

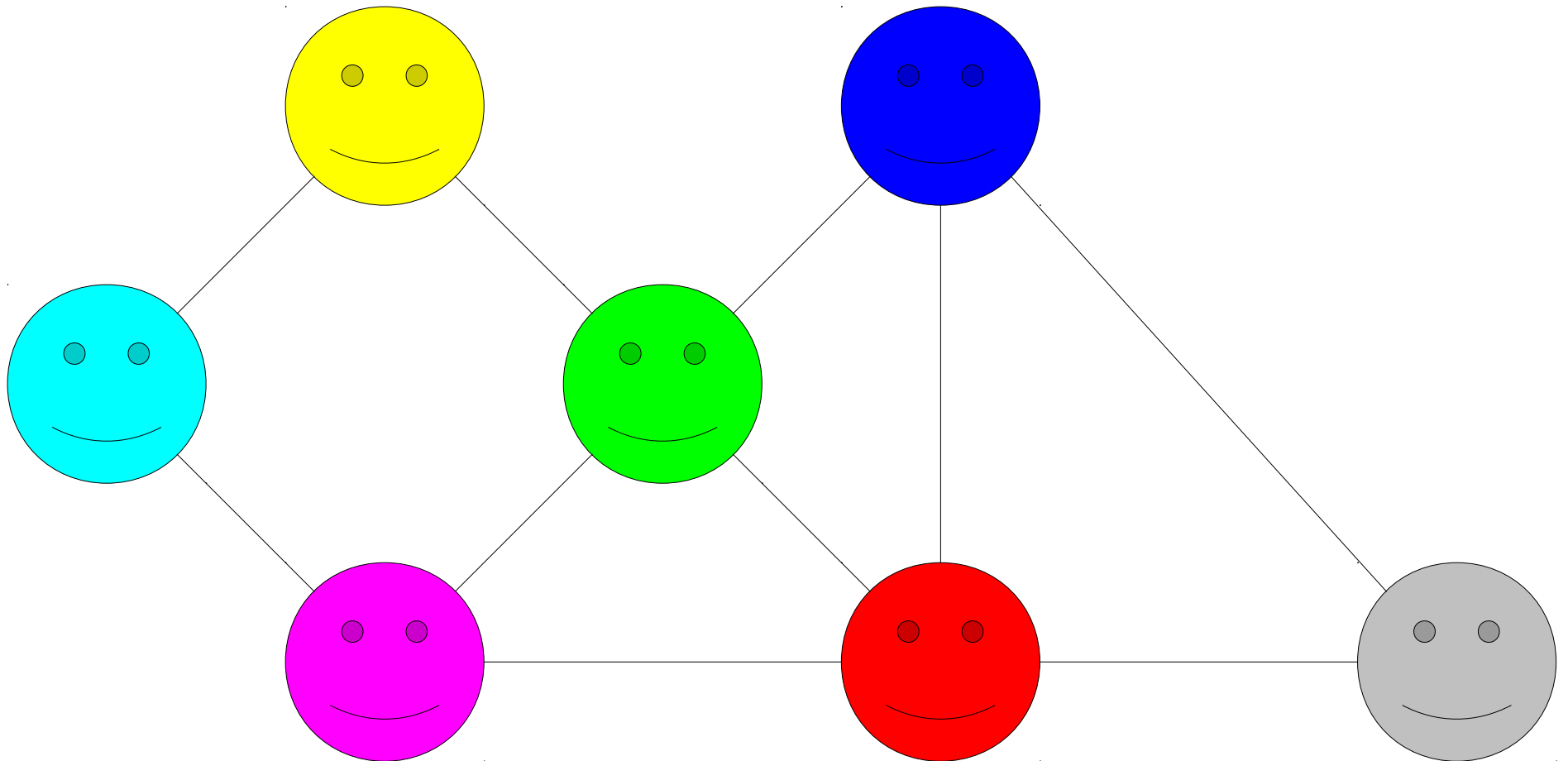
# Graphs and Networks

# Announcements

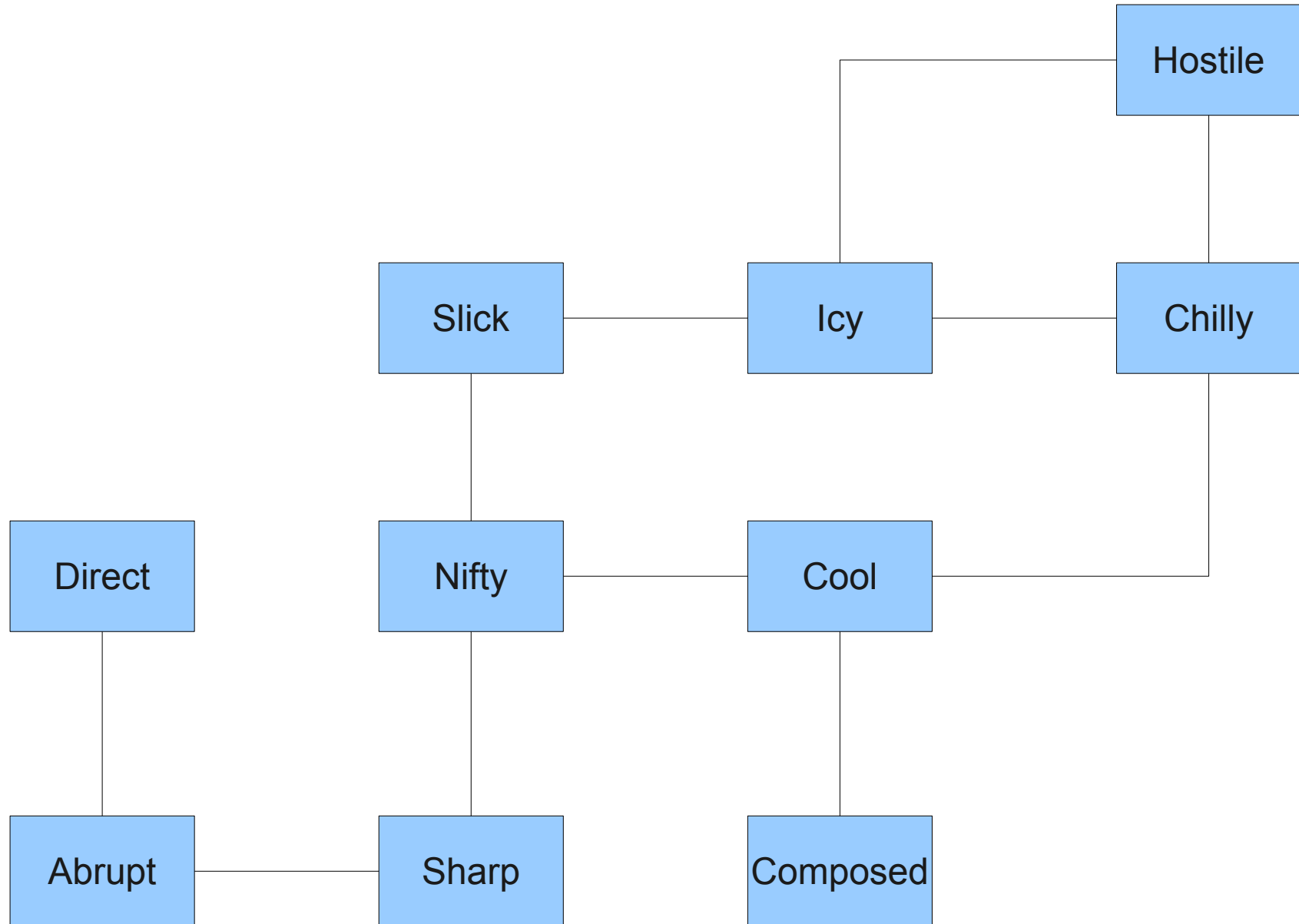
- Assignment 6 due right now.
- Assignment 7 (**FacePamphlet**) out, due next **Friday, March 16 at 3:15PM**
  - Build your own social network!
  - **No late days may be used.**
- YEAH hours tonight from 7-8PM in Herrin T175.
  - Get a jump on FacePamphlet!

# FacePamphlet Demo

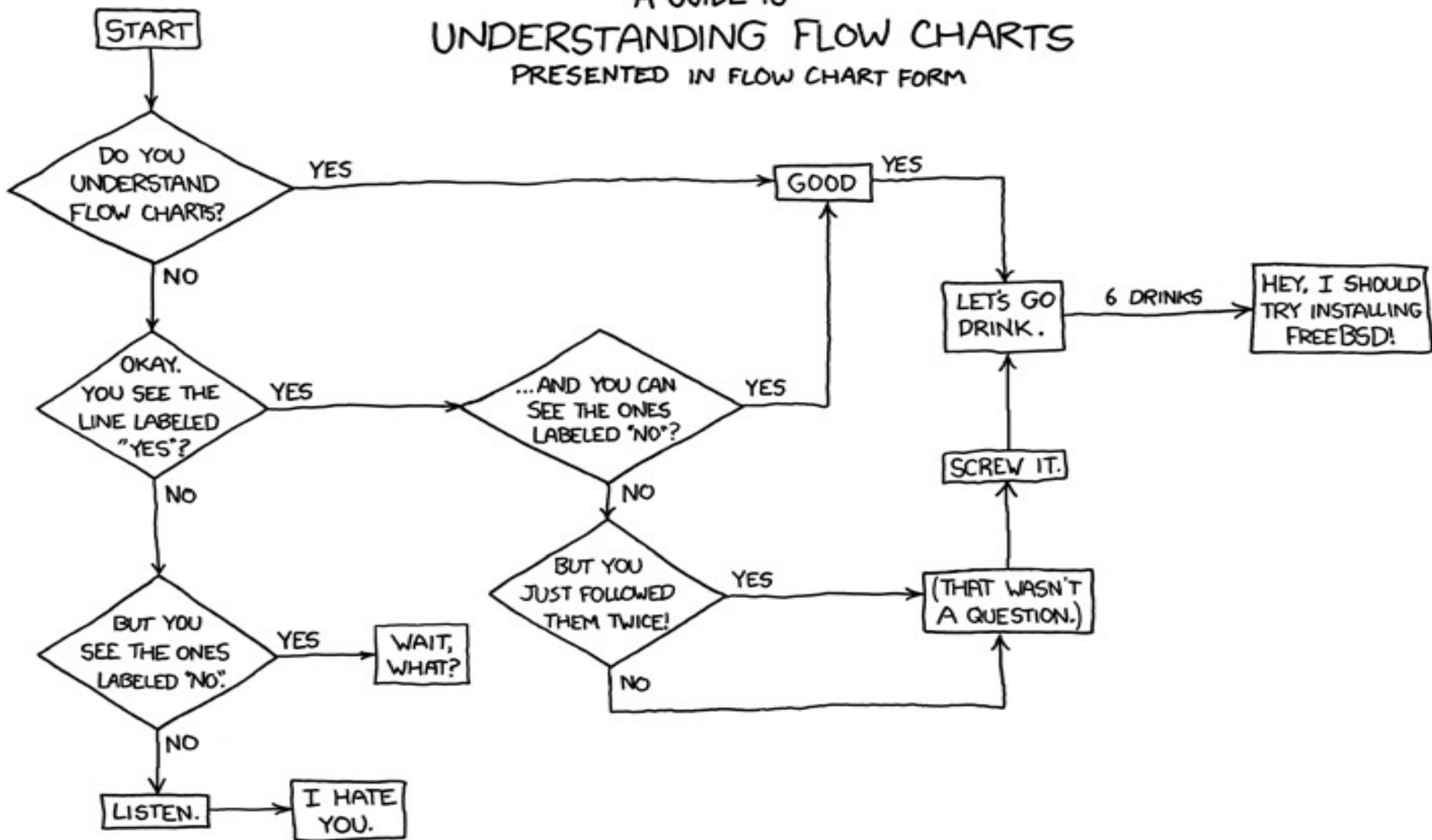
# A Social Network



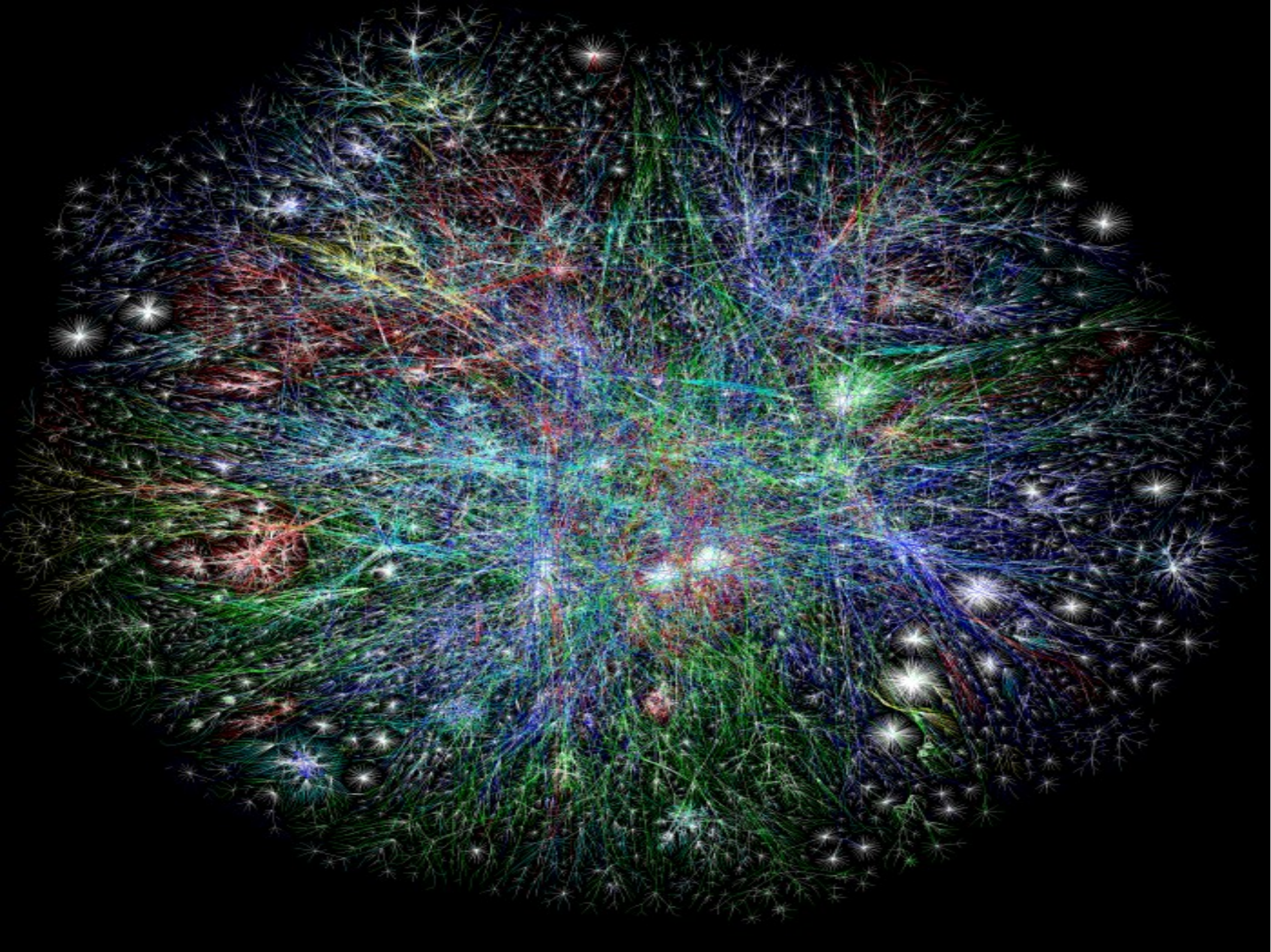
# Synonyms



# A GUIDE TO UNDERSTANDING FLOW CHARTS PRESENTED IN FLOW CHART FORM







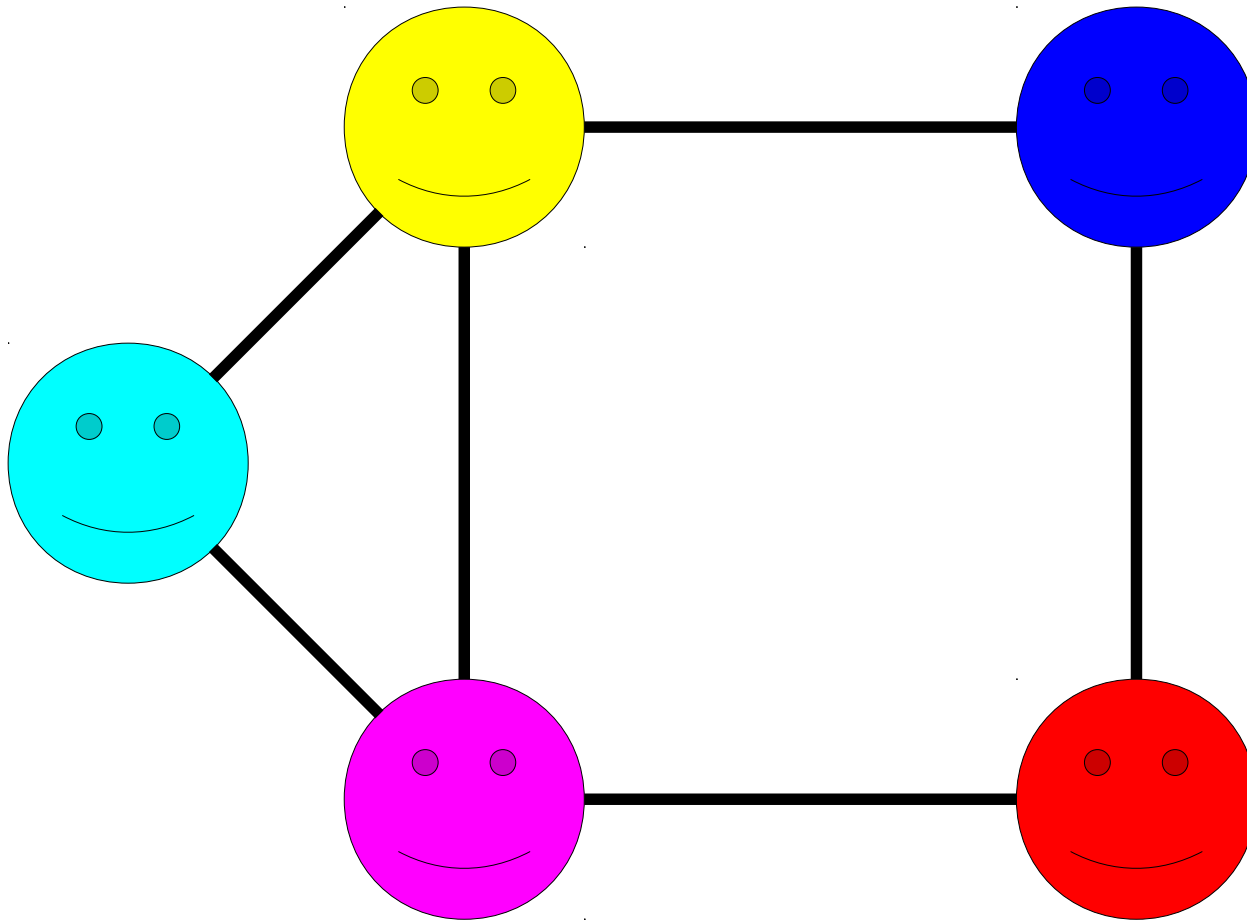


A **graph** is a mathematical structure for representing relationships.

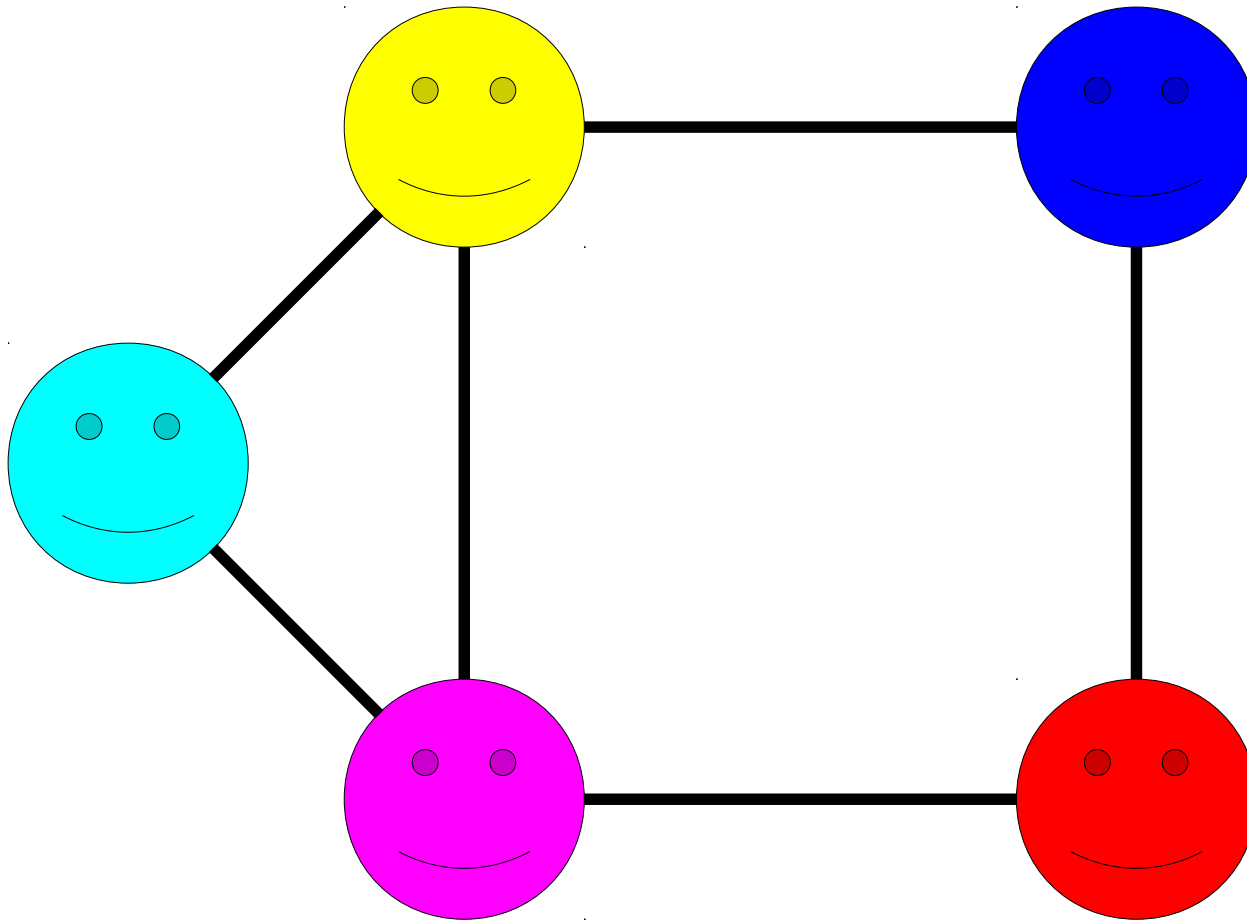


A **graph** is a mathematical structure for representing relationships.

