiamo con $G - V'$ il grafo ottenuto rimuovendo da G tutti i nodi in V' e tutti gli archi incidenti su nodi in V'.

Teorema 1. sia $G = (V, E)$ un grafo e siano $V_0 \subseteq V$ l'insieme di iniziatori e q la soglia di adozione di A: V_0 non genera una cascata completa se e soltanto se $G - V_0$ contiene un cluster di densità maggiore di $1 - q$.

Dimostrazione (se V_0 non genera...)

Se V_0 non genera una cascata completa, allora, esistono nodi che non adottano A. Ricordiamo che indichiamo con V_i i nodi che adottano A al passo i: sia t il passo tale che $V_t \neq \emptyset$ e $V_{t+1} = \emptyset$. Ossia, $t + 1$ è il primo passo in cui A non si diffonde più. Poniamo $V_A = \bigcup_{0 \le i \le t} V_i$: V_A contiene tutti e soli i nodi che adottano A. Poiché esistono nodi che non adottano A, allora, $V - V_A \neq \emptyset$.

Poiché i nodi in $V - V_A$ non adottano A, allora per ogni $v \in V - V_A$:

$$\frac{|N(v) \cap V_A|}{|N(v)|} < q$$

e, quindi, poiché

$$\frac{|N(v) \cap V_A|}{|N(v)|} + \frac{|N(v) \cap (V - V_A)|}{|N(v)|} = 1$$

allora

$$\frac{|N(v) \cap (V - V_A)|}{|N(v)|} > 1 - q$$

ossia, $V - V_A$ è un cluster di densità maggiore di $1 - q$ e $V - V_A$ è contenuto in $G - V_0$.

Dimostrazione (se $G - V_0$ contiene...)

Se $G - V_0$ contiene un cluster di densità maggiore di $1 - q$, cioè, esiste $C \subseteq V - V_0$ tale che, per ogni $v \in C$:

$$\frac{|N(v) \cap C|}{|N(v)|} > 1 - q$$

Supponiamo per assurdo che si generi una cascata completa: allora i nodi in C, prima o poi, adotteranno A. Sia t il primo passo tale che $V_t \cap C \neq \emptyset$ ossia, per $i < t$, $V_i \cap C = \emptyset$ e esiste $u \in V_t \cap C$. Poniamo $V' = \bigcup_{0 \le i \le t-1} V_i$: V' sono tutti i nodi che al passo $t - 1$ sono nello stato A e V' non contiene nodi di C. Allora:

a) poiché C è un cluster di densità $> 1 - q$ e $u \in C$, allora $\frac{|N(u) \cap C|}{|N(u)|} > 1 - q$

b) poiché $u \in V_t$, allora $\frac{|N(u) \cap V'|}{|N(u)|} \ge q$

ossia, poiché $V' \cap C = \emptyset$,

$$1 = \frac{|N(u)|}{|N(u)|} \ge \frac{|N(u) \cap (C \cup V')|}{|N(u)|} = \frac{|N(u) \cap C|}{|N(u)|} + \frac{|N(u) \cap V'|}{|N(u)|} > (1 - q) + q = 1$$

$1 > 1$, un assurdo.

7 Il ruolo dei weak ties

Il teorema appena dimostrato mette in luce un nuovo aspetto della dicotomia strong ties / weak ties:

- le innovazioni si diffondono con relativa facilità all'interno dei cluster
  - ossia, quando viaggia lungo strong ties l'innovazione ha un impatto significativo sui nodi che raggiunge
- invece, incontrano difficoltà ad uscire dai cluster

– ossia, quando viaggia lungo weak ties l'innovazione ha un impatto debole sui nodi che raggiunge

Perciò, possiamo concludere che:

- mentre l'esperimento di Granovetter ha permesso di mettere in luce la forza dei weak ties
  – in quanto fonte di vantaggi informativi
- lo studio dei processi di diffusione ne evidenzia la debolezza
  – in quanto ostacolo alla diffusione

8 Clusters e Marketing virale

Il teorema appena dimostrato mostra che i cluster ostacolano la diffusione virale di un'innovazione.

Ci domandiamo, ora, se e come è possibile superare lo stallo nel quale, a causa della presenza di cluster sufficientemente densi, uno stato A smette di diffondersi.

- Possiamo scegliere gli iniziatori in posizioni tali che ciascun cluster contenga almeno un iniziatore
  – e questa è la lezione che abbiamo imparato dal teorema
- Oppure, come abbiamo già visto, è possibile aumentare l'appetibilità di A
  – per esempio, abbassando il suo prezzo di vendita
  – così da aumentare il beneficio a della sua adozione e diminuire la soglia di adozione

Ma, dal punto di vista del venditore, abbassare il prezzo di vendita non è una strategia che adotterebbe con molto piacere...

