Going Viral with the Subscription Business Model

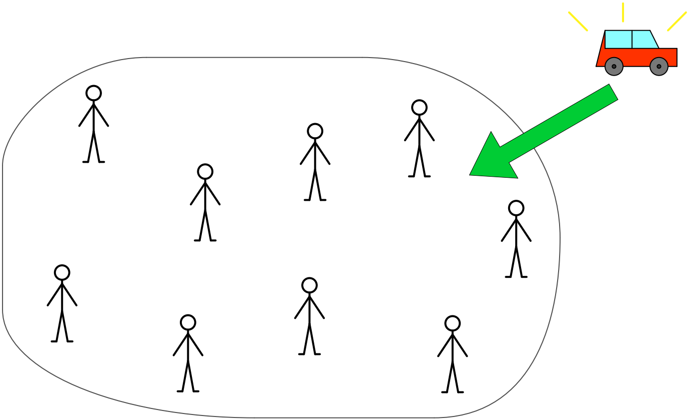
Influence Spreading in Social Networks

Giovanni Viglietta

(in collaboration with Joseph Peters)

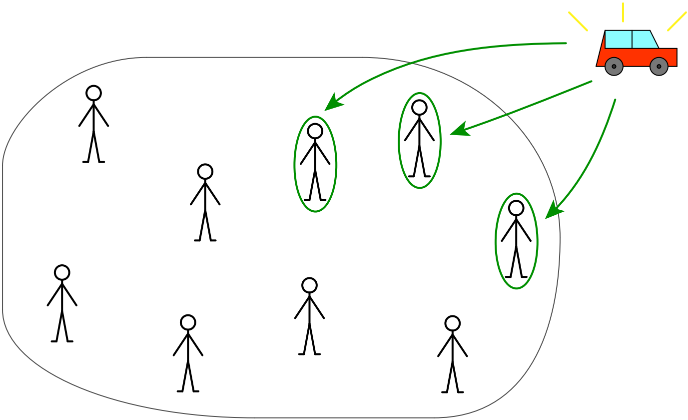
JAIST – September 11, 2019

Marketing in a social network



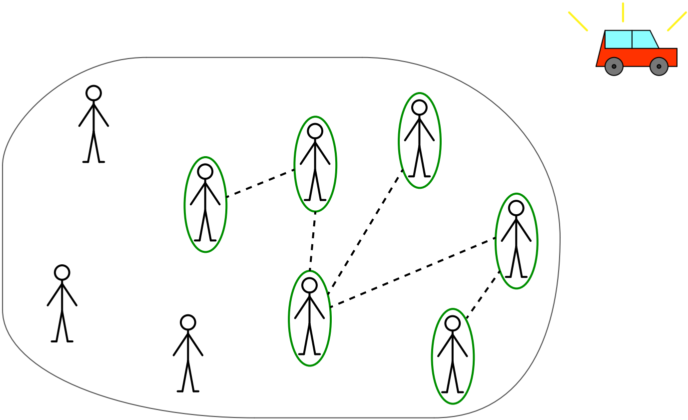
A company wants to introduce a new product to a group of people.

Marketing in a social network



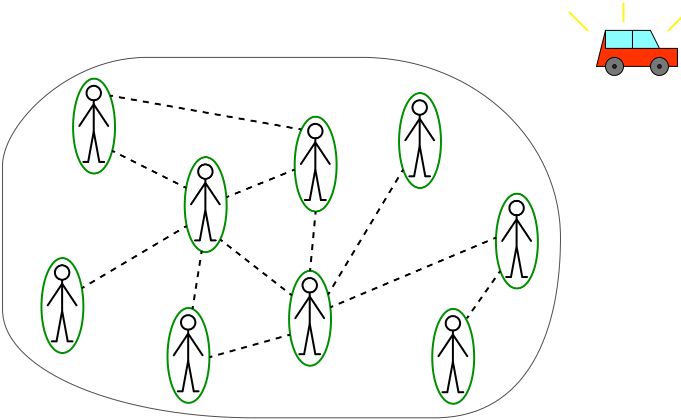
Instead of marketing it to everyone, it chooses a smaller target set.