A **graph** is a mathematical structure for representing relationships.

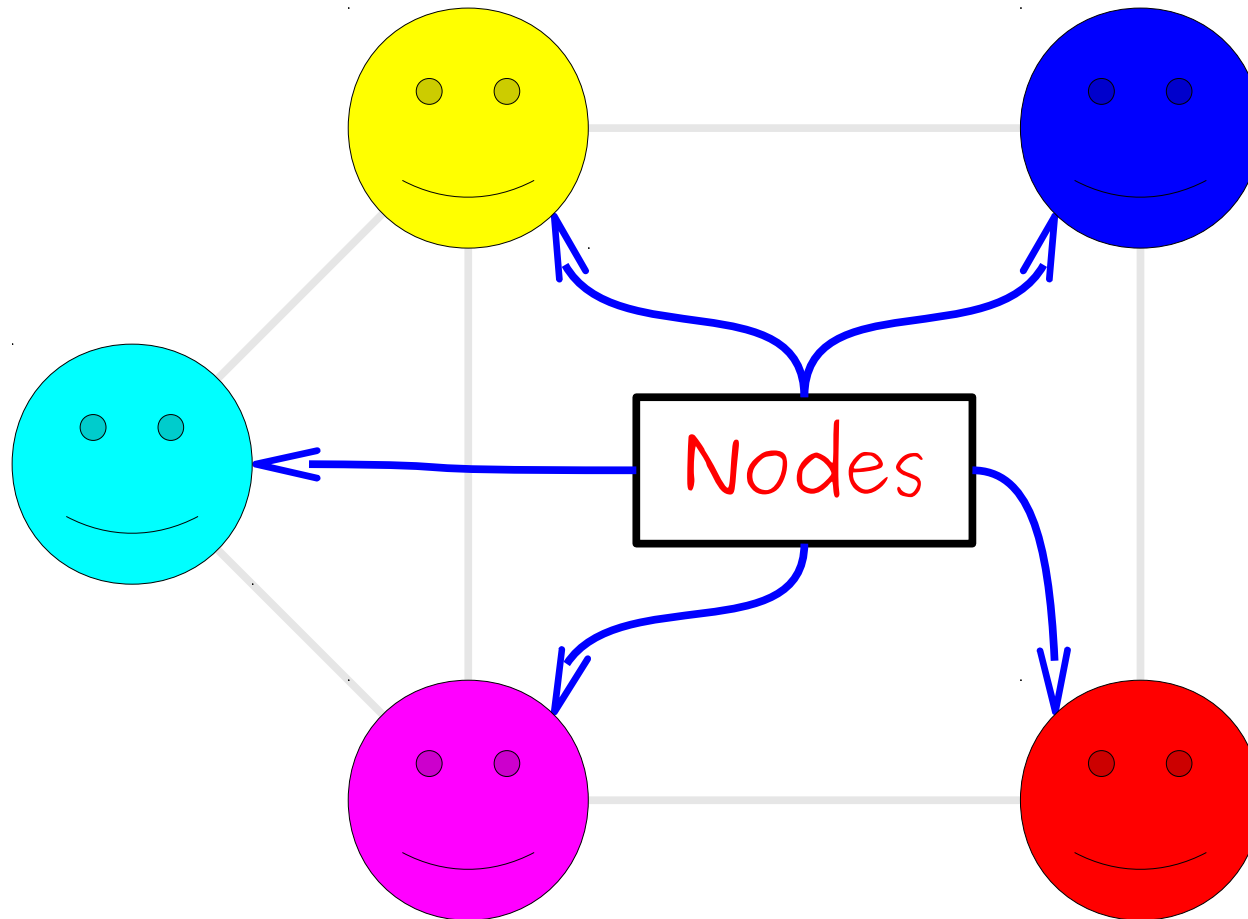


A **graph** is a mathematical structure for representing relationships.



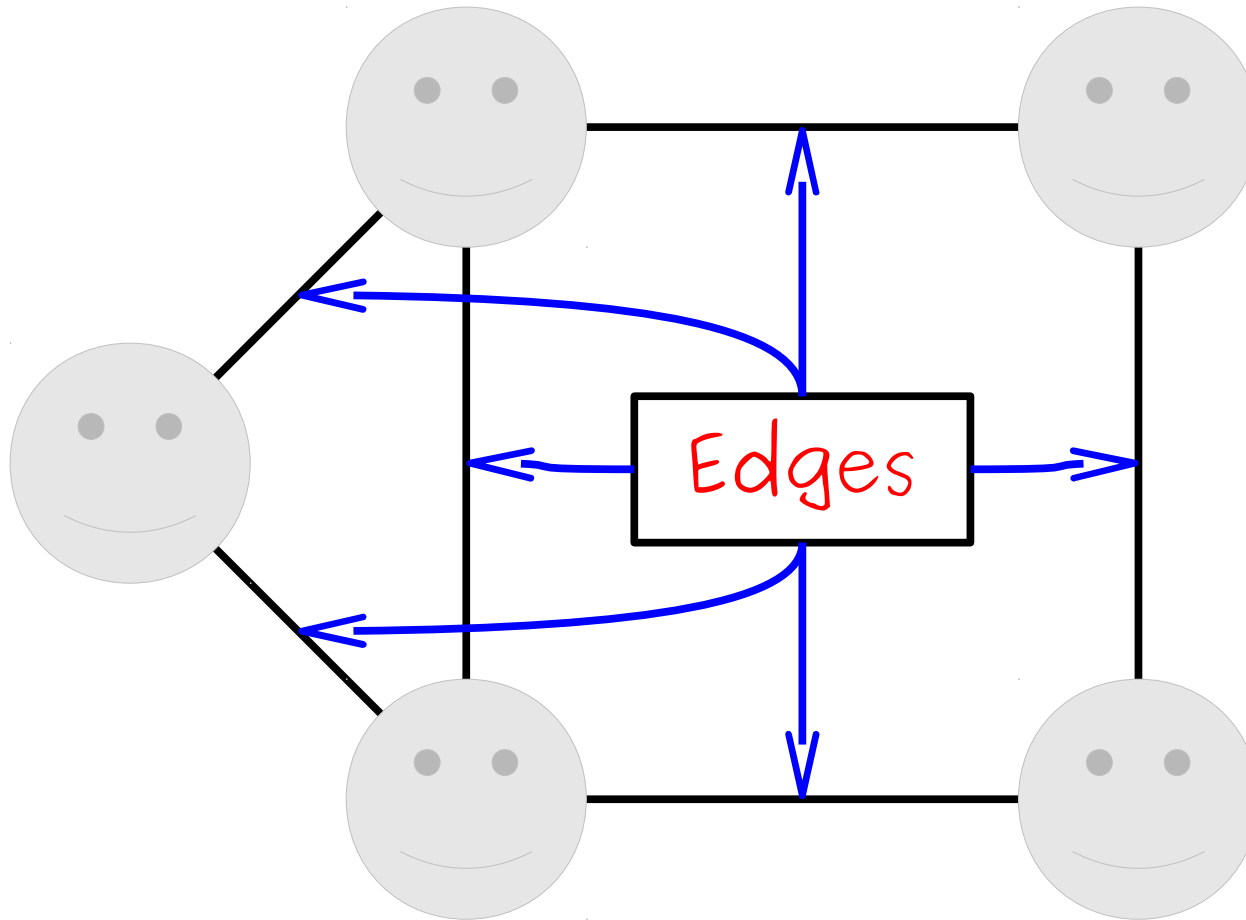
A graph consists of a set of **nodes** connected by **edges**.

A **graph** is a mathematical structure for representing relationships.



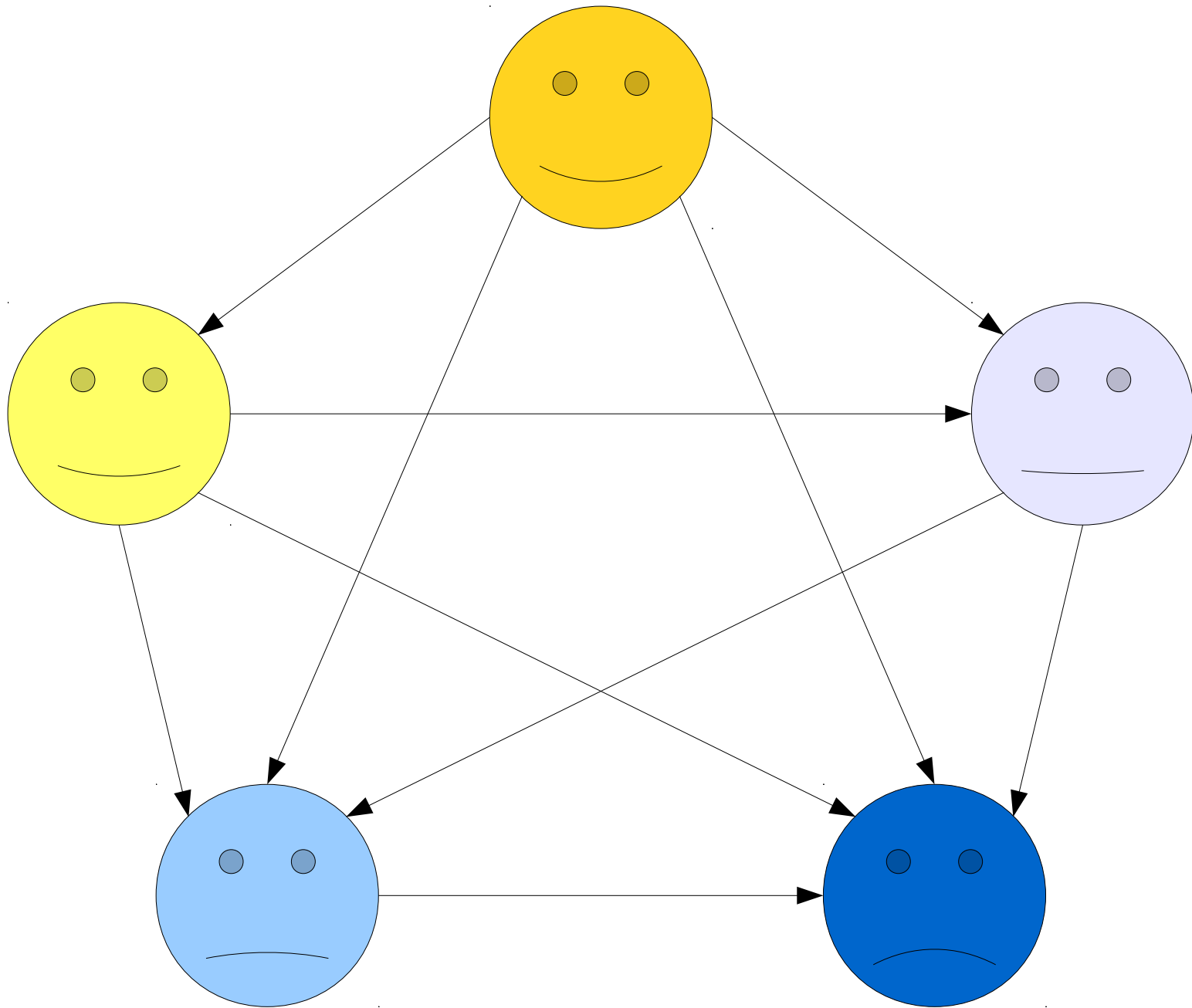
A graph consists of a set of **nodes** connected by **edges**.

A **graph** is a mathematical structure for representing relationships.



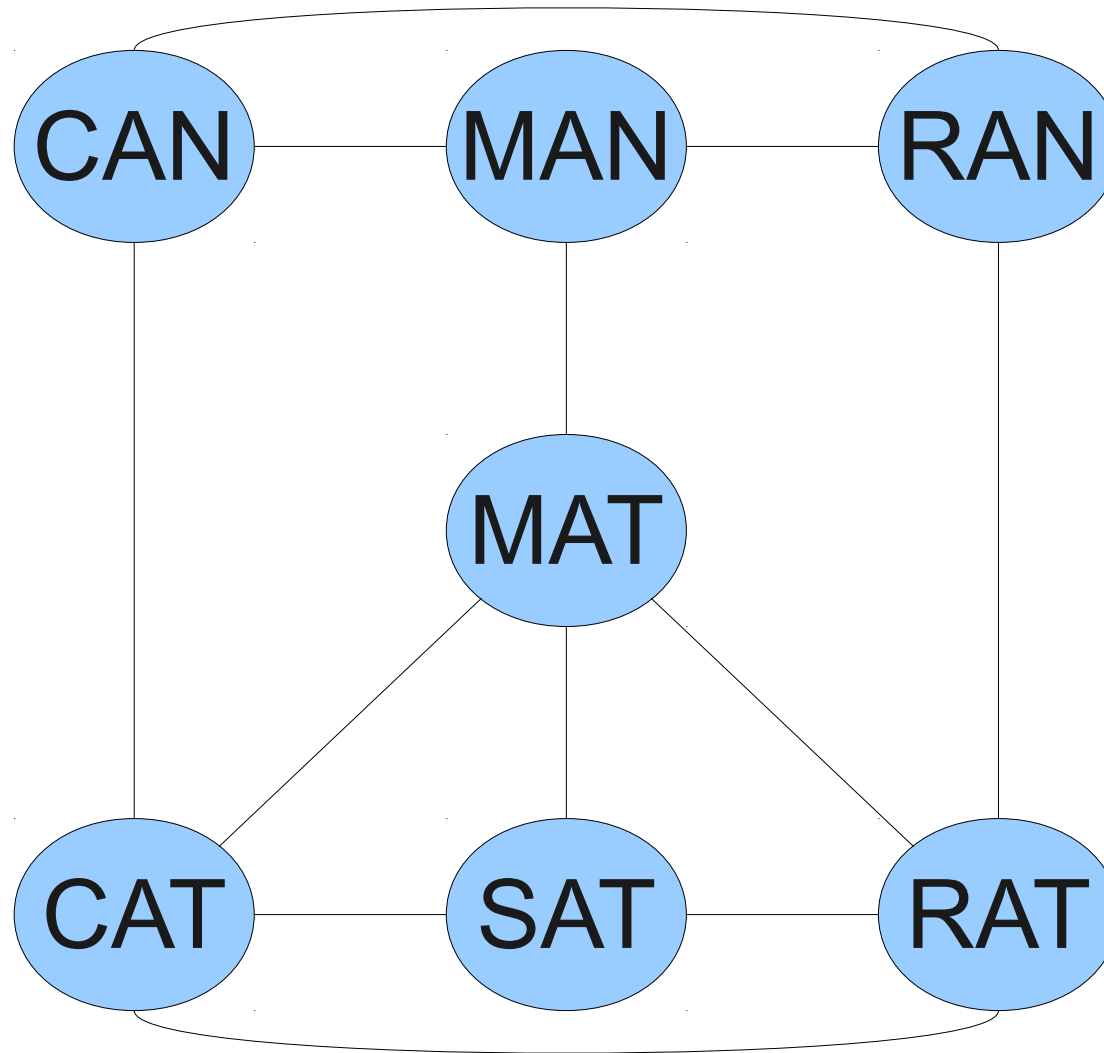
A graph consists of a set of **nodes** connected by **edges**.

Some graphs are **directed**.

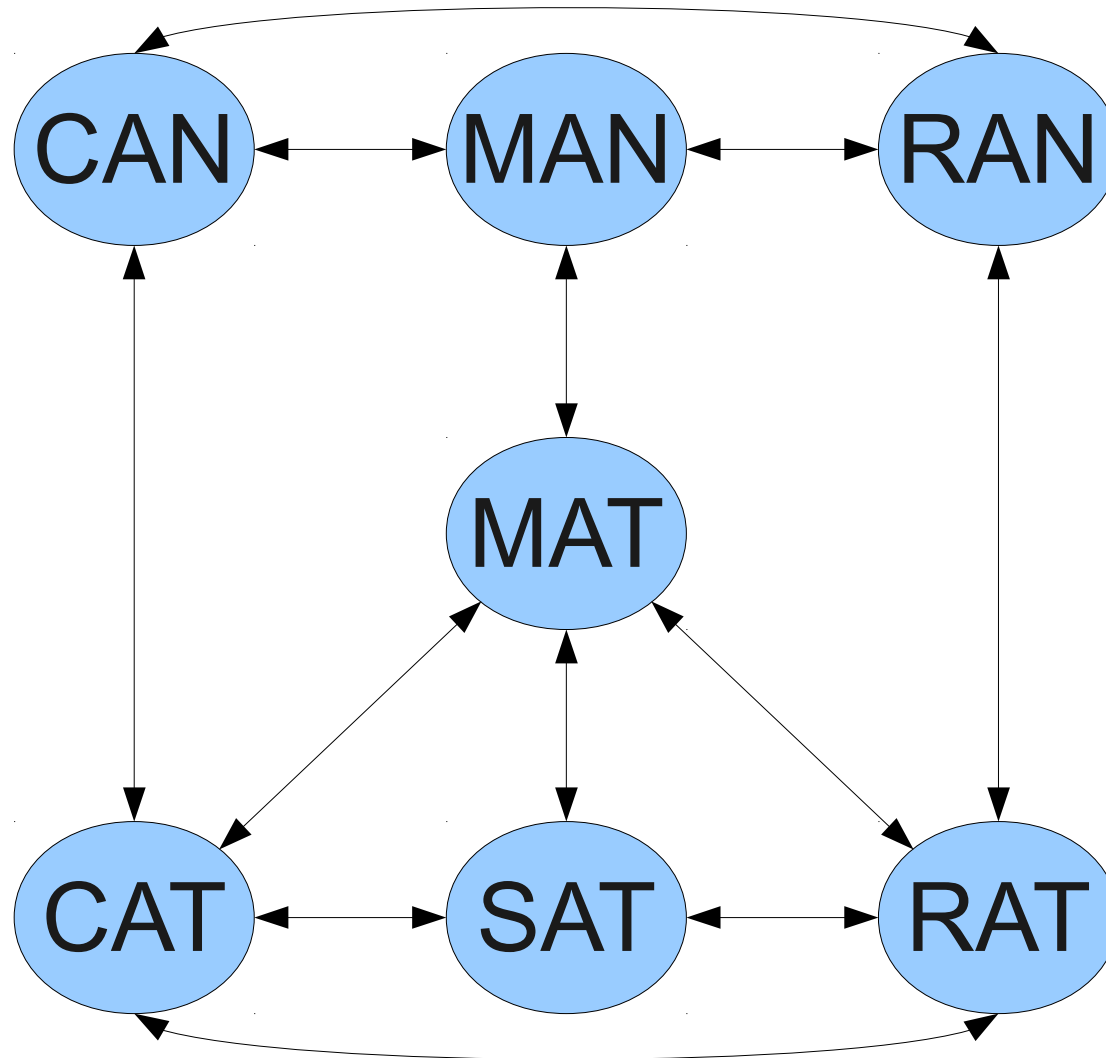




Some graphs are **undirected**.



Some graphs are **undirected**.

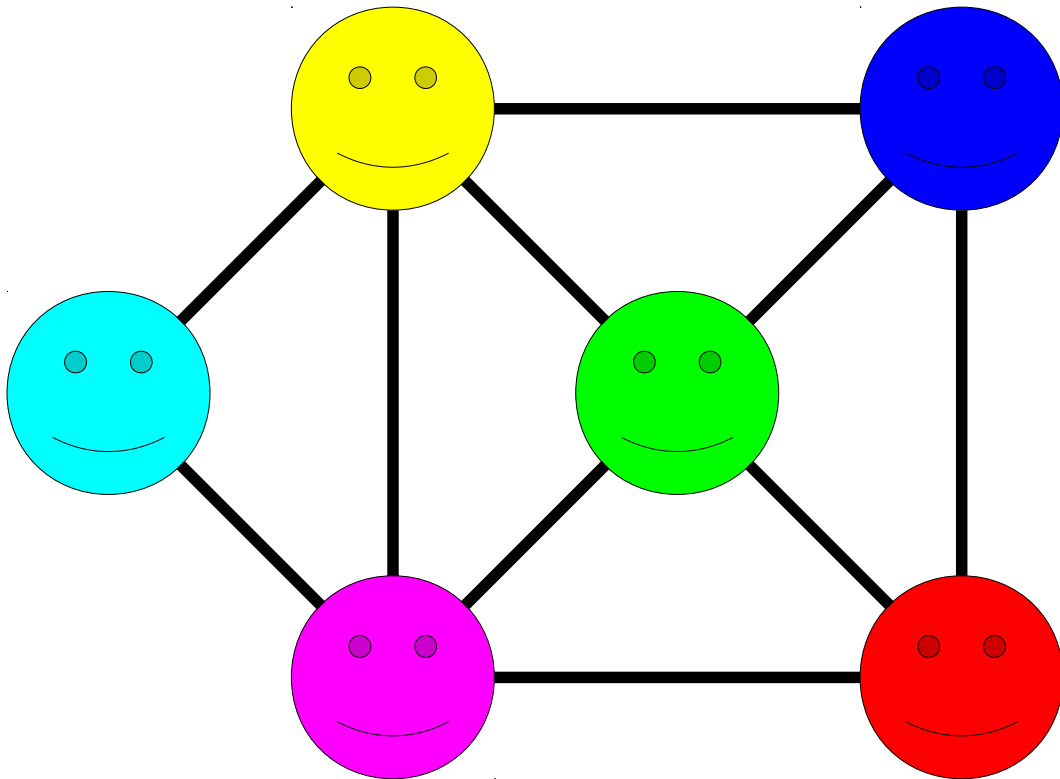


You can think of them as directed graphs with edges both ways.