Allora, ci domandiamo: quanto alto può essere mantenuto il prezzo (quanto bassa può essere l'attrattiva) di A perché la diffusione sia comunque virale? O meglio, quanto alta può essere tenuta la soglia di adozione perché si generi una cascata completa?

9 Capacità di cascata

Quanto alta può essere tenuta la soglia di adozione q perché si generi una cascata completa? (naturalmente, $q = \frac{b}{a+b} \leq 1$)

Innanzi tutto, dipende dalla struttura della rete: alcune strutture ostacolano maggiormente di altre la generazione di cascate complete.

PROBLEMA: dato un grafo $G = (V, E)$, qual è la soglia di adozione massima q_{MAX} in G affinché un "piccolo" insieme V_0 di iniziatori di un nuovo stato A generi una cascata completa?

Il valore q_{MAX} prende il nome di **capacità di cascata** di G. Ma, nel descriverlo siamo stati un po' imprecisi: cosa significa che V_0 deve essere "piccolo"? Quanto "piccolo"?!

Osservazione: se scegliamo $V_0 = V$, allora, la soglia $q = 1$ genera una cascata completa: tutti i nodi sono forzati nello stato A! Ma ciò non ha molto senso... Richiedendo che V_0 sia "piccolo", intendiamo che vogliamo che $|V_0| \ll |V|$.

9.1 Ipotesi di lavoro (Grafi infiniti)

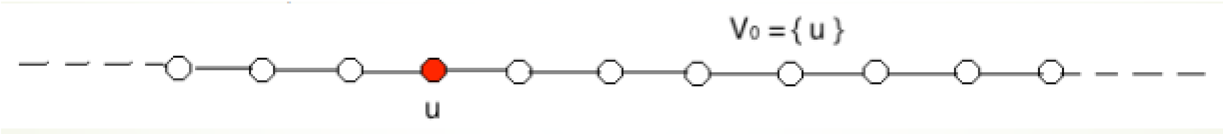
- G è un grafo infinito, ossia, contiene un numero infinito di nodi
- Vedremo degli esempi in cui G è un grafo regolare, ossia, la struttura del vicinato di un nodo è uguale a quella di tutti gli altri
- ESEMPIO: G è una catena di infiniti nodi, oppure una griglia di infiniti nodi
- l'insieme V_0 degli iniziatori può essere un qualsiasi insieme finito

Cioè: dato un grafo (regolare) infinito, qual è la soglia di adozione massima q_{MAX} in G affinché esista un insieme finito V_0 di iniziatori di un nuovo stato A che generi una cascata completa?

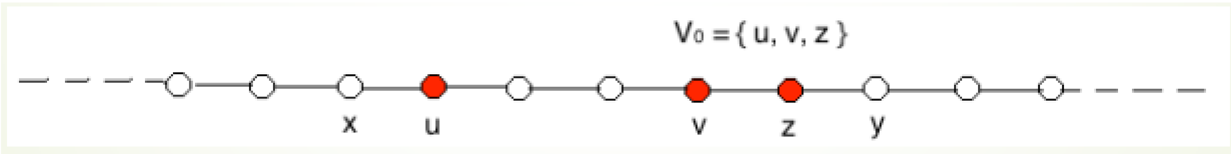
9.2 Esempi

ESEMPIO: G è una catena infinita

- se $|V_0| = 1$ ossia, V_0 contiene un solo nodo, allora occorre $q = \frac{1}{2}$ per generare una cascata completa



- ma, scegliendo un insieme più grande, è possibile generare una cascata con $q > \frac{1}{2}$?
- No: perché i nodi "al confine" con V_0, (i nodi x e y in figura), hanno comunque bisogno di $q = \frac{1}{2}$ per passare ad A

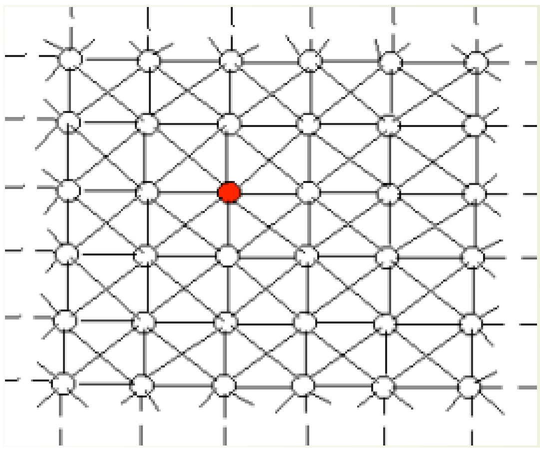


Allora, in una catena infinita, $q_{MAX} = \frac{1}{2}$.