Marketing in a social network



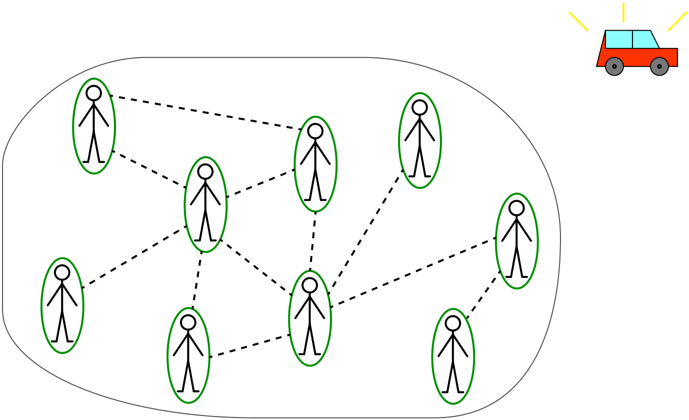
The initial adopters will spontaneously advertize it to their friends.

Marketing in a social network



This starts a cascade of adoptions spreading through the network.

Marketing in a social network

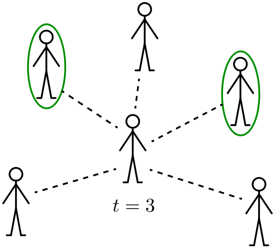


Customers have an individual value, but also a network value!

Basic influence-spreading model

D. Kempe, J. Kleinberg, and É. Tardos
Maximizing the spread of influence through a social network
KDD 2003

Individuals may be more or less reluctant to emulate their peers.
This is measured by a parameter called threshold.

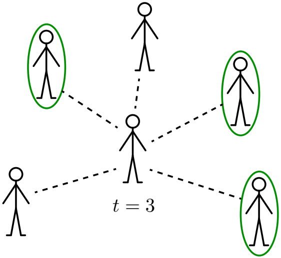


If an individual with threshold t has at least t friends who adopted the product, he will adopt the product as well.

Basic influence-spreading model

D. Kempe, J. Kleinberg, and É. Tardos
Maximizing the spread of influence through a social network
KDD 2003

Individuals may be more or less reluctant to emulate their peers.
This is measured by a parameter called threshold.

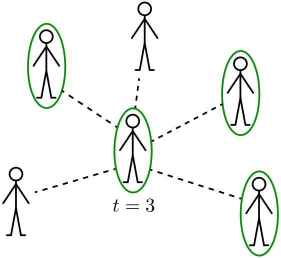


If an individual with threshold t has at least t friends who adopted the product, he will adopt the product as well.

Basic influence-spreading model

D. Kempe, J. Kleinberg, and É. Tardos
Maximizing the spread of influence through a social network
KDD 2003

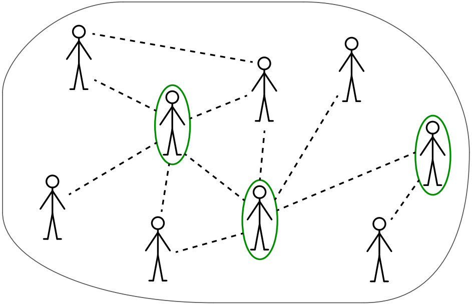
Individuals may be more or less reluctant to emulate their peers.
This is measured by a parameter called threshold.



If an individual with threshold t has at least t friends who adopted the product, he will adopt the product as well.

The subscription business model

What if the product is *subscription-based*?

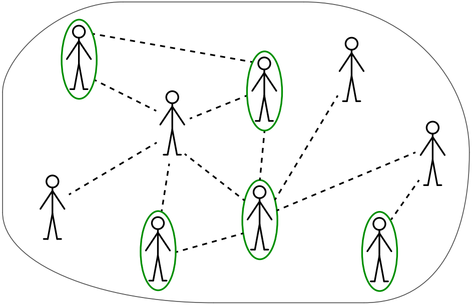


Every month, each individual decides whether to subscribe or not, based on his friends' choices and his own threshold.

Only the current subscribers exert influence on their friends.

The subscription business model

What if the product is *subscription-based*?