How can we represent graphs in Java?

# Representing Graphs

We can represent a graph as a map from nodes to the list of nodes each node is connected to.



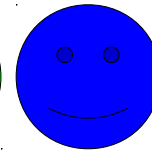
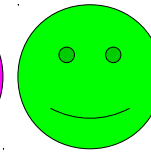
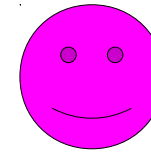
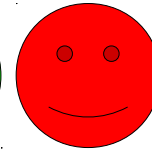
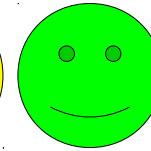
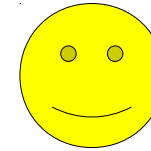
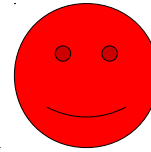
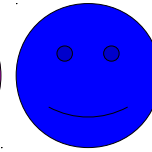
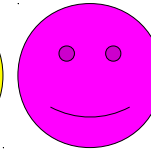
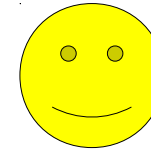
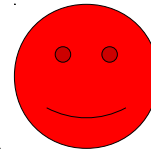
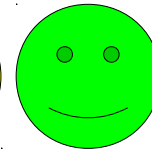
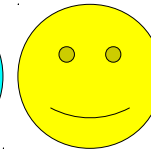
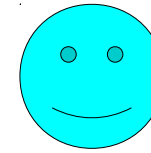
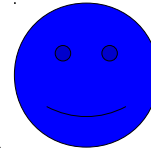
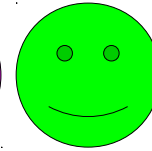
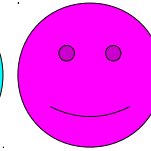
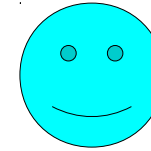
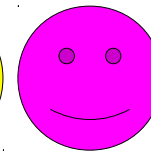
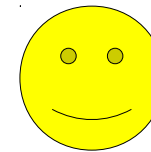
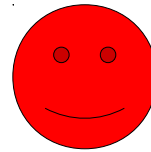
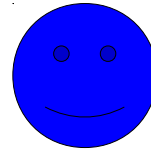
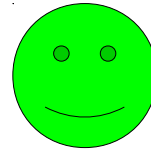
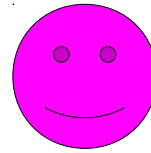
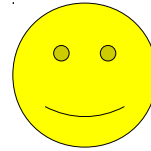
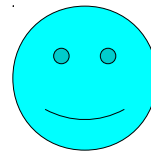
Map<**Node**, List<**Node**>>

**Node**

List<**Node**>

**Node**

**Connected To**



# The Wikipedia Graph



**WIKIPEDIA**  
*The Free Encyclopedia*

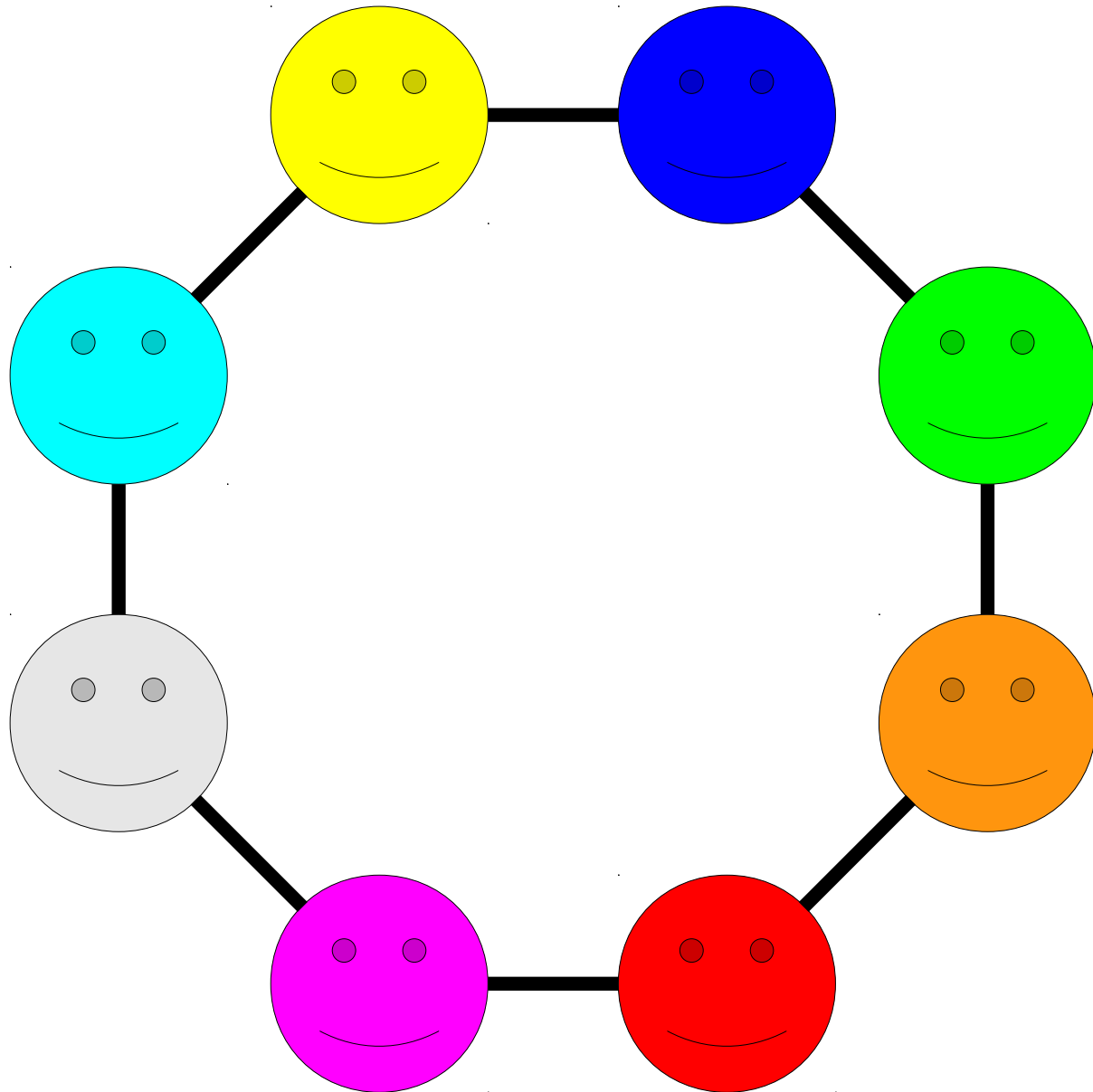
- Wikipedia (and the web in general) is a graph!
- Each page is a node.
- There is an edge from one page to another if the first page links to the second.

# Network Analysis

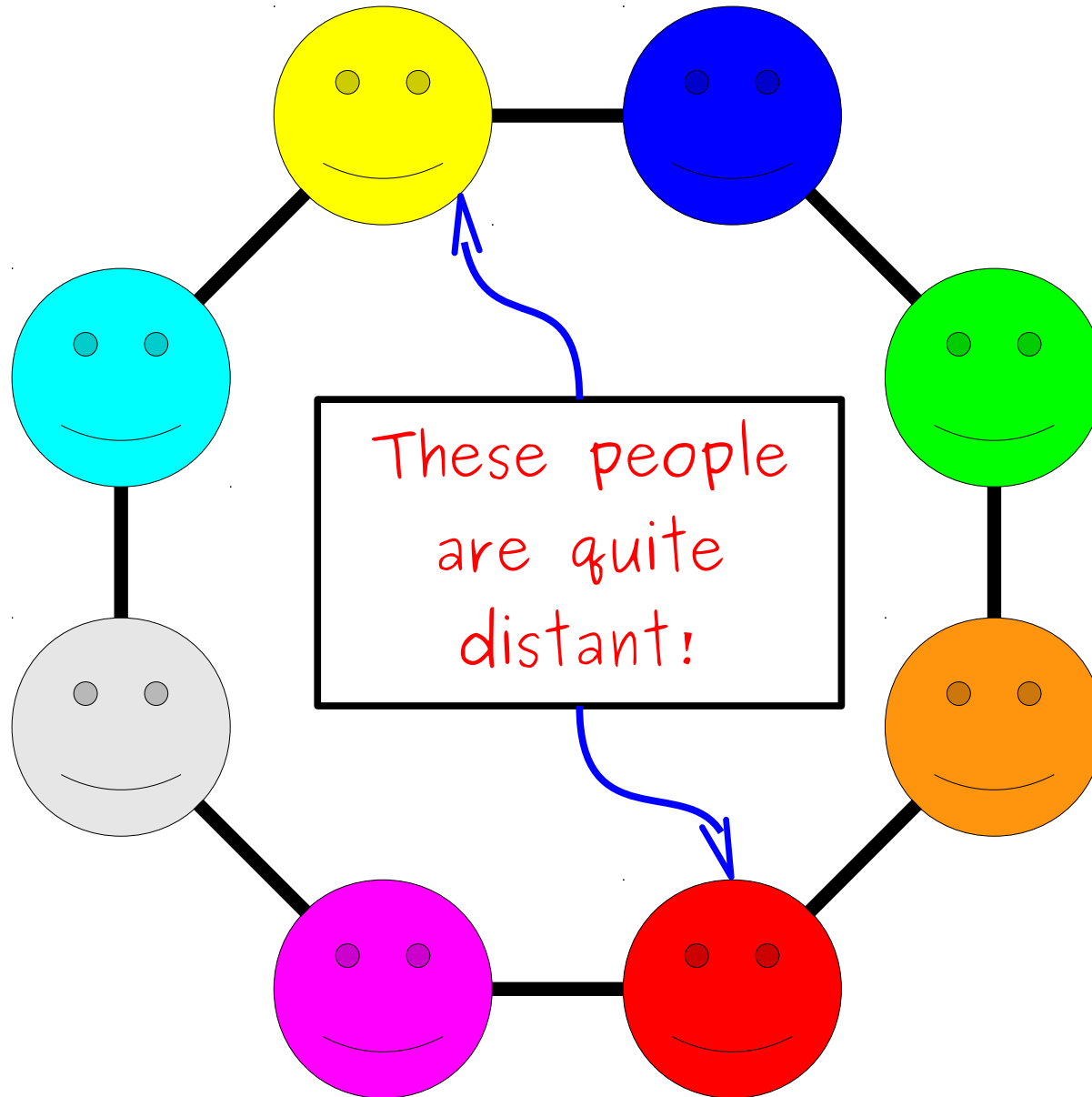
- We can analyze how nodes in a graph are connected to learn more about the graph.
- How connected are the nodes in the graph?
- How important is each node in the graph?



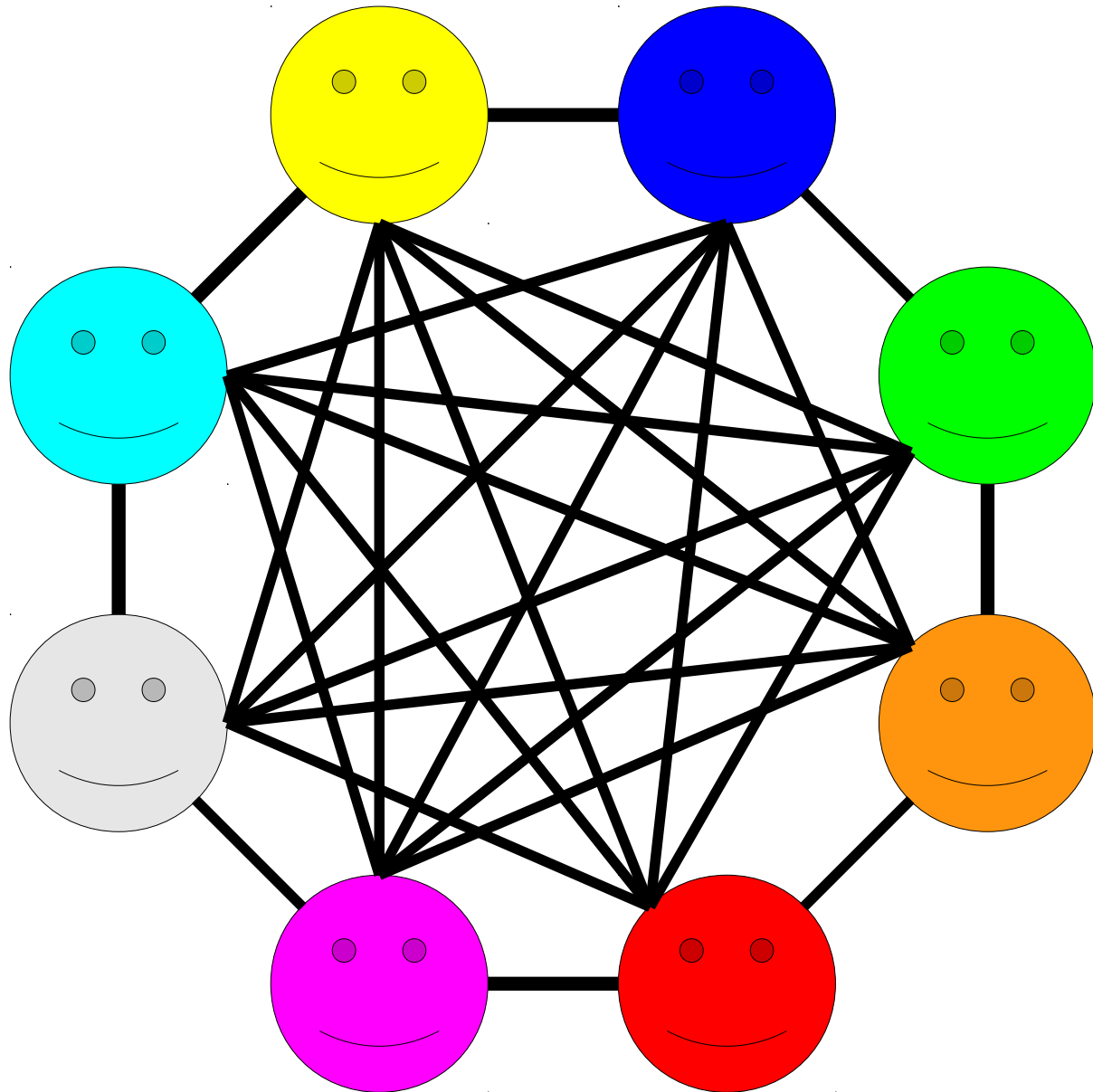
# Connectivity



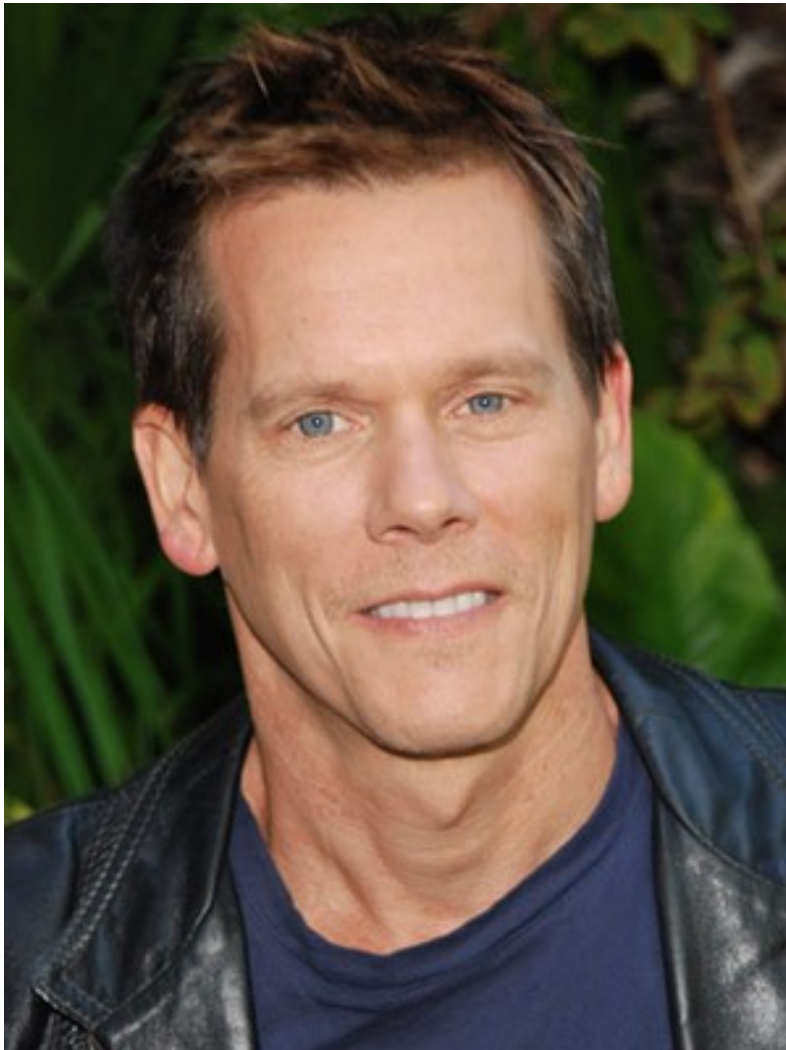
# Connectivity



# Connectivity



# Network Connectivity



- All actors and actresses have a **Bacon number** describing how removed they are from Kevin Bacon.
- Less than 1% of all actors and actresses have a Bacon number greater than six.

# Finding Important Nodes

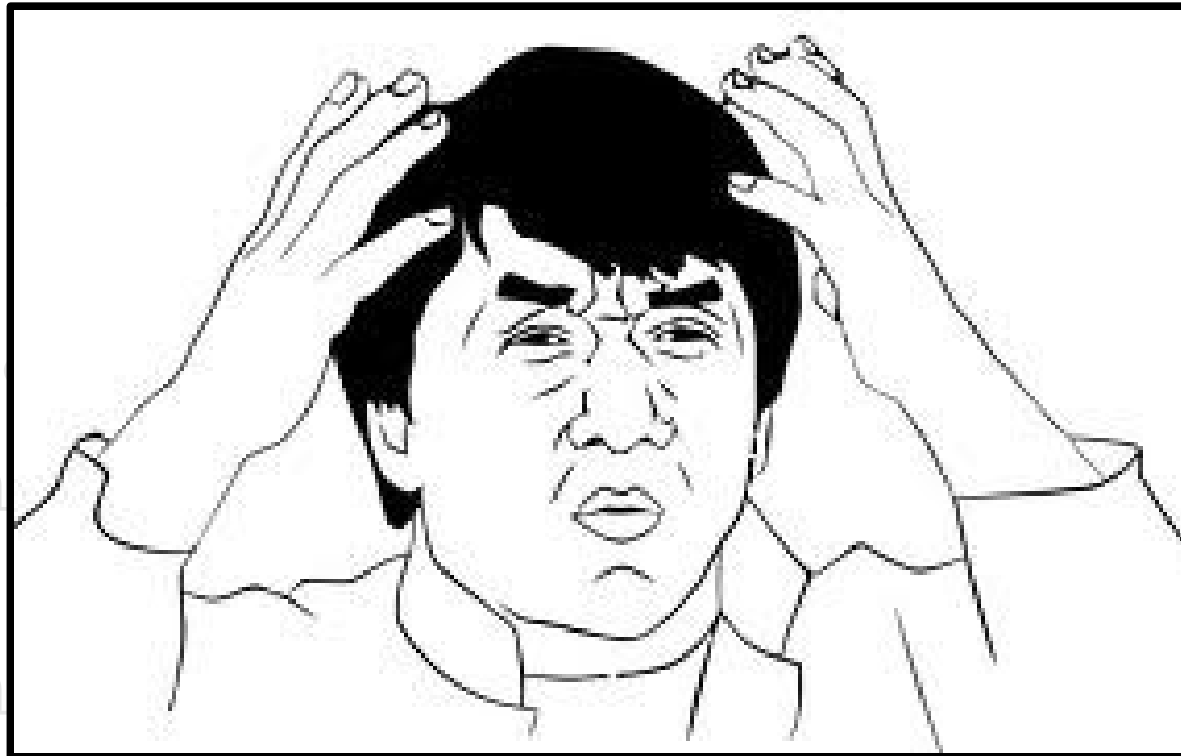
- Suppose that we want to have the computer find “important” articles on Wikipedia.
- We just have the link structure, not the text of the page, the number of edits, the length of the article, etc.
- How might we do this?