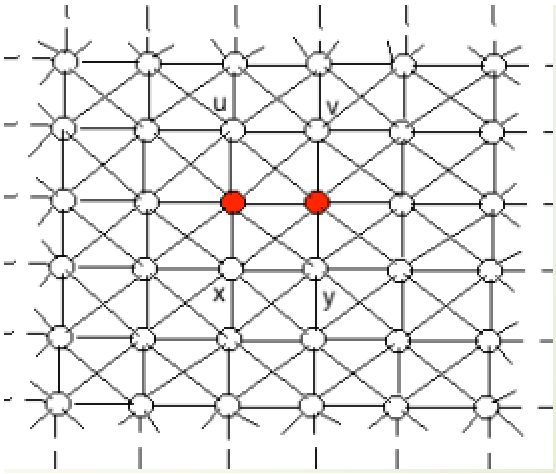
ESEMPIO: G è una griglia infinita

(Si assume una griglia con 8 vicini per nodo: 4 ortogonali e 4 diagonali)

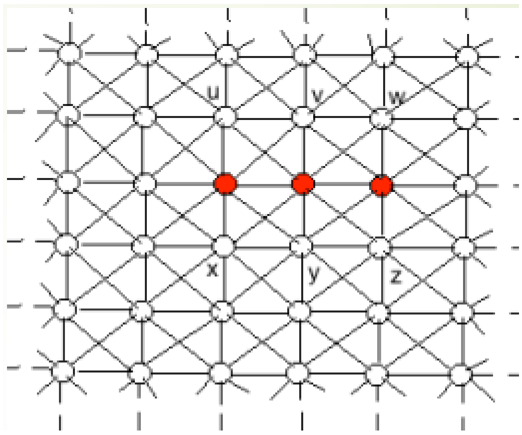
- se $|V_0| = 1$, ossia, V_0 contiene un solo nodo, allora occorre $q = \frac{1}{8}$ per generare una cascata completa



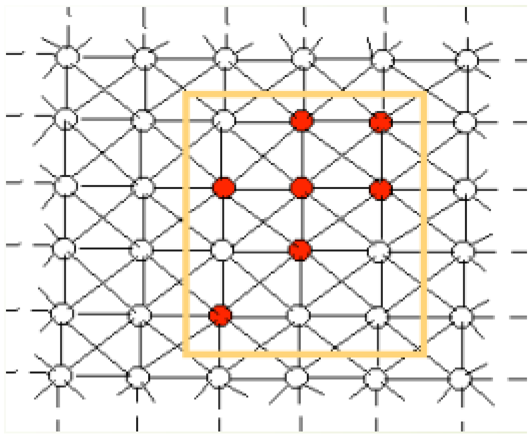
- se $|V_0| = 2$, ossia, V_0 contiene due nodi, allora, scegliendo i due nodi in V_0 come nodi adiacenti, con $q = \frac{2}{8} = \frac{1}{4}$ si riesce a influenzare u, v, x, y e poi a generare una cascata completa



- se $|V_0| = 3$, allora, scegliendo i tre nodi in V_0 come nodi adiacenti, con $q = \frac{3}{8}$ si riesce a influenzare v e y, poi u, w, x, z, e così via fino a generare una cascata completa



- aumentando $|V_0|$ non si riesce ad aumentare la soglia di adozione: una volta influenzati tutti i nodi nel rettangolo (giallo) che contiene gli iniziatori, occorre uscire da esso



- e per farlo è necessario $q = \frac{3}{8}$

Allora, in una griglia infinita (a 8 vicini), $q_{MAX} = \frac{3}{8}$.

9.3 Conclusioni (Capacità di cascata)

Dai due esempi della catena e della griglia possiamo concludere quanto segue:

- la soglia di adozione massima è più bassa nella griglia (che ha una topologia più ricca) che non nella catena (che ha una topologia più povera)

- in entrambi i casi, essa non supera $\frac{1}{2}$
- cioè, affinché si generi una cascata completa, A deve essere almeno tanto appetibile di B

Poiché la soglia di adozione massima è una caratteristica della rete, ovvero, della sua topologia, la domanda sorge, a questo punto, spontanea: esistono topologie nelle quali la soglia di adozione massima è maggiore di $\frac{1}{2}$?

Ossia, esistono topologie nelle quali innovazioni di qualità mediocre soppiantino uno status quo di qualità migliore?

Fortunatamente, no...