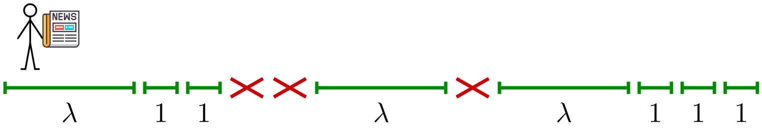


Every month, each individual decides whether to subscribe or not, based on his friends' choices and his own threshold.

Only the current subscribers exert influence on their friends.

The subscription business model

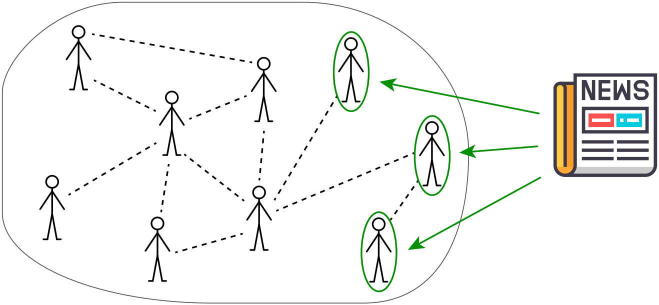
There may be a "promotional offer" for new subscribers:



If you become a subscriber, your subscription will remain active for λ months; then, you will decide whether to subscribe again or not.

Influencing the whole network

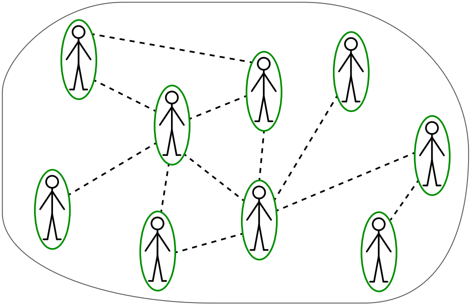
The company wants everyone to become a stable subscriber, only marketing the product to a small initial set of individuals.



Problem: How hard is it to determine such an initial target set?

Influencing the whole network

The company wants everyone to become a stable subscriber, only marketing the product to a small initial set of individuals.



Problem: How hard is it to determine such an initial target set?

Hardness of approximation

We have a reduction from the NP-complete problem Set Cover.

$$U = \{1, 2, 3, 4, 5, 6, 7\}$$

$$S_1 = \{1, 2, 4\}$$
$$S_2 = \{2, 3, 5\}$$
$$S_3 = \{4, 6\}$$
$$S_4 = \{3, 4, 6, 7\}$$
$$S_5 = \{1, 5, 7\}$$

Choose the smallest number of sets whose union is the universe.

Hardness of approximation

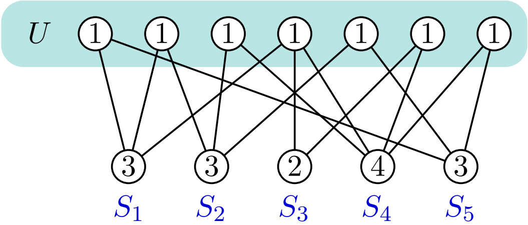
We have a reduction from the NP-complete problem Set Cover.

$$U = \{1, 2, 3, 4, 5, 6, 7\}$$

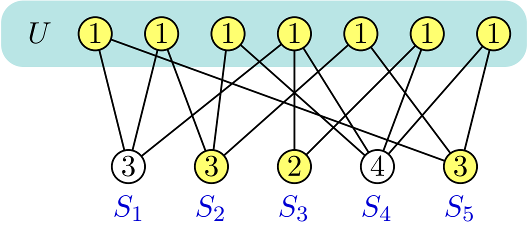
$$S_1 = \{1, 2, 4\}$$
$$S_2 = \{2, 3, 5\} \longleftarrow$$
$$S_3 = \{4, 6\} \longleftarrow$$
$$S_4 = \{3, 4, 6, 7\}$$
$$S_5 = \{1, 5, 7\} \longleftarrow$$

Choose the smallest number of sets whose union is the universe.