# Link Analysis

- To find important Wikipedia pages, let's look at the links between pages.
- We'll make two assumptions:
  - The more important an article is, the more pages will link to it.
  - The more important an article is, the more that its links matter.
- **An article is important if other important articles link to it.**



# Link Analysis



- **An article is important if other important articles link to it.**

# The Random Surfer Model

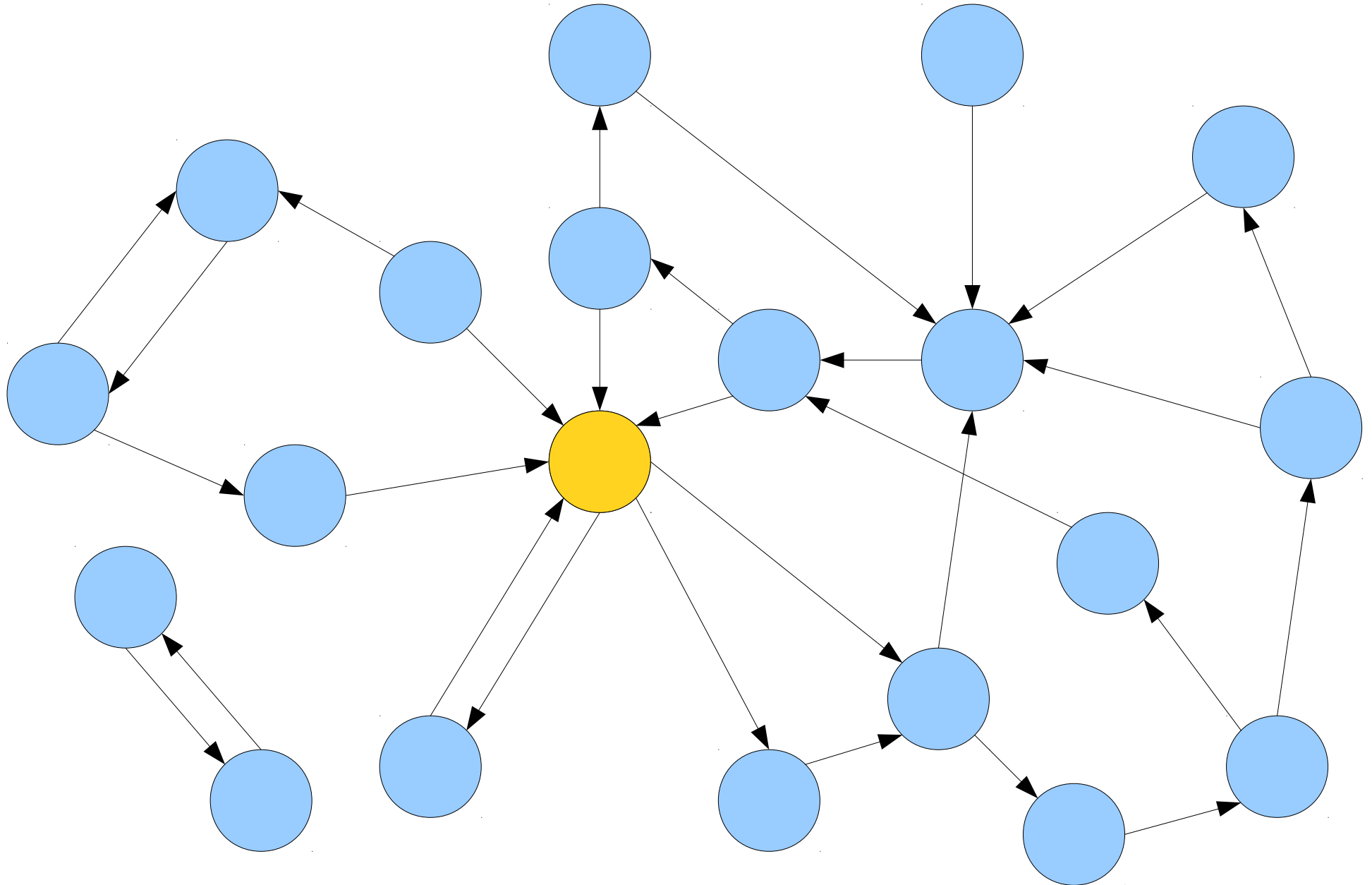


(seriously though)

# The Random Surfer Model

- Think about the behavior of a Wikipedia reader who randomly surfs Wikipedia.
- Visits some initial page at random.
- From there, the user either
  - Clicks a random link on the page to some other article, or
  - hits the “random page” link to visit a totally random page.