Reduction:

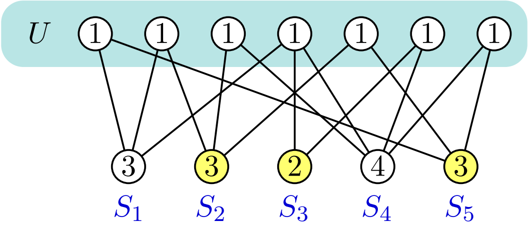


Reduction:



If $\lambda = 1$, choose the universe and some other elements.

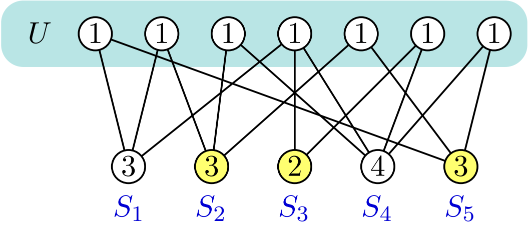
Reduction:



If $\lambda > 1$, ignore the universe elements.

Hardness of approximation

Reduction:



If $\lambda > 1$, ignore the universe elements.

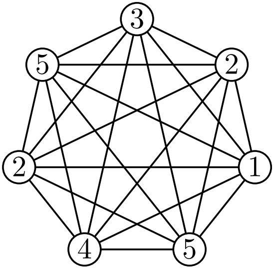
Note: Set Cover is hard to approximate within a factor of $o(\log n)$, and we gave an approximation-preserving reduction.

Theorem

For bipartite networks, the smallest target set is NP-hard to approximate within a factor of $o(\log n)$.

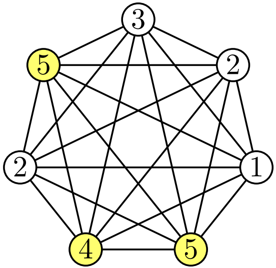
Complete graphs

What if the network is a complete graph?



Complete graphs

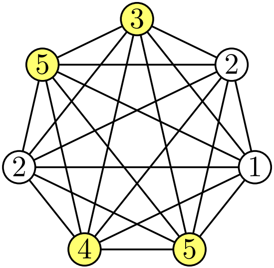
What if the network is a complete graph?



Note: it is convenient to pick the nodes of highest threshold.

Complete graphs

What if the network is a complete graph?

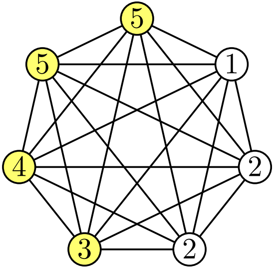


Note: it is convenient to pick the nodes of highest threshold.

Note: picking more nodes is better than picking fewer nodes.

Complete graphs

What if the network is a complete graph?



Note: it is convenient to pick the nodes of highest threshold.

Note: picking more nodes is better than picking fewer nodes.

So, sort the nodes in $O(n \log n)$ time and use binary search.

Complete graphs

We can find the smallest target set if we can predict whether a given initial set will influence the whole network.

Note: after λ rounds, the number of influenced nodes will be monotonic (increasing or decreasing).

So, the influenced set stabilizes in $O(n)$ rounds.

Since the steps of the binary search are $O(\log n)$, this yields $O(n \log n)$ time.

Theorem

If the network is a complete graph, the smallest target set is computable in $O(n \log n)$ time.

Paths and cycles

What if the network is a path or a cycle?

Paths and cycles

What if the network is a path or a cycle?

Note: if $\lambda = 1$, we must choose all nodes of threshold 2, as well as their neighbors.

Paths and cycles

What if the network is a path or a cycle?

Note: if $\lambda = 1$, we must choose all nodes of threshold 2, as well as their neighbors.

Note: if $\lambda > 1$, if two neighboring nodes of threshold 2 are both inactive, they will never become active.

Paths and cycles

What if the network is a path or a cycle?

Note: if $\lambda = 1$, we must choose all nodes of threshold 2, as well as their neighbors.

Note: if $\lambda > 1$, if two neighboring nodes of threshold 2 are both inactive, they will never become active.

Note: activating a single node in a chain of nodes of threshold 1 will activate the whole chain.

Paths and cycles

What if the network is a path or a cycle?

Note: if $\lambda = 1$, we must choose all nodes of threshold 2, as well as their neighbors.

Note: if $\lambda > 1$, if two neighboring nodes of threshold 2 are both inactive, they will never become active.

Note: activating a single node in a chain of nodes of threshold 1 will activate the whole chain.

Paths and cycles

What if the network is a path or a cycle?

Note: if $\lambda = 1$, we must choose all nodes of threshold 2, as well as their neighbors.

Note: if $\lambda > 1$, if two neighboring nodes of threshold 2 are both inactive, they will never become active.

Note: activating a single node in a chain of nodes of threshold 1 will activate the whole chain.

Paths and cycles

What if the network is a path or a cycle?

Note: if $\lambda = 1$, we must choose all nodes of threshold 2, as well as their neighbors.

Note: if $\lambda > 1$, if two neighboring nodes of threshold 2 are both inactive, they will never become active.

Note: activating a single node in a chain of nodes of threshold 1 will activate the whole chain.

These observations lead to a linear-time algorithm.

Theorem

If the network is a path or a cycle, the smallest target set is computable in $O(n)$ time.

Activation periods

Each node can only have one of $\lambda + 1$ possible states:

- it is inactive,
- it has been active for 1 round,
- it has been active for 2 rounds,
- ...
- it has been active for λ rounds or more.

So, the whole network has $(\lambda + 1)^n$ possible states.

After $(\lambda + 1)^n$ rounds, the activation pattern will become periodic.

Problem: What are the possible periods?

Activation periods

Note: If the period is > 1, at least one particle must become active and then inactive: this takes $\lambda + 1$ rounds.

So, all the periods between 2 and λ are impossible.

Period 1 is possible (e.g., empty target set).

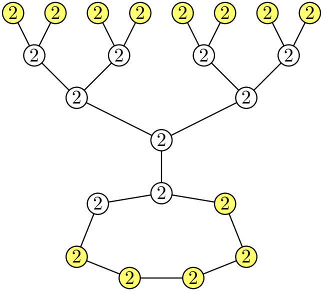
Activation periods

Note: If the period is > 1, at least one particle must become active and then inactive: this takes $\lambda + 1$ rounds.

So, all the periods between 2 and λ are impossible.

Period 1 is possible (e.g., empty target set).

If $\lambda > 1$, all periods $> \lambda$ are possible.



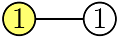
Activation periods

Note: If the period is > 1, at least one particle must become active and then inactive: this takes $\lambda + 1$ rounds.

So, all the periods between 2 and λ are impossible.

Period 1 is possible (e.g., empty target set).

If $\lambda > 1$, all periods $> \lambda$ are possible.



If $\lambda = 1$, period 2 is possible.

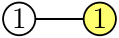
Activation periods

Note: If the period is > 1, at least one particle must become active and then inactive: this takes $\lambda + 1$ rounds.

So, all the periods between 2 and λ are impossible.

Period 1 is possible (e.g., empty target set).

If $\lambda > 1$, all periods $> \lambda$ are possible.



If $\lambda = 1$, period 2 is possible.

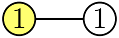
Activation periods

Note: If the period is > 1, at least one particle must become active and then inactive: this takes $\lambda + 1$ rounds.

So, all the periods between 2 and λ are impossible.

Period 1 is possible (e.g., empty target set).

If $\lambda > 1$, all periods $> \lambda$ are possible.



If $\lambda = 1$, period 2 is possible.

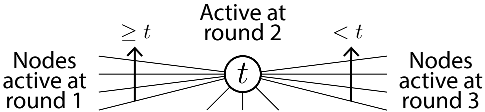
No large periods for $\lambda = 1$

If $\lambda = 1$, all periods > 2 are impossible.

Proof for a special case:

Suppose that each node is only active for one round per period.

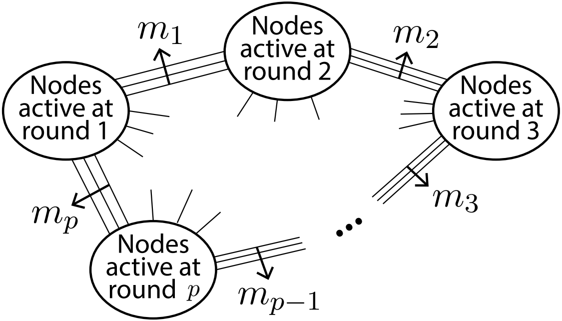
If a node is active only at round i, it must receive more influence at round $i - 1$ than in any other round in the period.