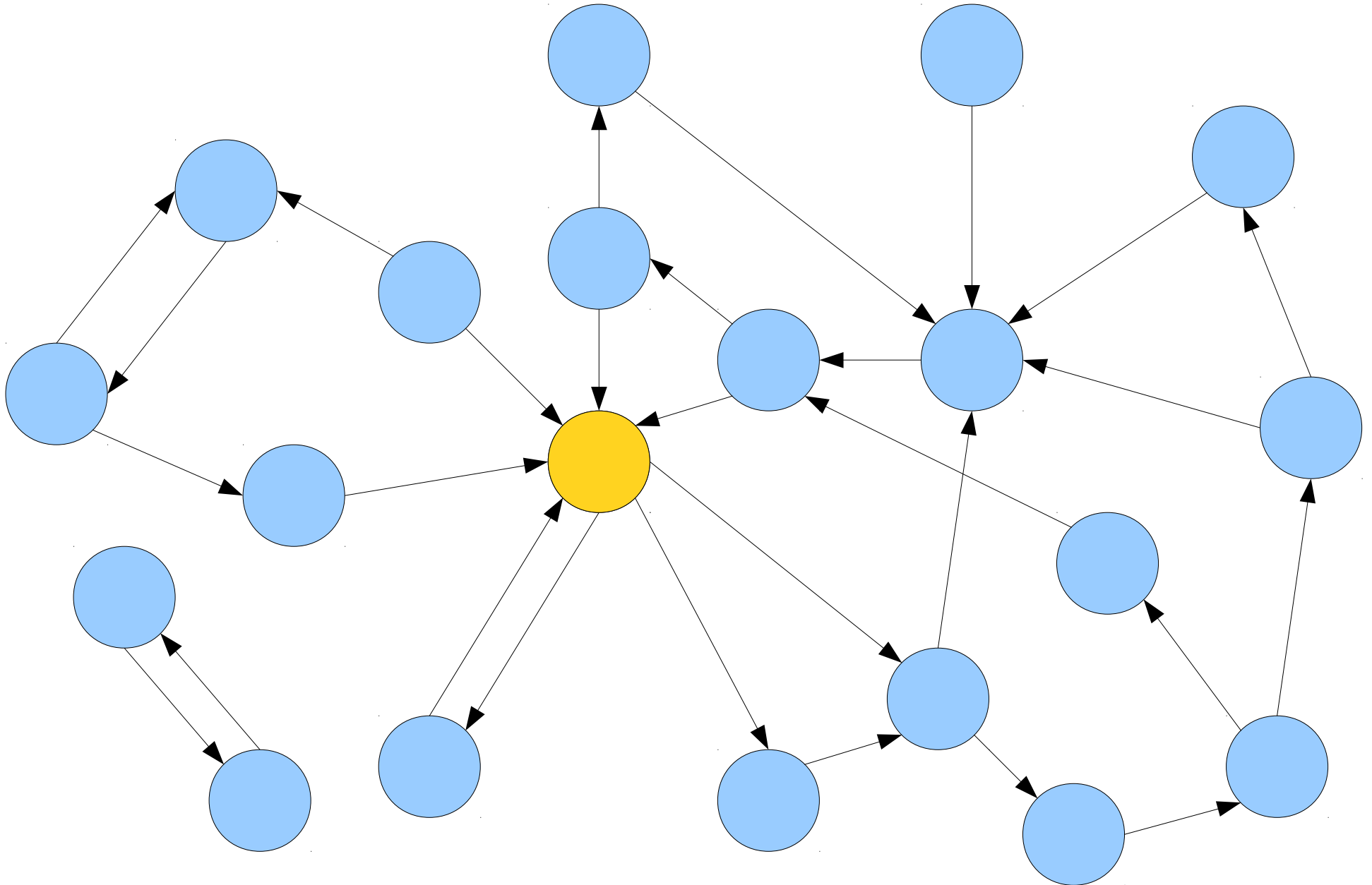
# The Random Surfer Model



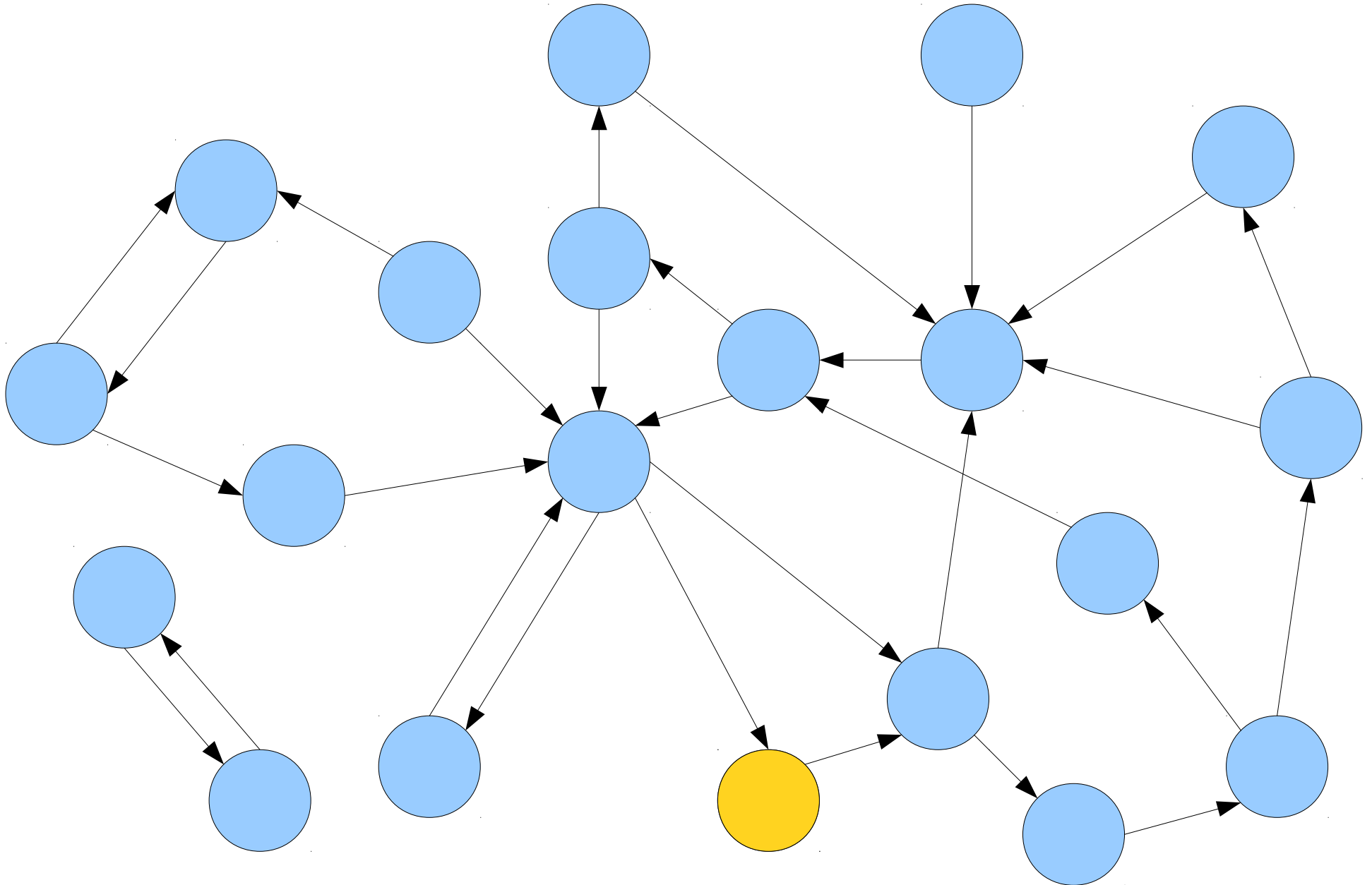




# The Random Surfer Model

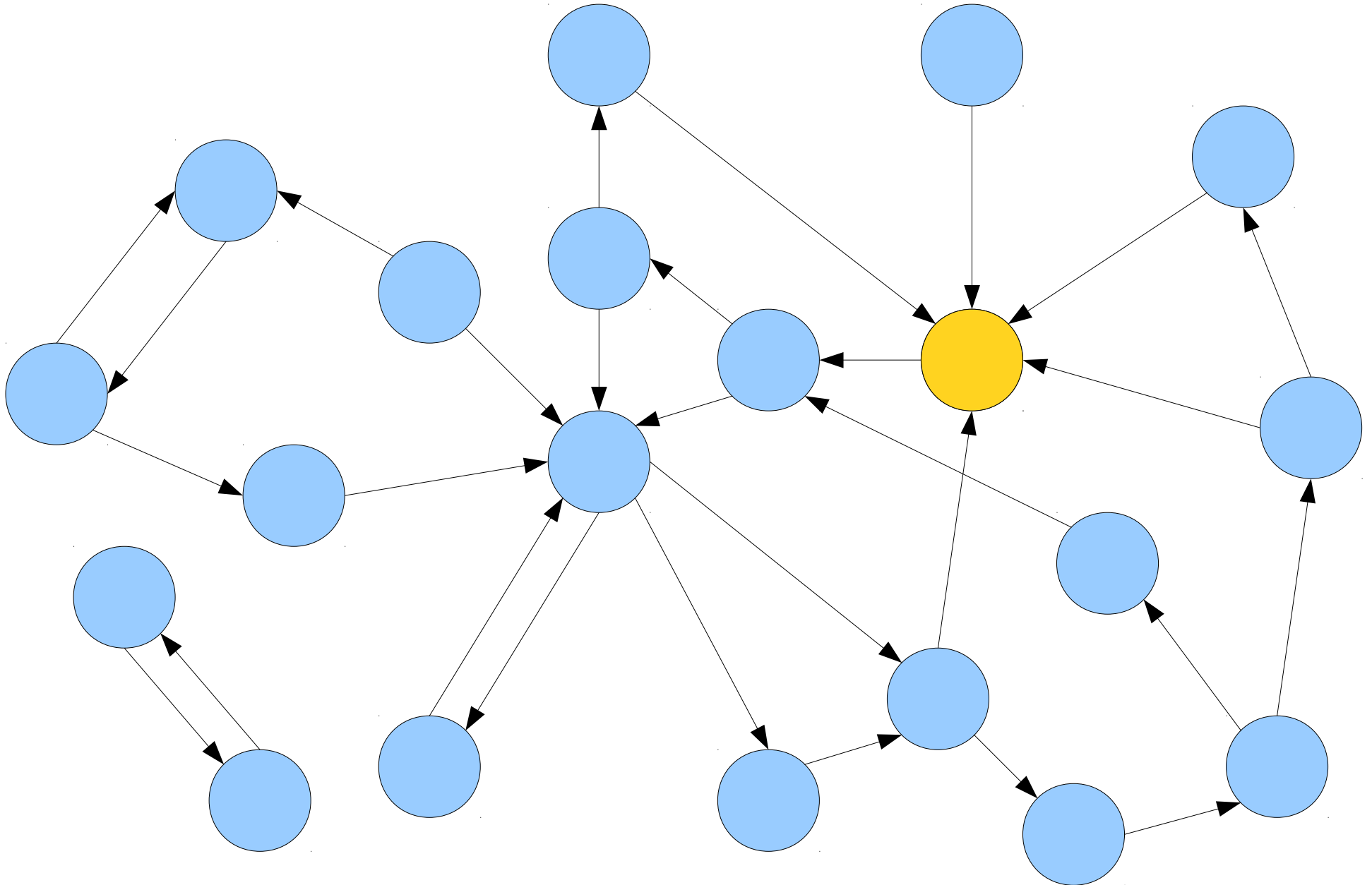


# The Random Surfer Model



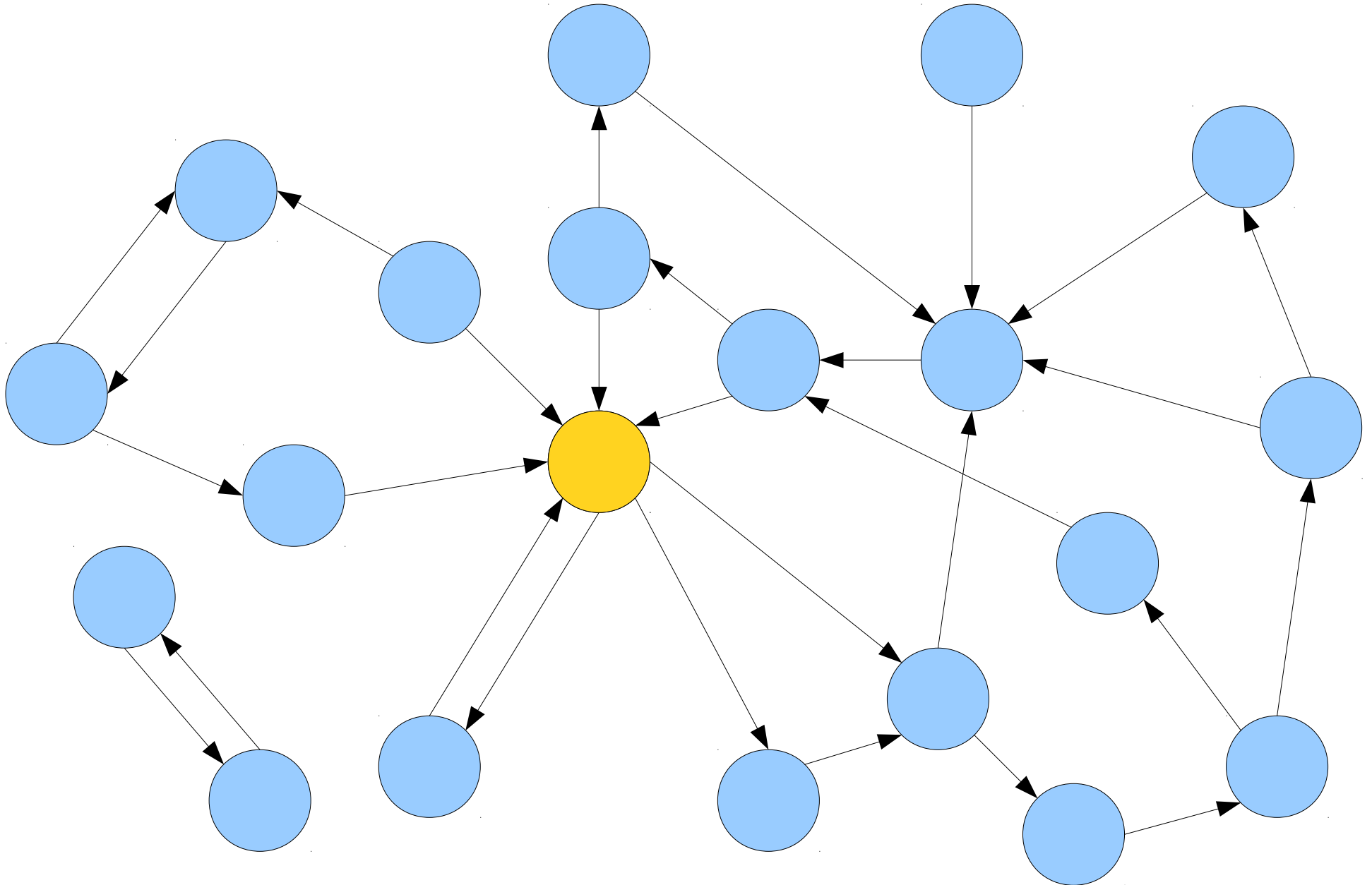


# The Random Surfer Model

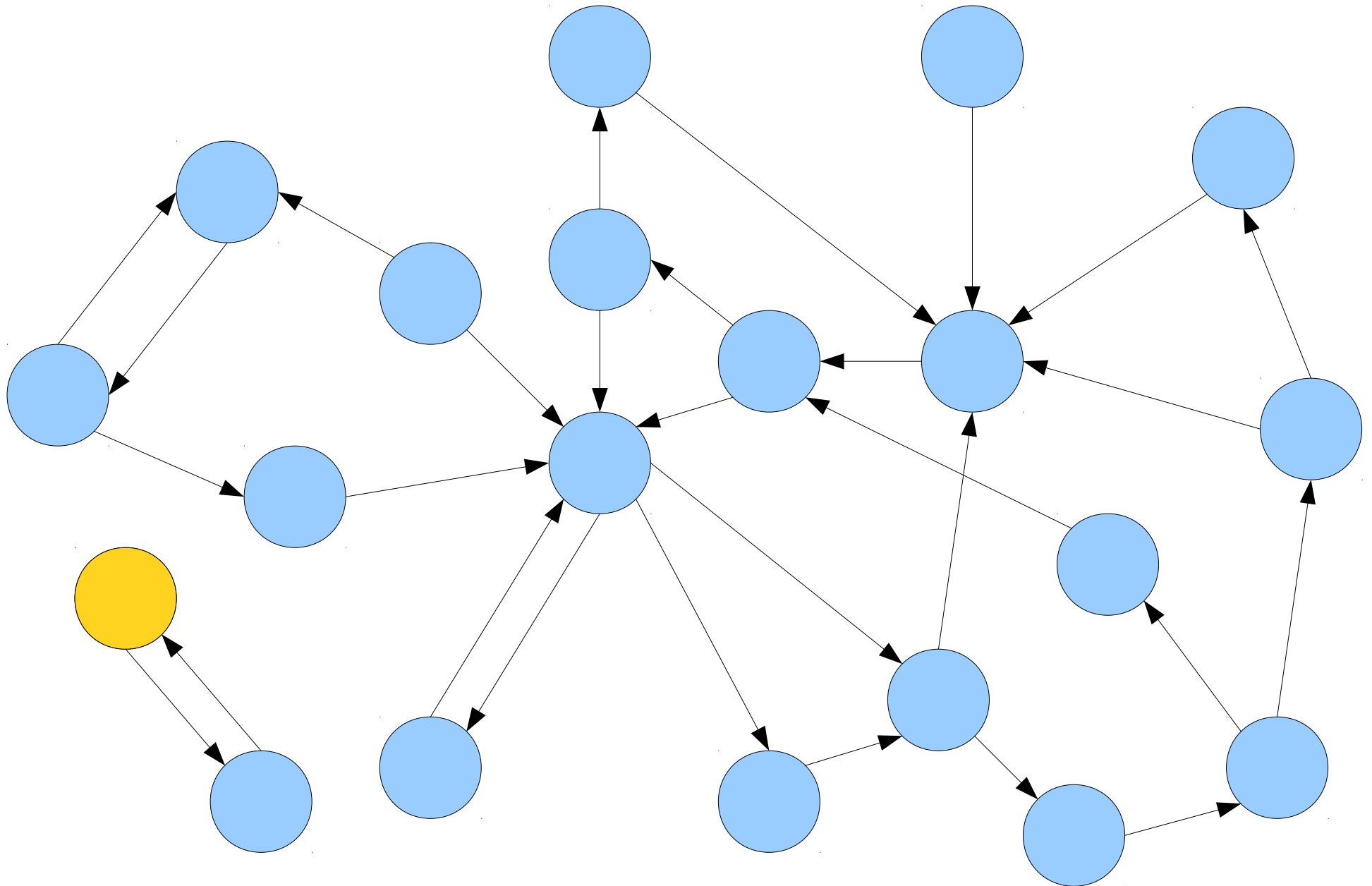




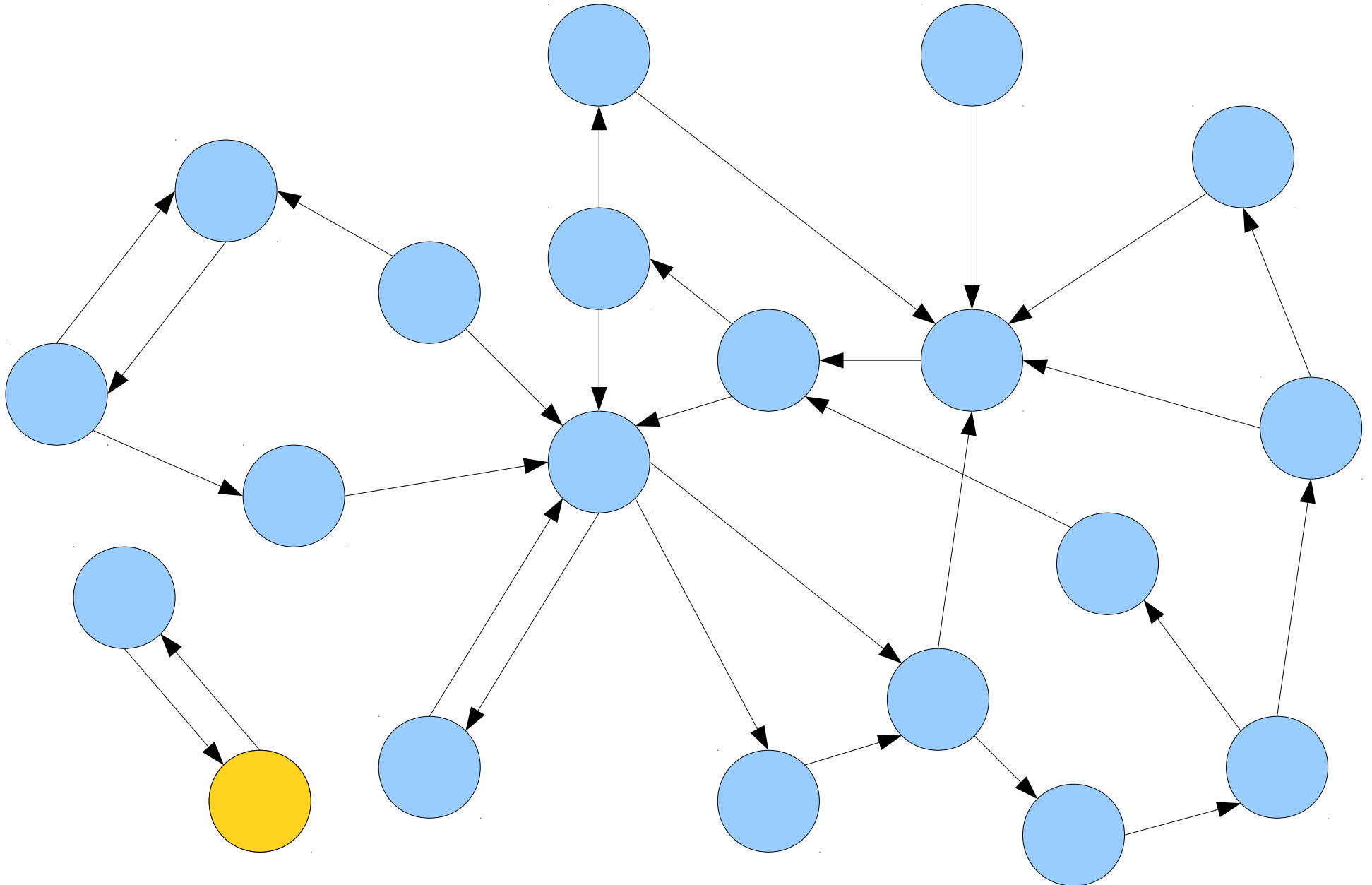
# The Random Surfer Model



# The Random Surfer Model

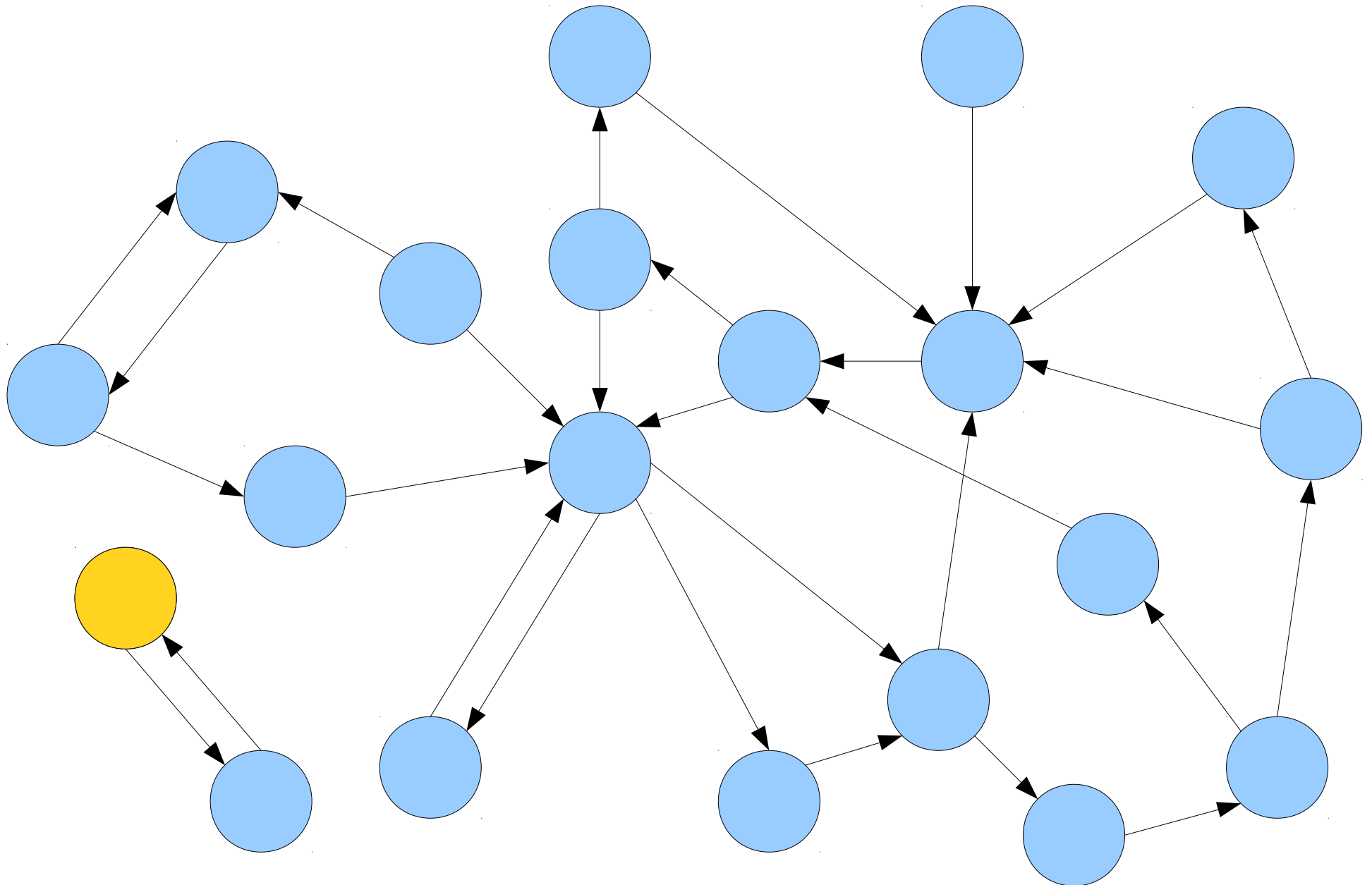


# The Random Surfer Model

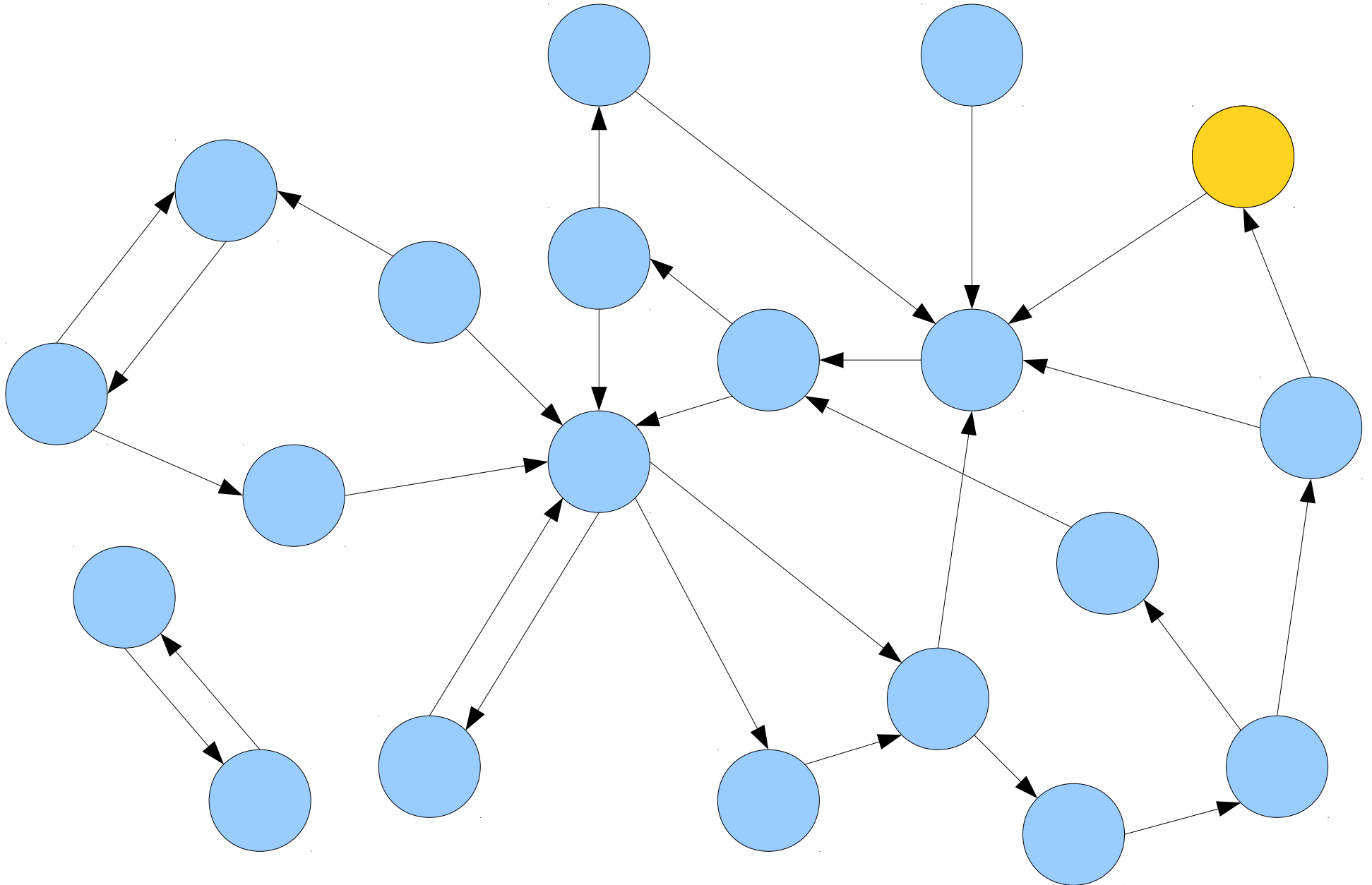




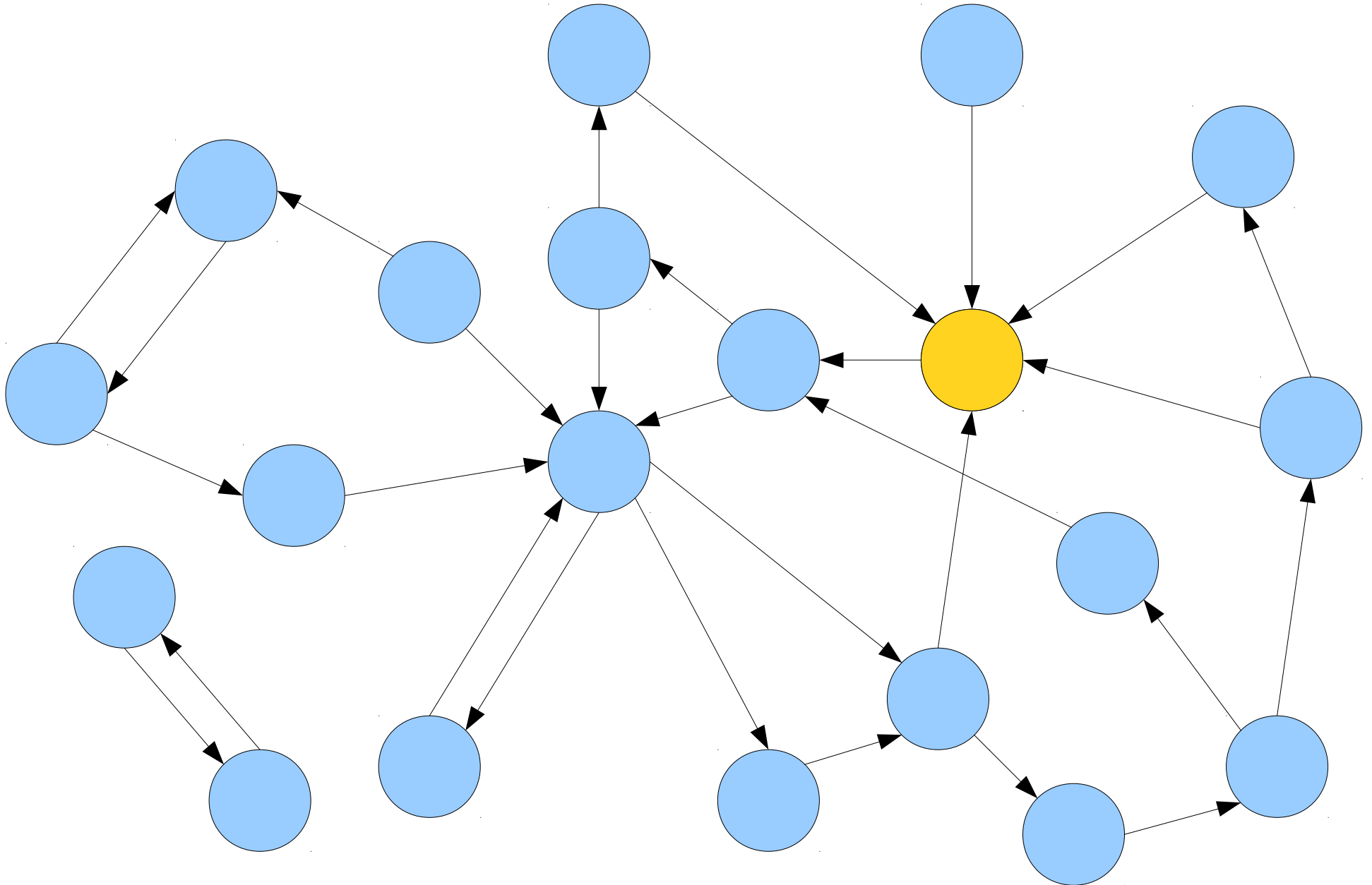
# The Random Surfer Model



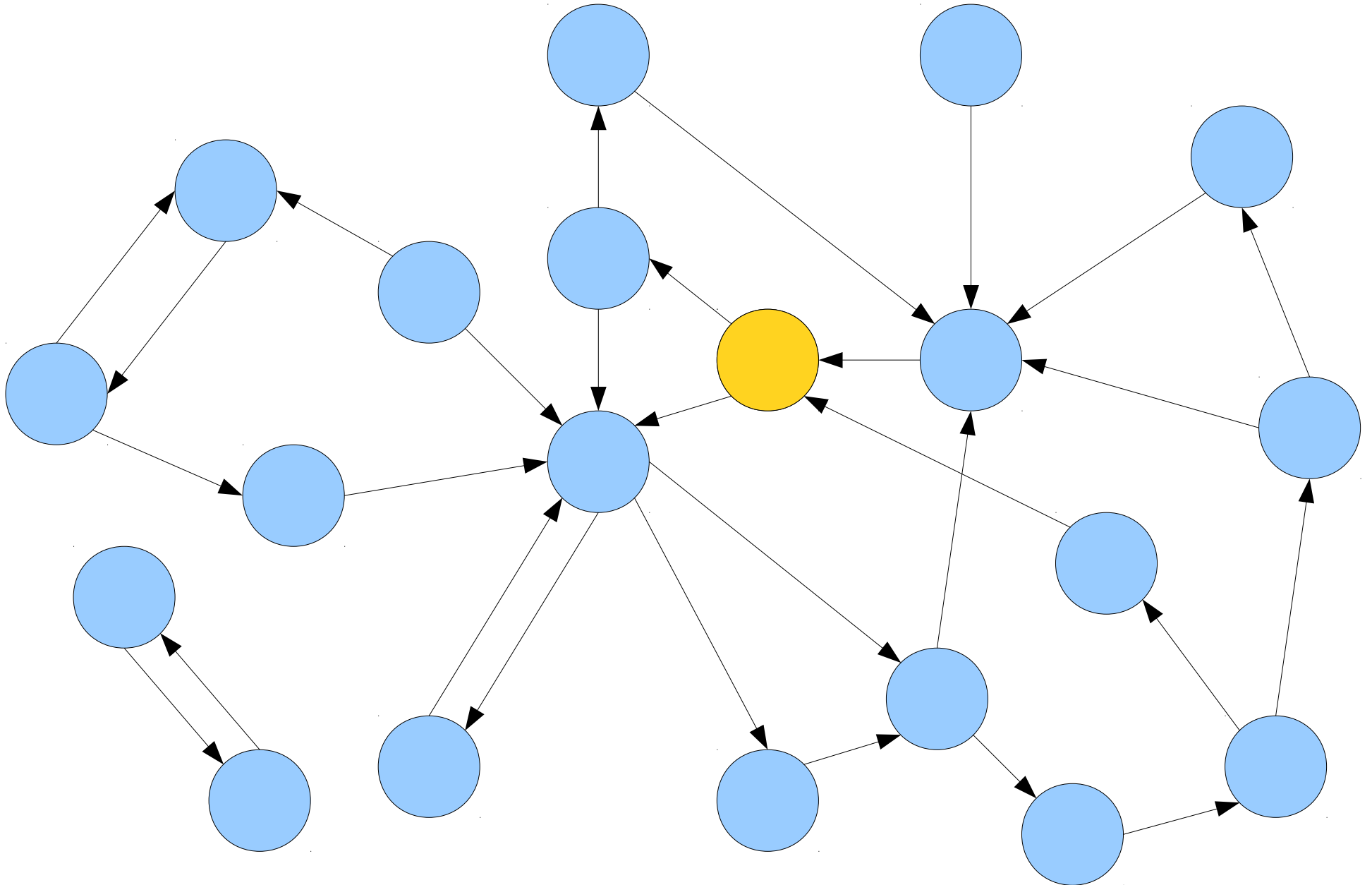
# The Random Surfer Model



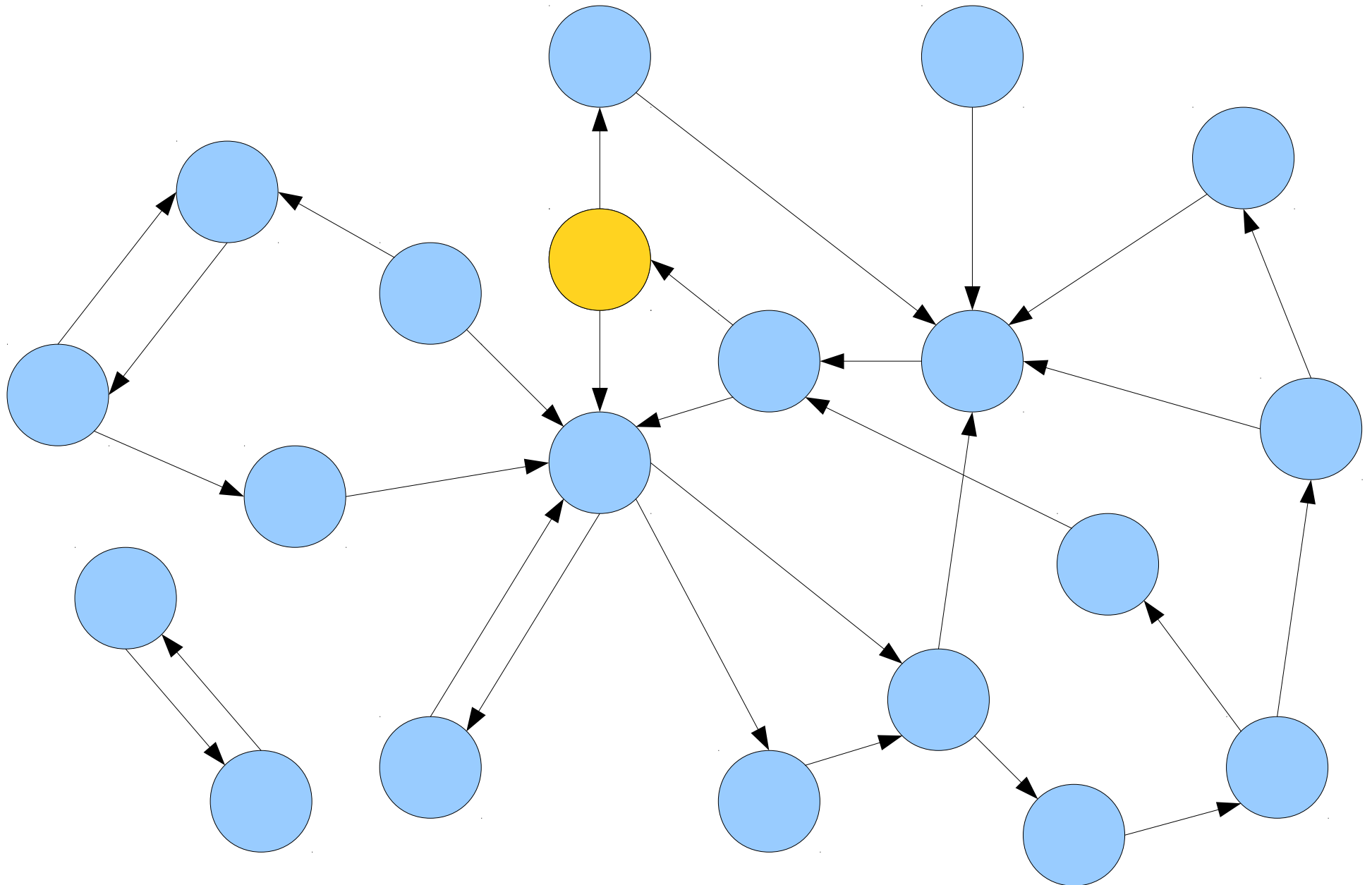
# The Random Surfer Model



# The Random Surfer Model

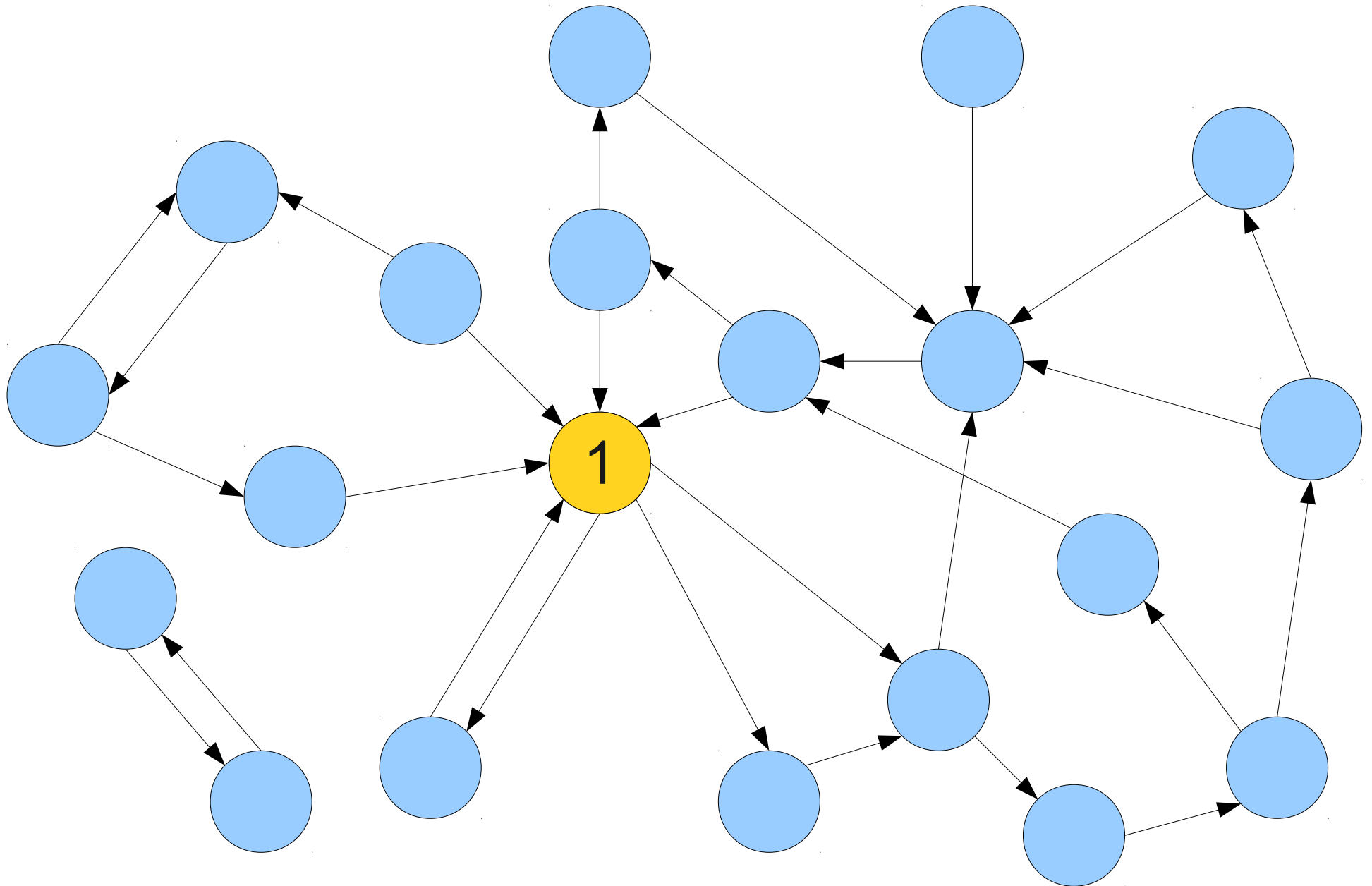


# The Random Surfer Model

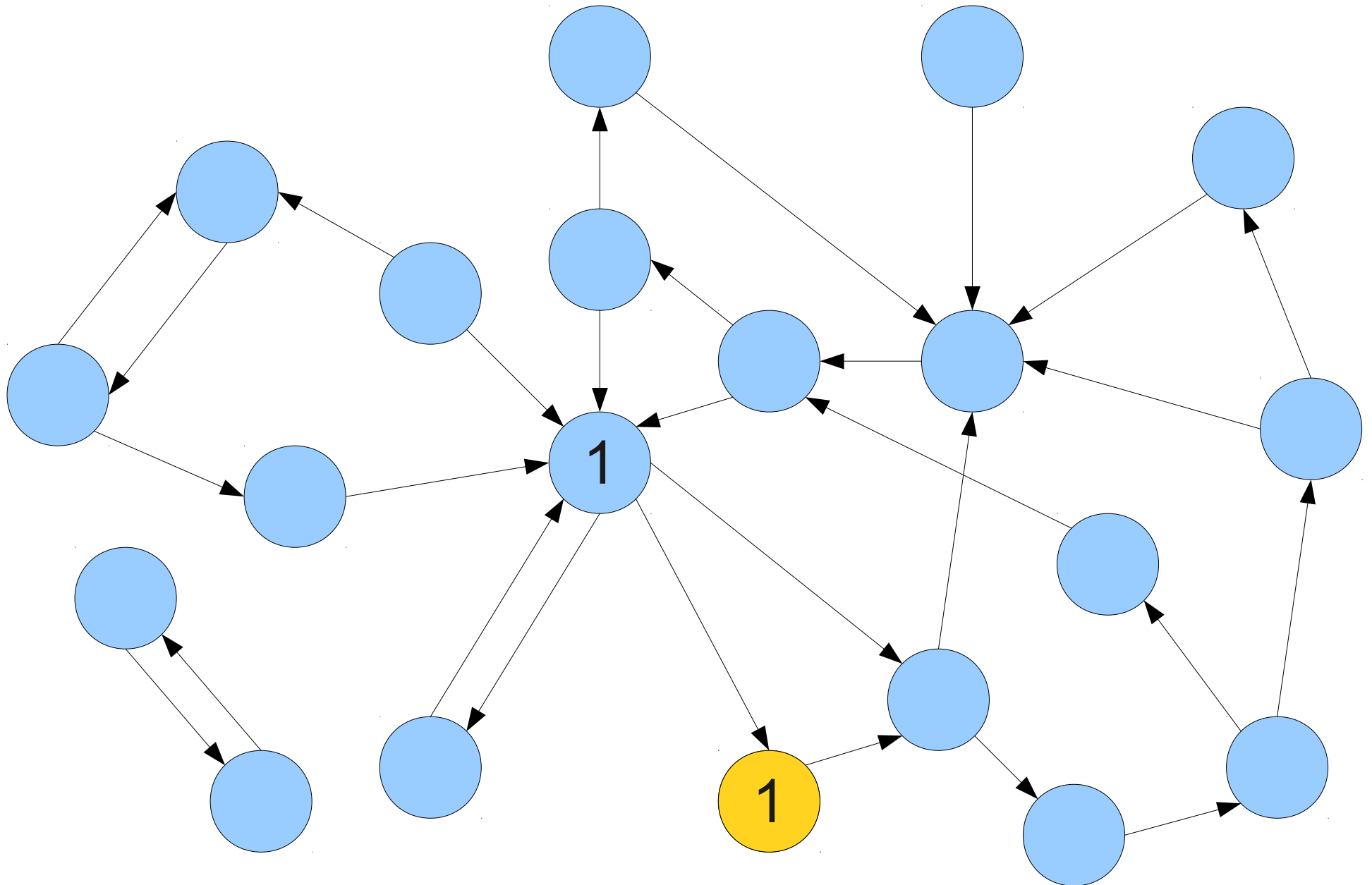




# The Random Surfer Model

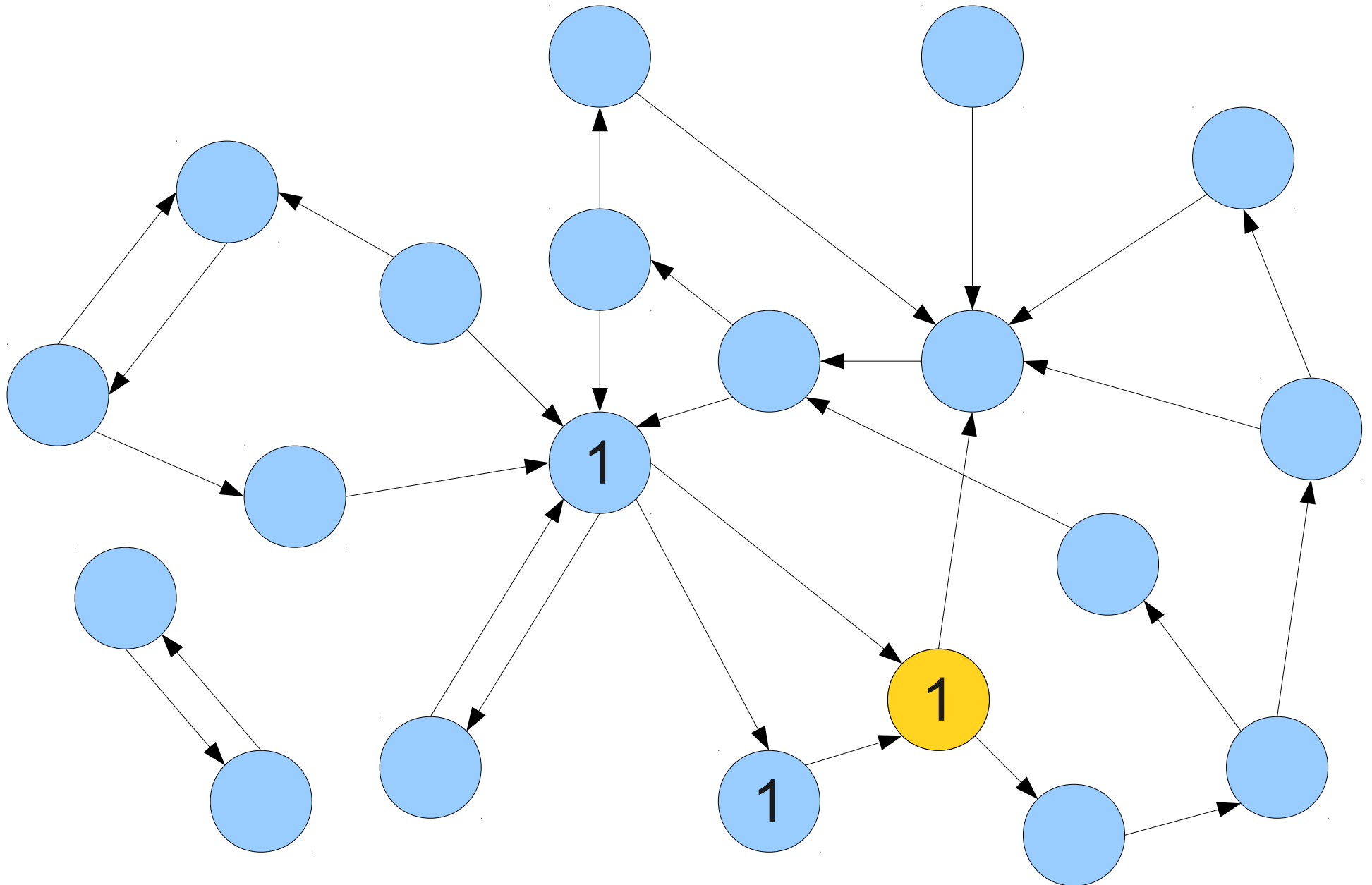