So, it receives more influence at round $i - 1$ than at round $i + 1$.

No large periods for $\lambda = 1$

Since this is true of all nodes, it leads to a cycle of inequalities.



$m_1 > m_2 > m_3 > \cdots > m_{p-1} > m_p > m_1$: impossible!

Activation periods for general graphs

Theorem

If $\lambda = 1$,

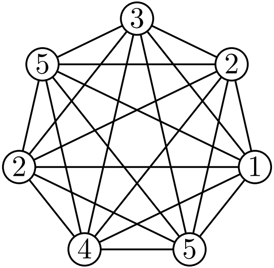
- *Periods* 1 *and* 2 *are possible*
- *All other periods are impossible*

If $\lambda > 1$,

- *Periods* $2, 3, \ldots, \lambda$ *are impossible*
- *All other periods are possible*

Activation periods for complete graphs

What if the network is a complete graph?



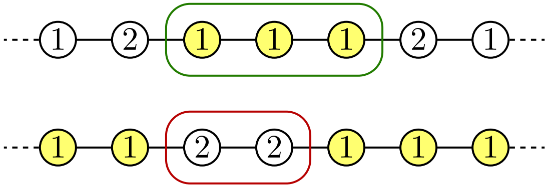
We know that the configuration of active nodes becomes stable after $O(n)$ rounds.

So, complete graphs have period 1.

Activation periods for paths and cycles

What if the network is a path or a cycle?

Some areas become permanently active or permanently inactive.



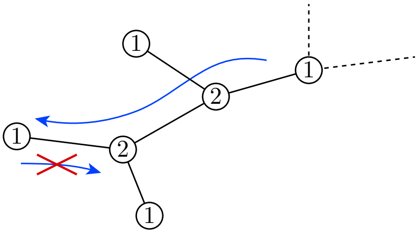
These areas act as "walls" for influence spreading.

We can apply induction on the left and right sub-paths and deduce that the only possible period is 1.

Activation periods for trees

What if the network is a tree?

A chain of deactivation events cannot "rebound" off leaves.



Hence all leaves will become stable; so we can prune the tree and apply induction.

We conclude that the only possible period is 1.

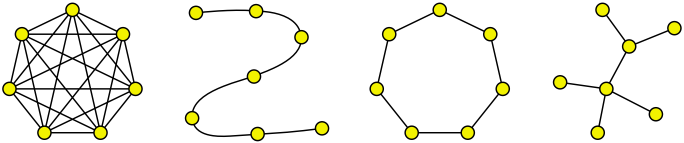
Activation periods for special classes of graphs

Theorem

If $\lambda > 1$ *and the network is*

- *a complete graph or*
- *a path or*
- *a cycle or*
- *a tree,*

then the activation period is 1*.*

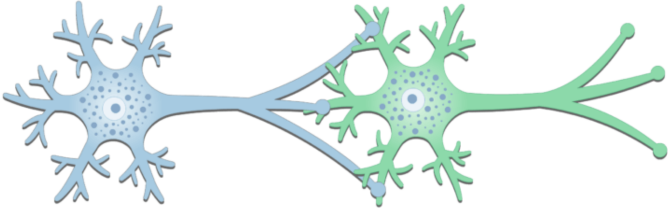


A change of perspective

Our influence-spreading model for social networks
works equally well for neural networks.

A change of perspective

A neuron becomes "excited" when the signals it receives from its neighbors exceeds a threshold.

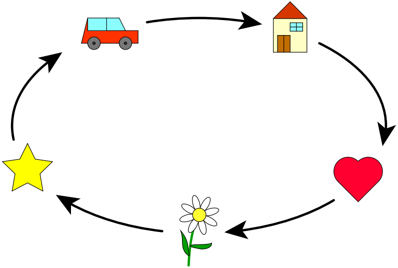


λ represents the "persistence of information" of neurons.