# The Random Surfer Model

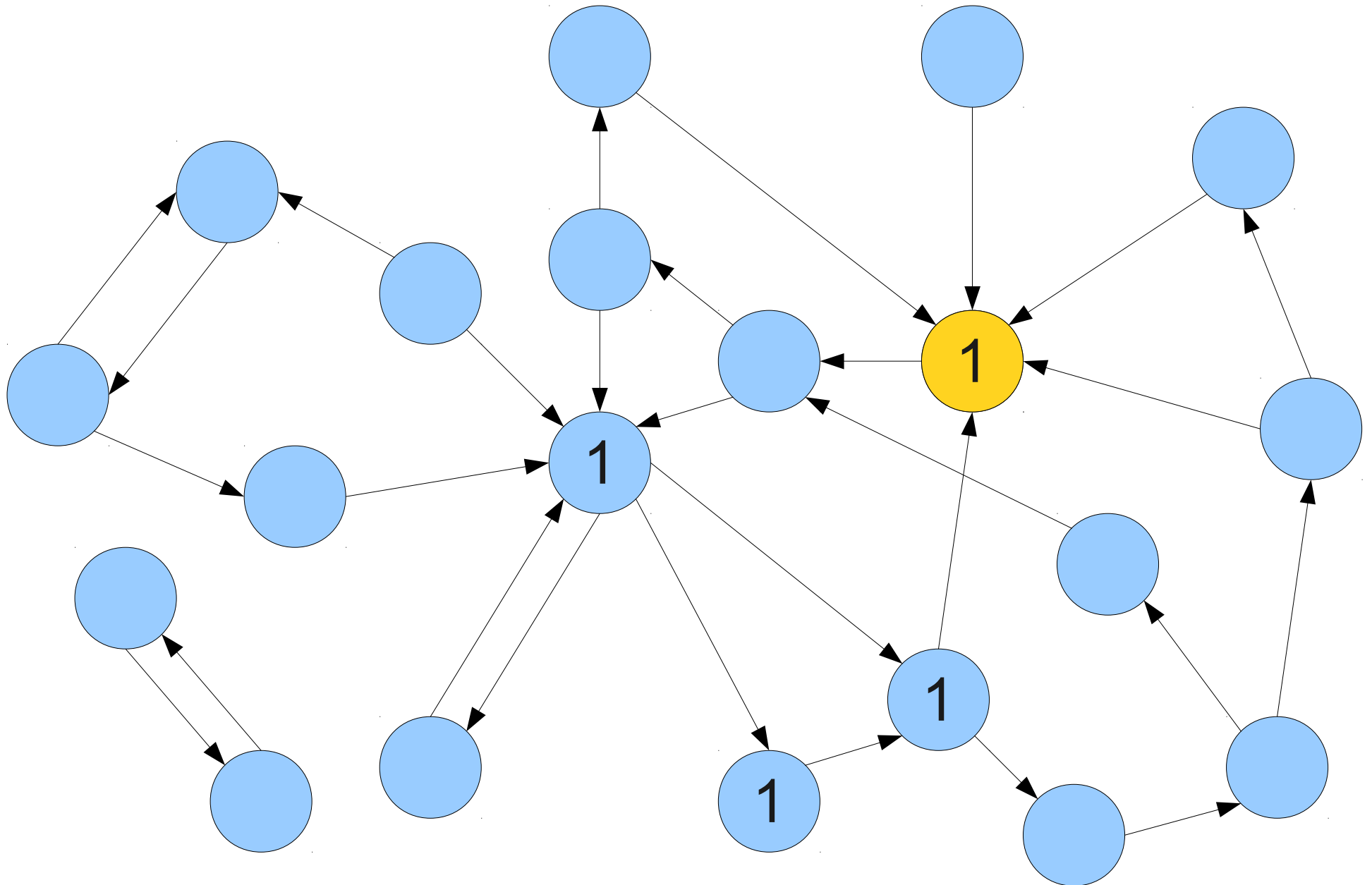




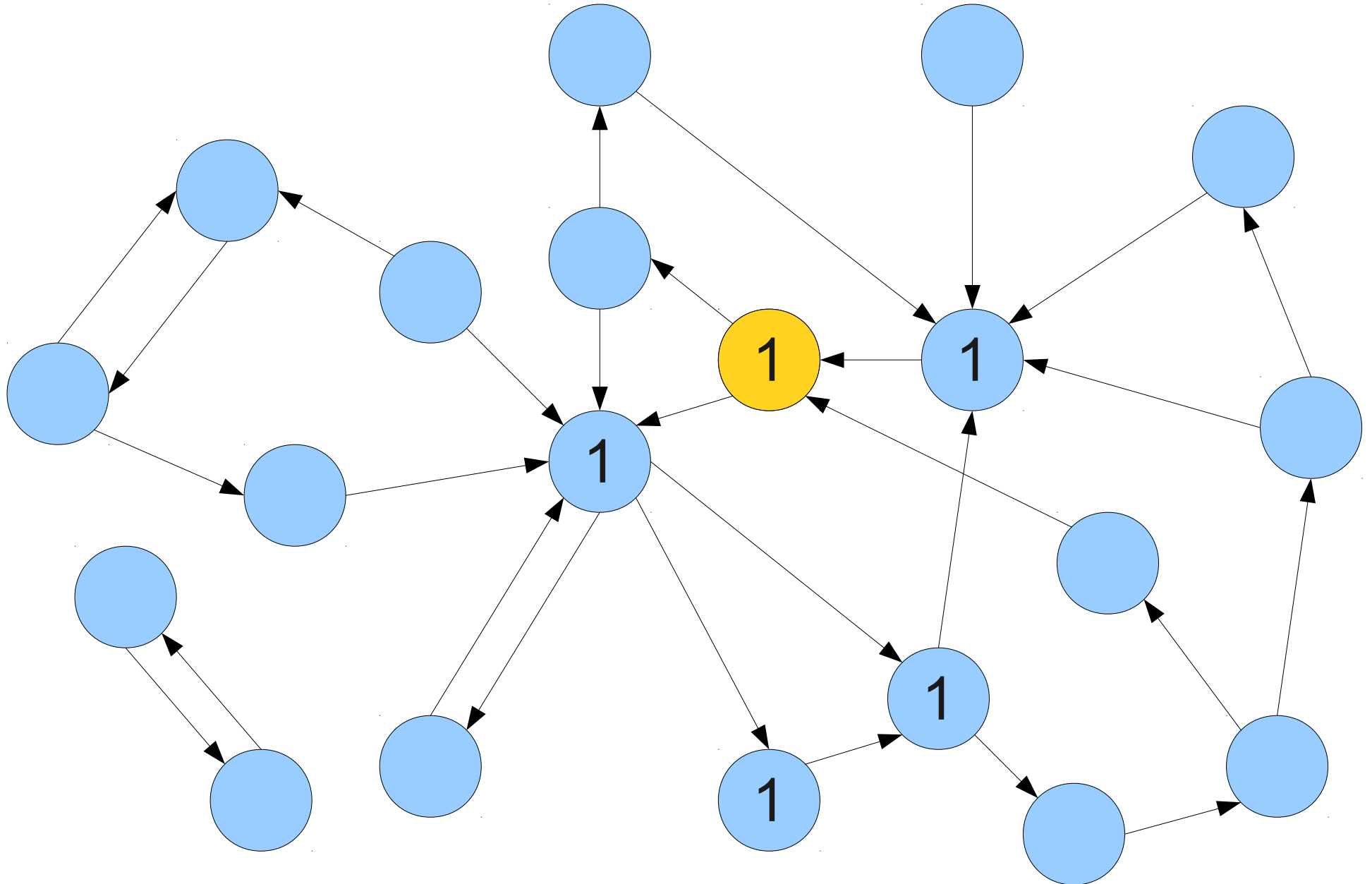
# The Random Surfer Model



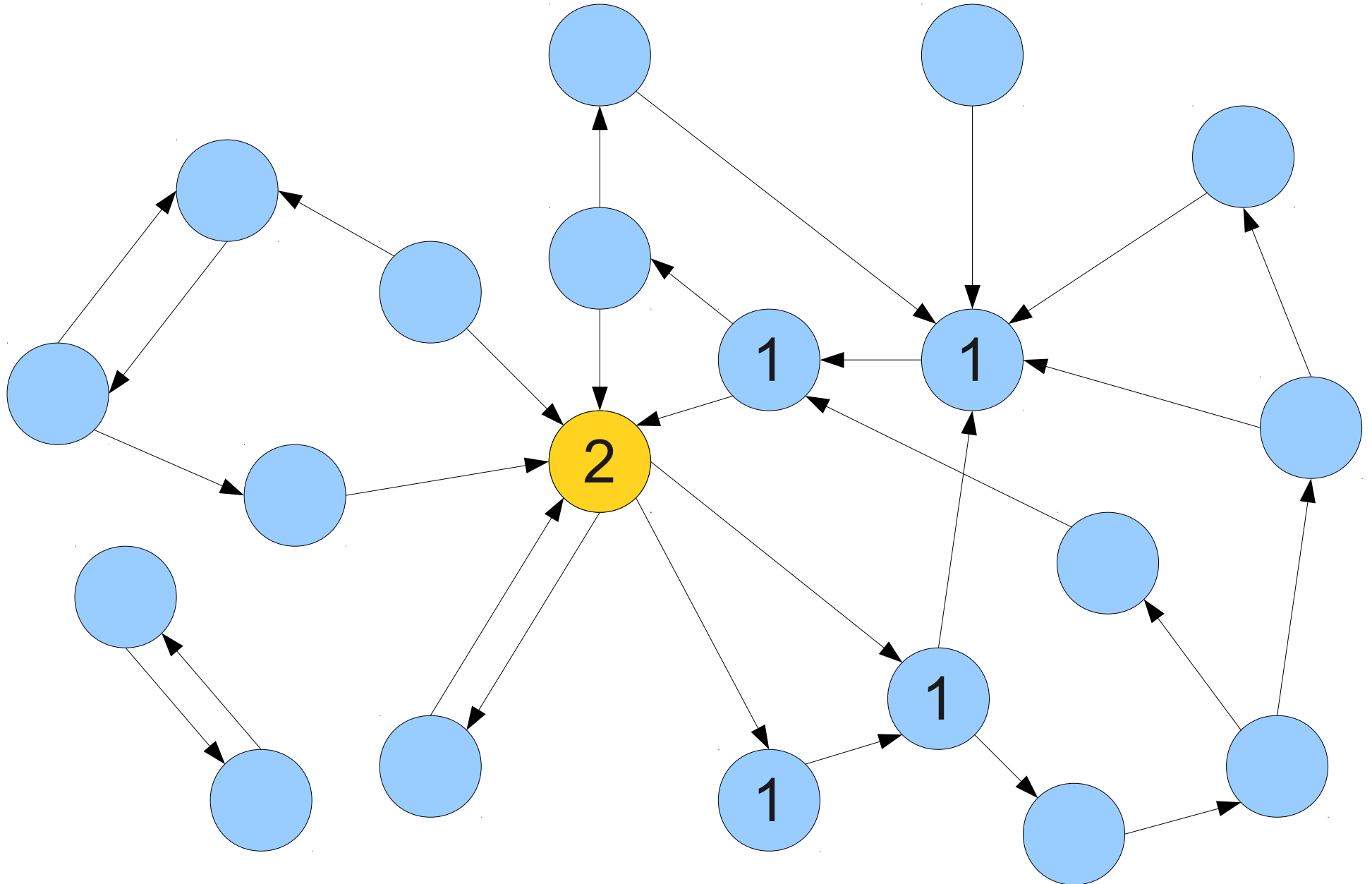
# The Random Surfer Model



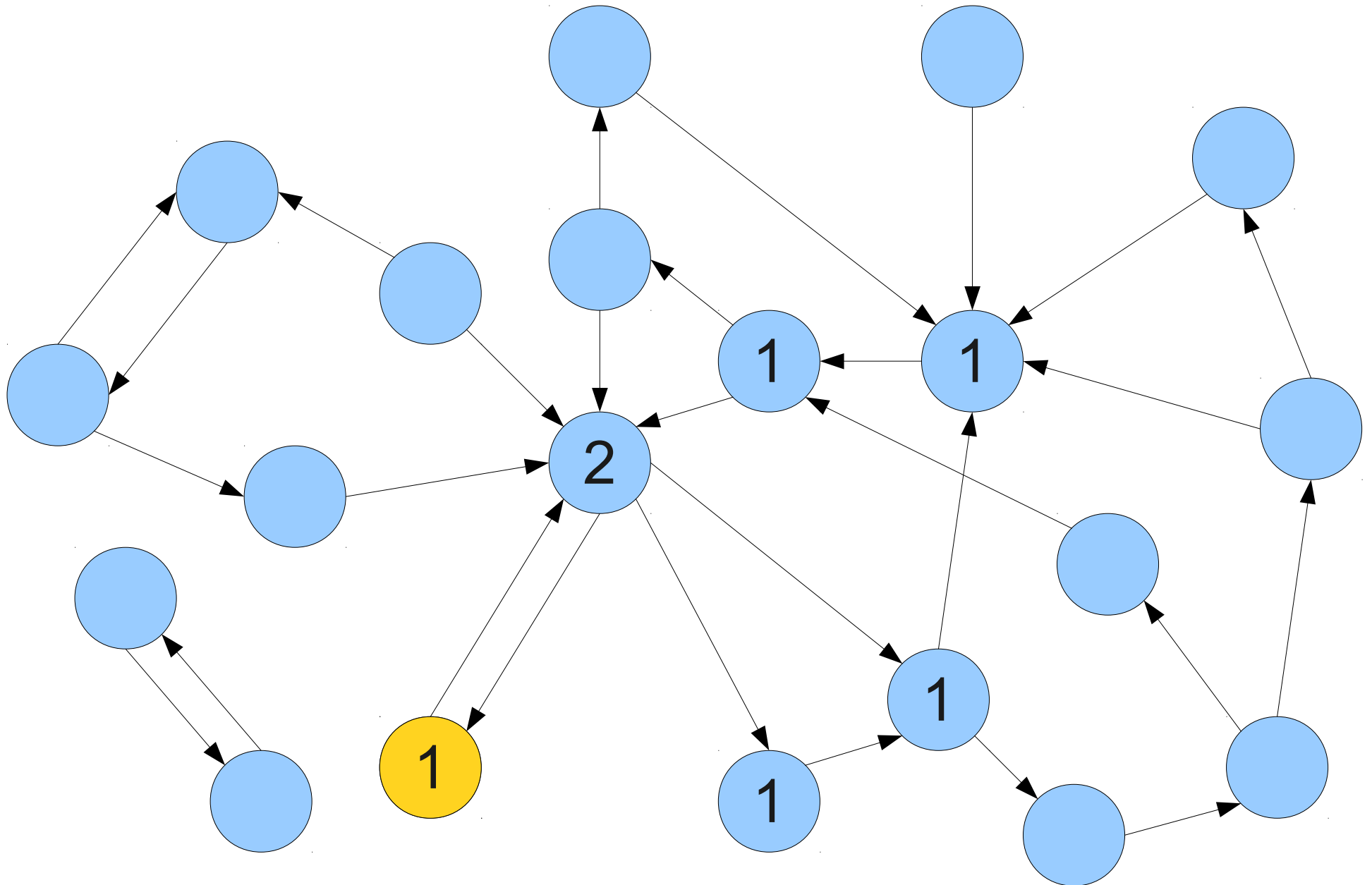
# The Random Surfer Model



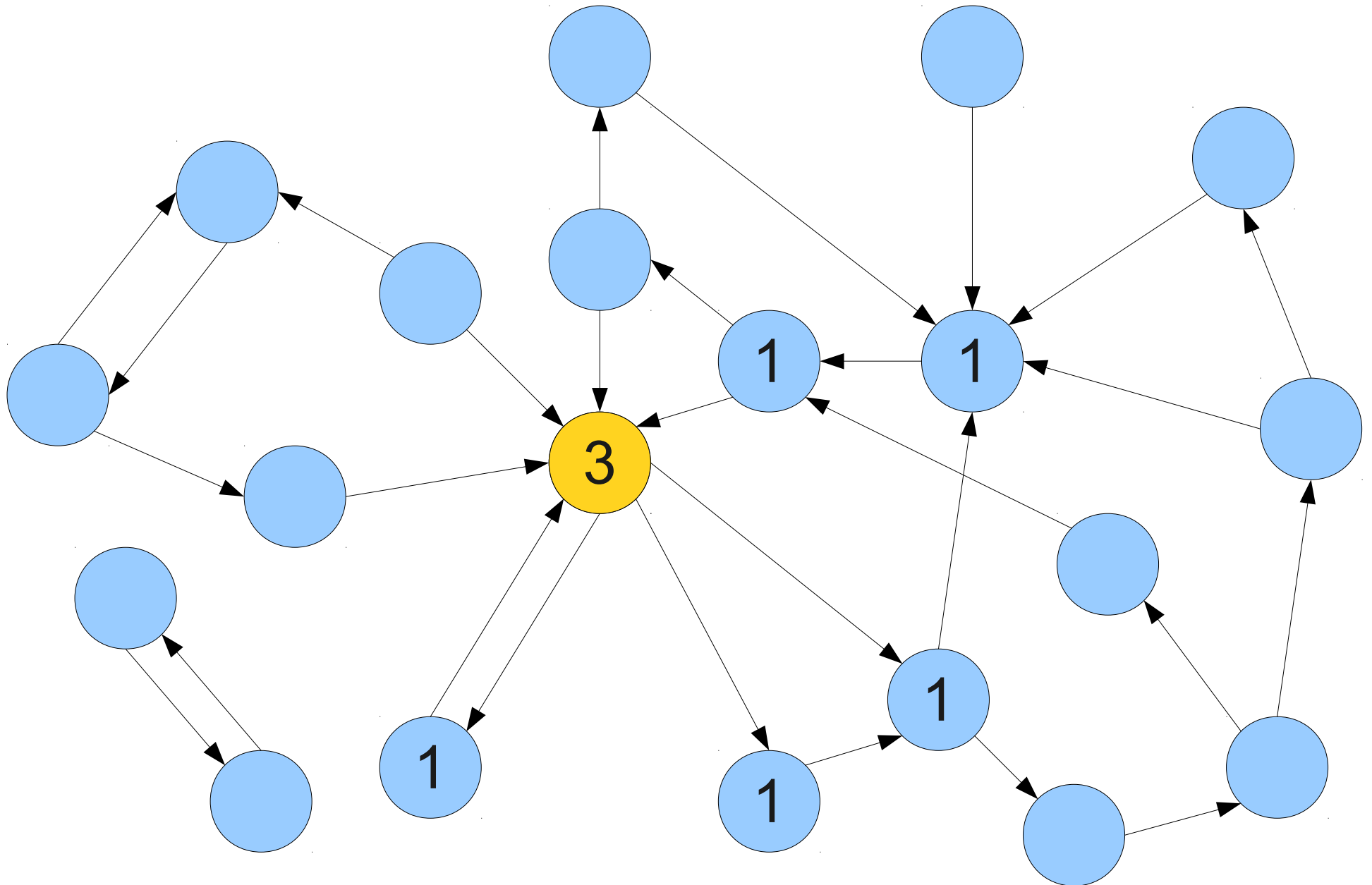
# The Random Surfer Model



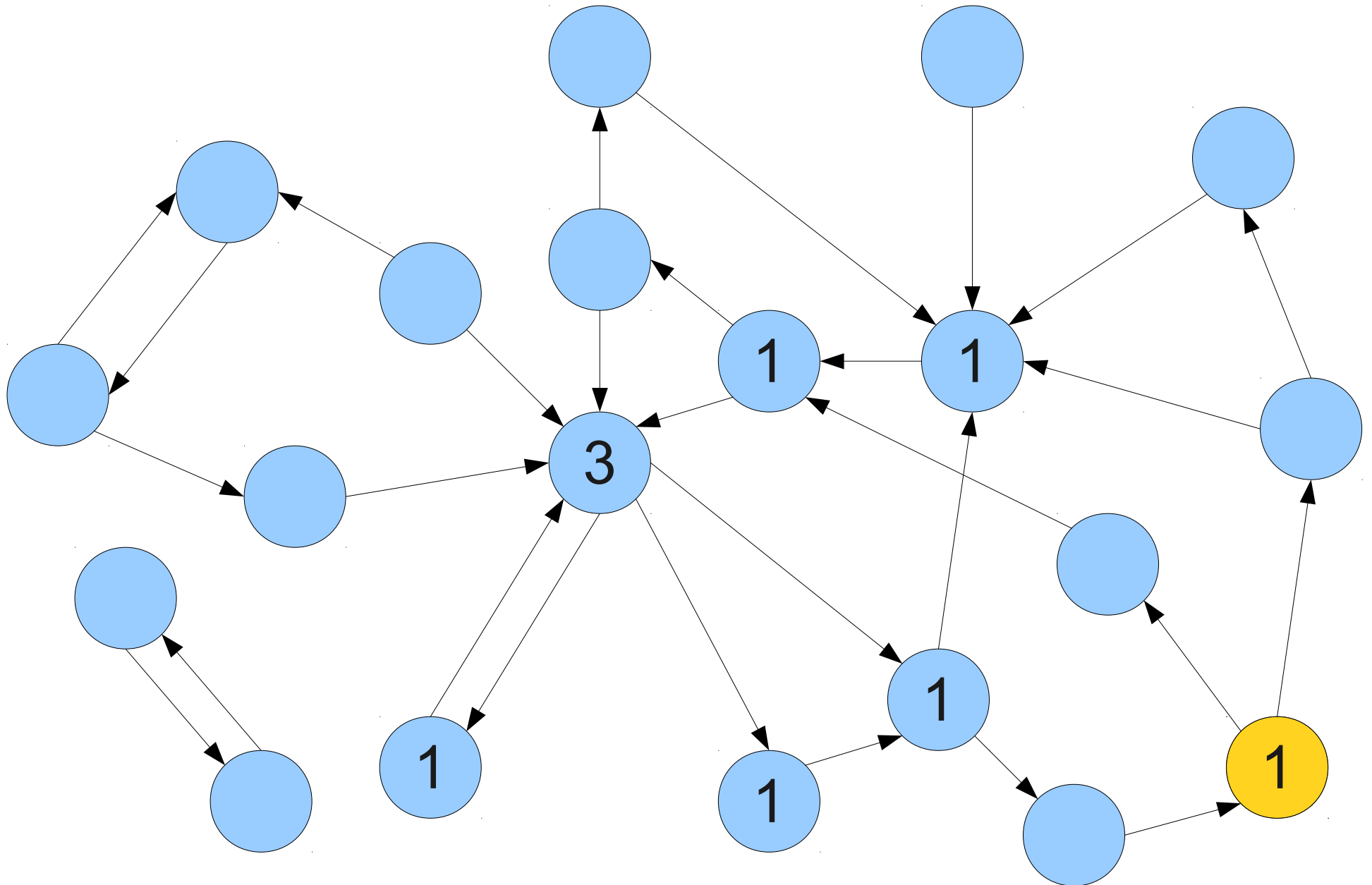
# The Random Surfer Model



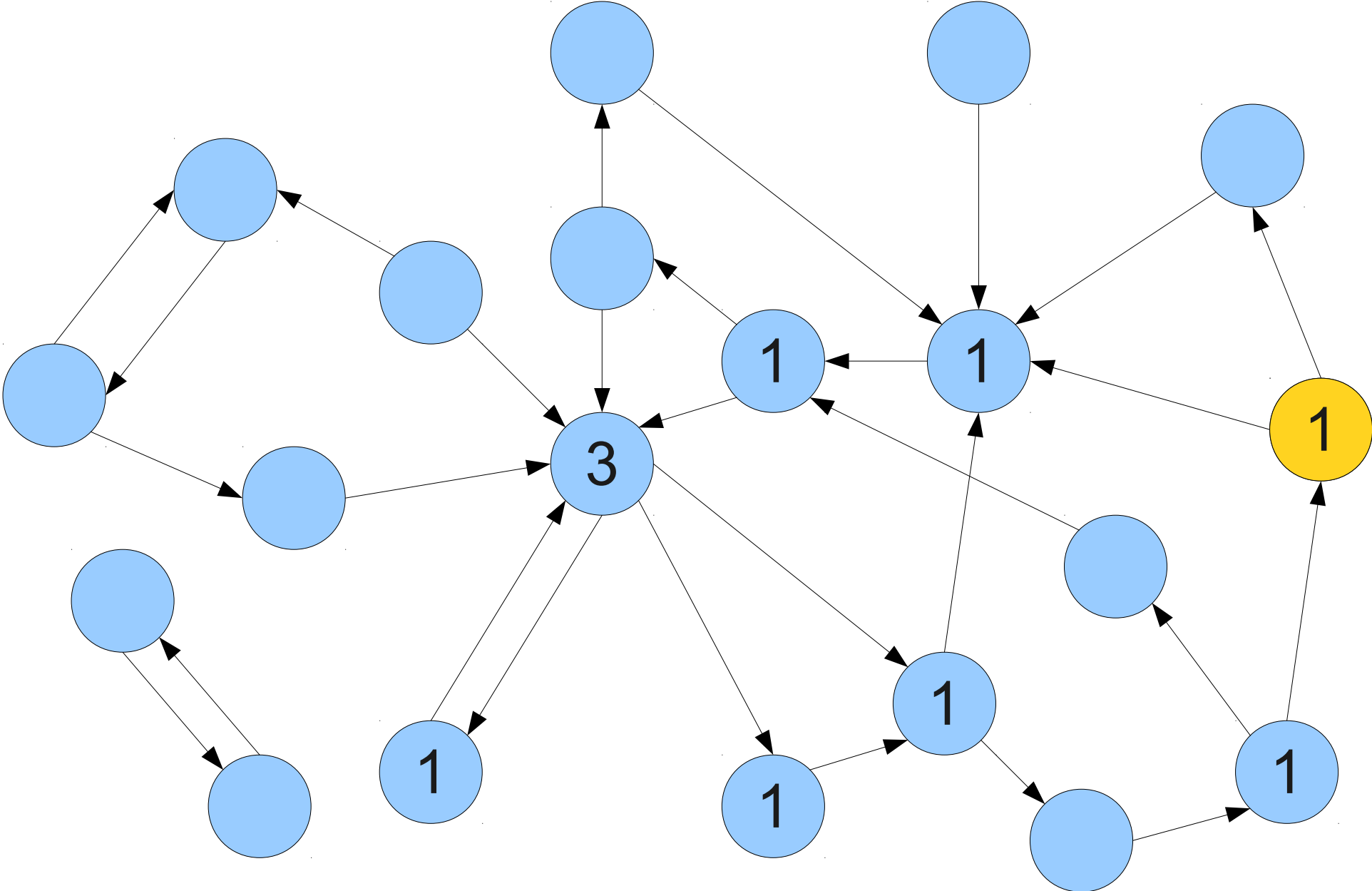
# The Random Surfer Model



# The Random Surfer Model



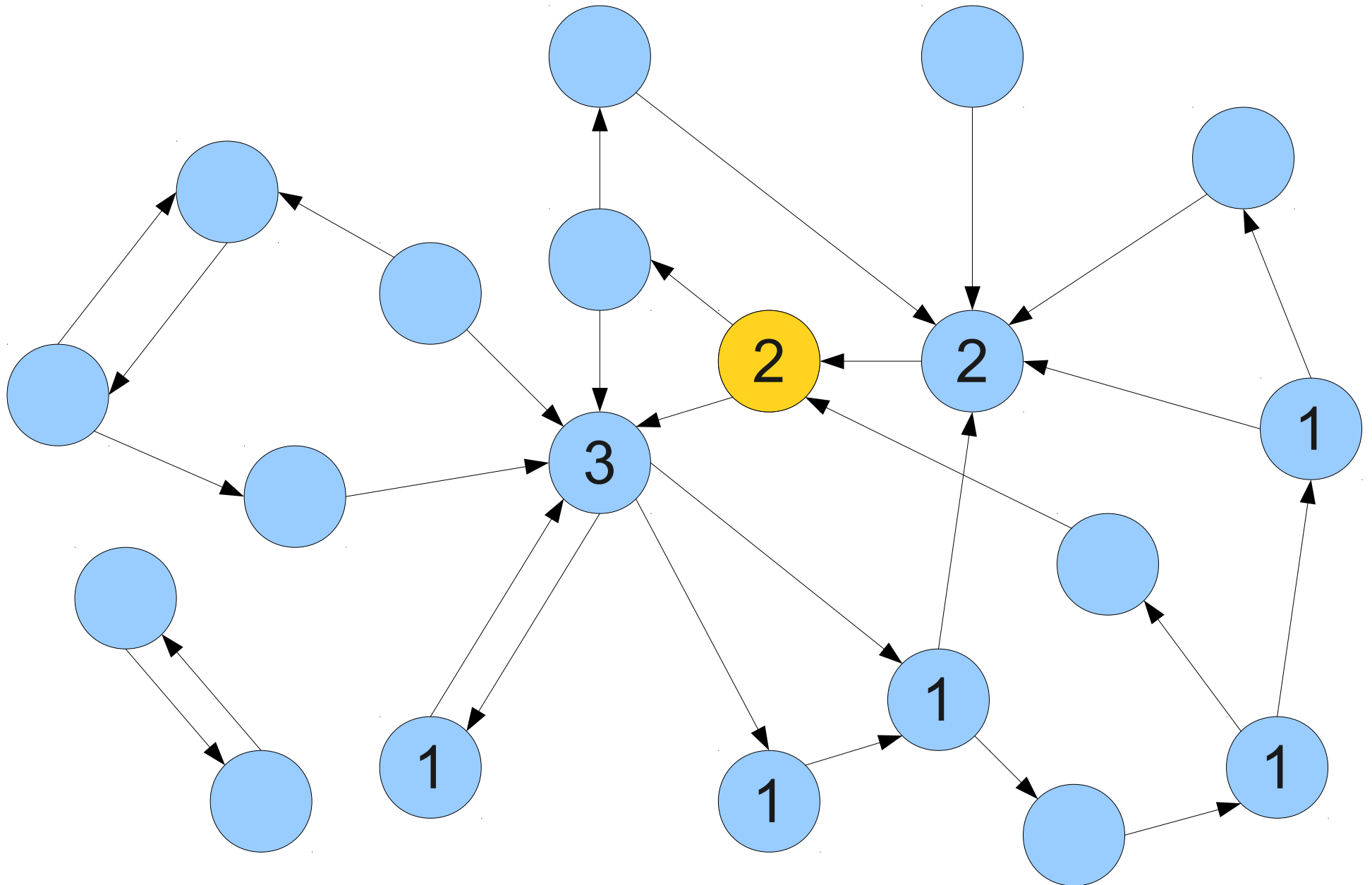