How can we model memory and intelligence?

A change of perspective

In the absence of external stimuli, intelligence is measured as the number of different "ideas" a brain can produce before repeating the same cycle of thoughts.



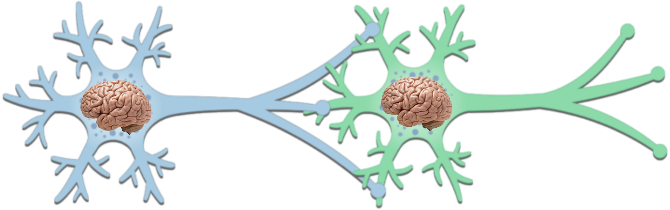
In the same context, memory is the number of different "concepts" a brain can repeat in a cycle.

Both are modeled by the activation period of the network!

A change of perspective

We know that high periods are possible only when $\lambda > 1$.

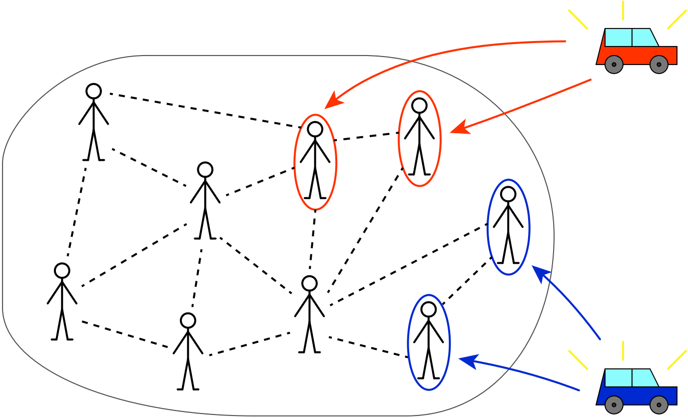
This means that memory is not an emergent behavior: in order to have a network with memory, we need neurons with some memory.



We also know that, if the network's topology is simple to describe (complete graph, path, cycle, tree), its behavior is not complex.

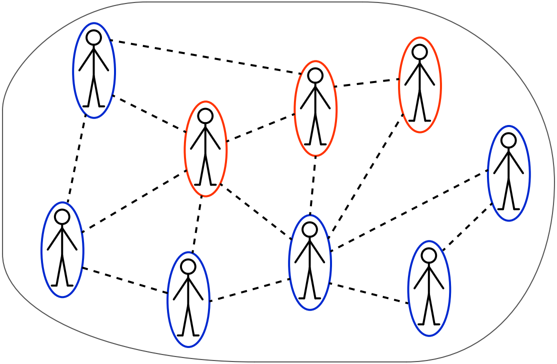
So, intelligence, memory and other complex behaviors require some inherent structural complexity of the network.

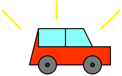
Open problem

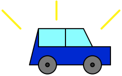


What if two companies release two competing products?
Whoever adopts a product cannot adopt the other.

Open problem

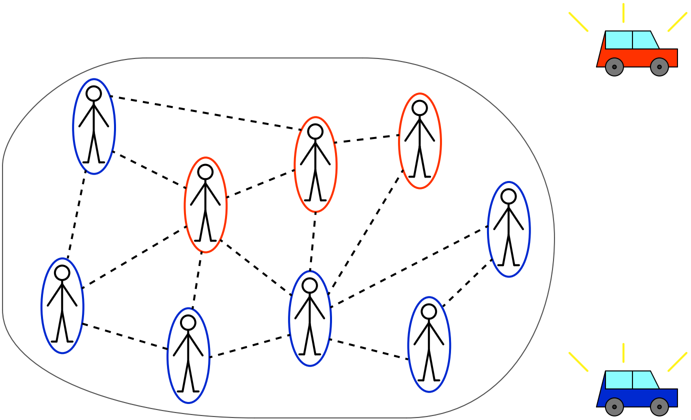






Both companies choose their target set with the goal of maximizing the number of adopters of their product.

Open problem



This game-like problem is solvable in PSPACE.
Conjecture: it is PSPACE-complete.