# The Random Surfer Model





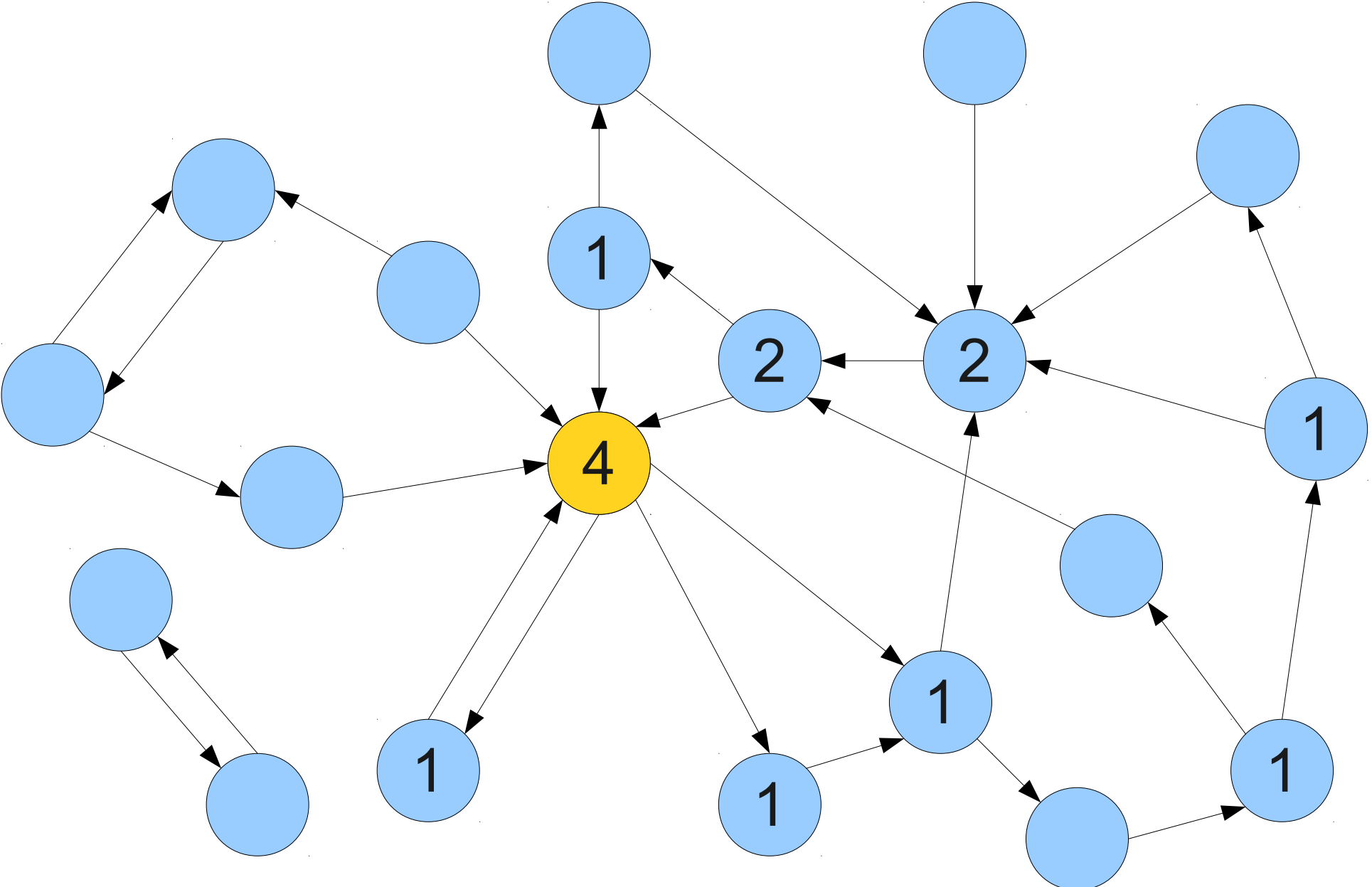


# The Random Surfer Model

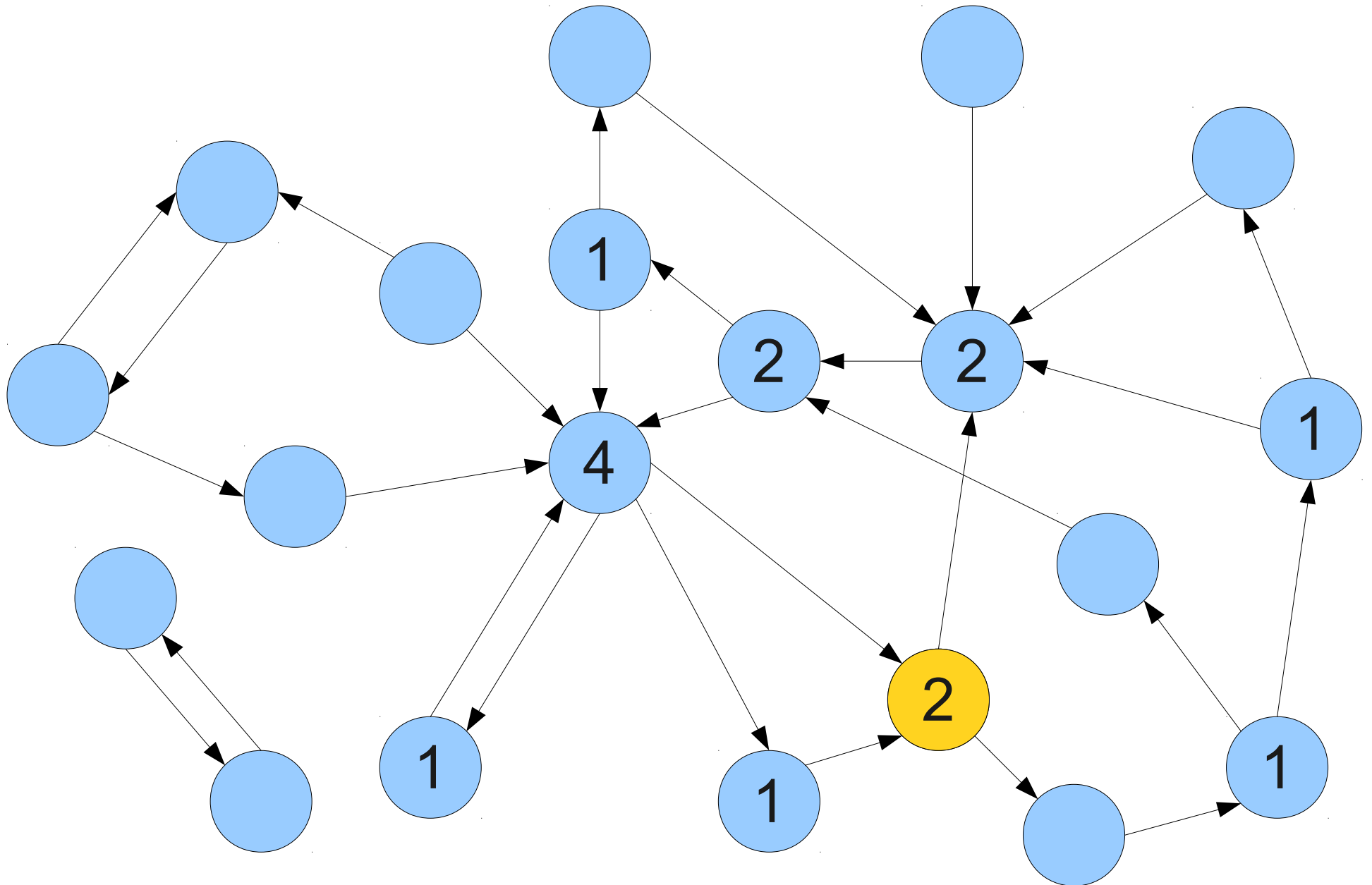




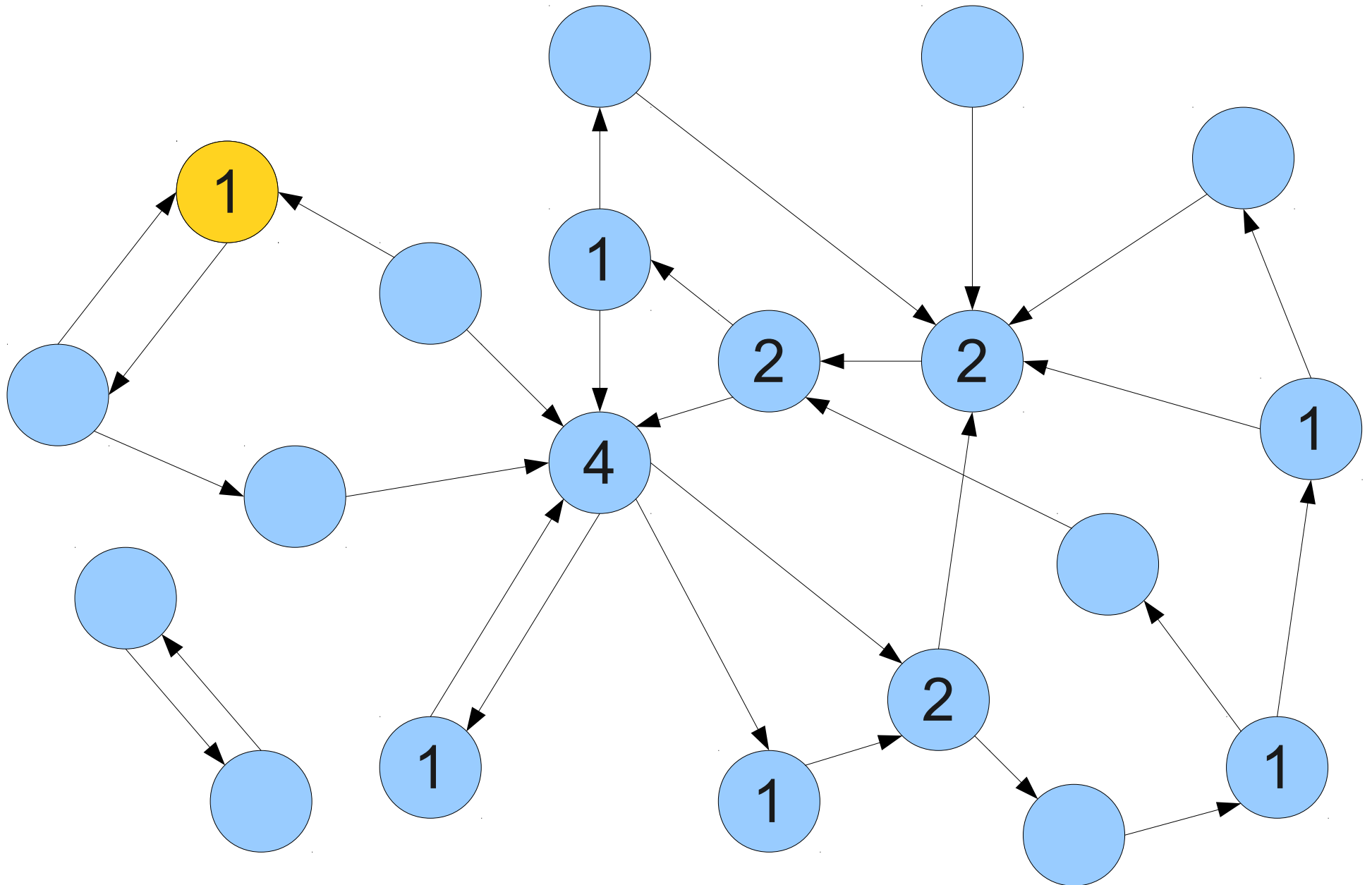
# The Random Surfer Model



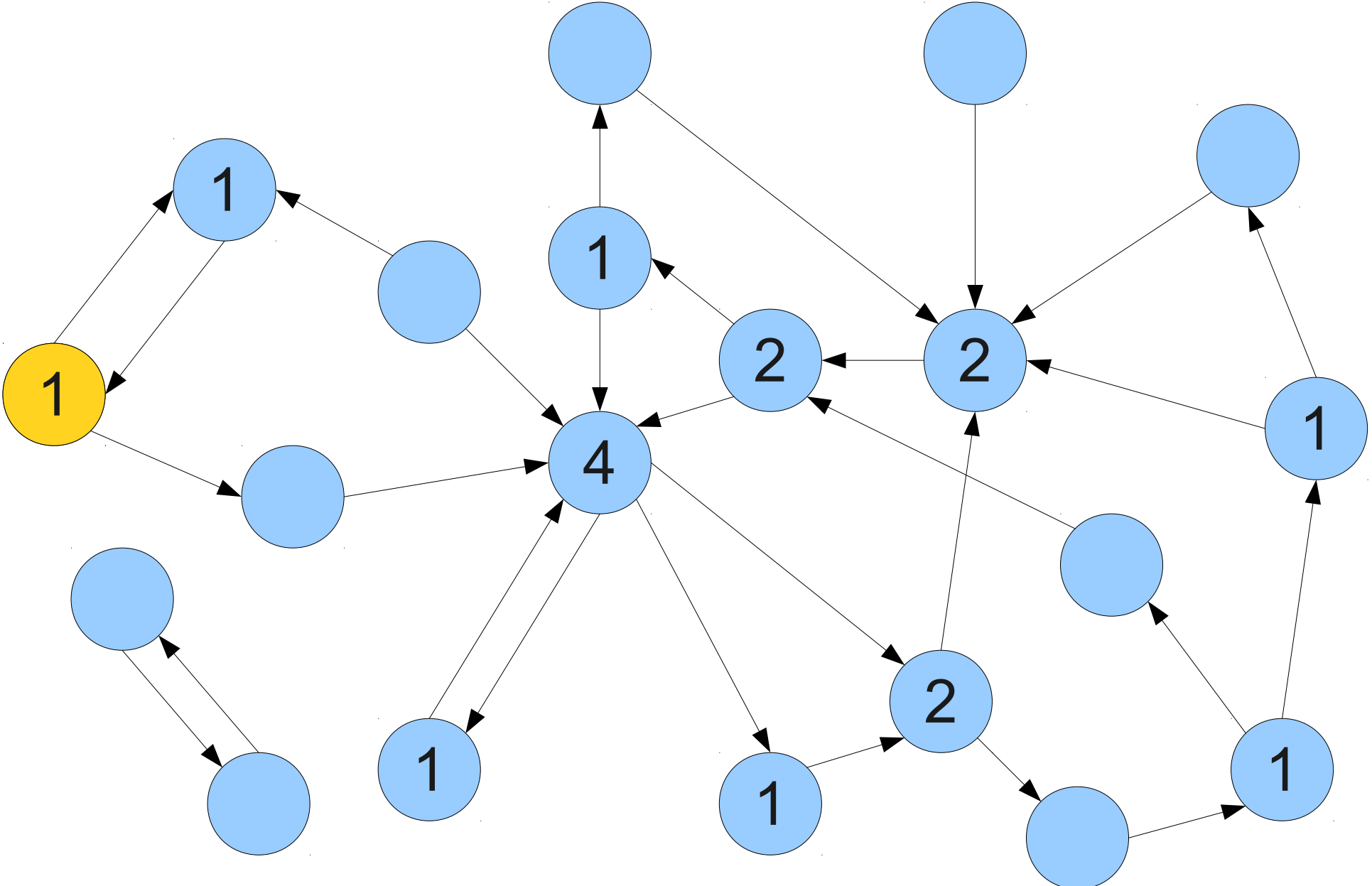
# The Random Surfer Model



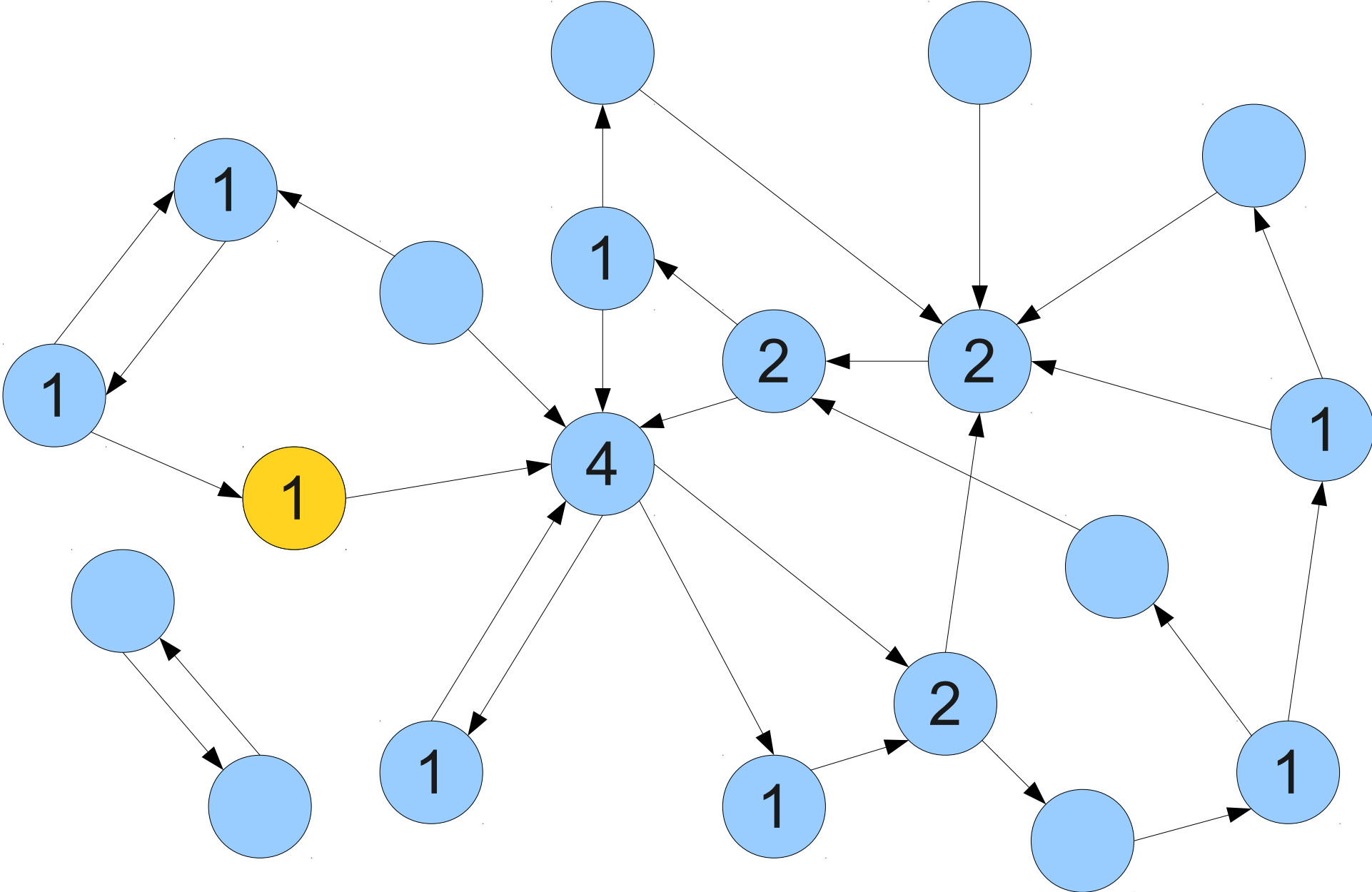
# The Random Surfer Model



# The Random Surfer Model

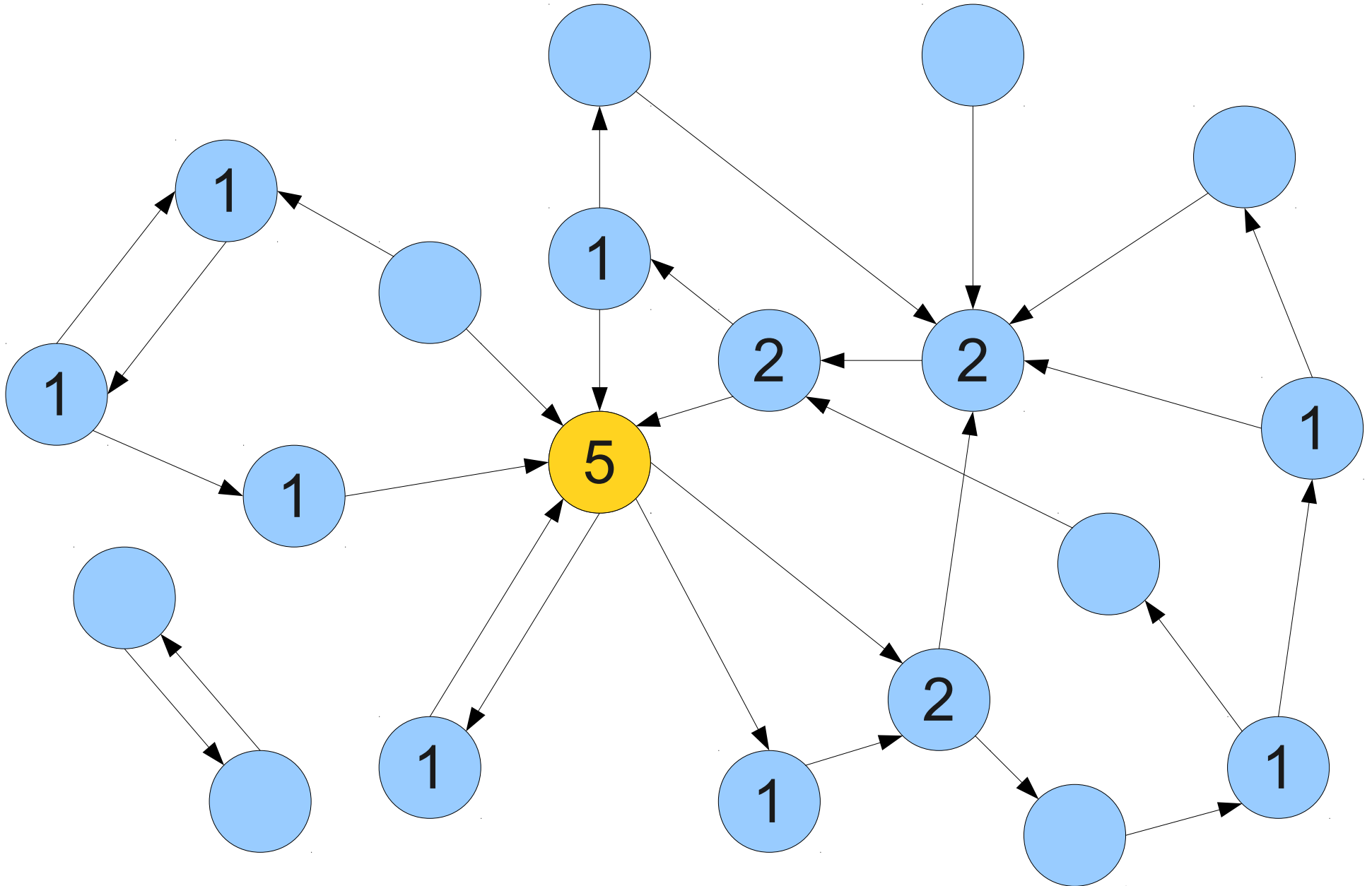


# The Random Surfer Model

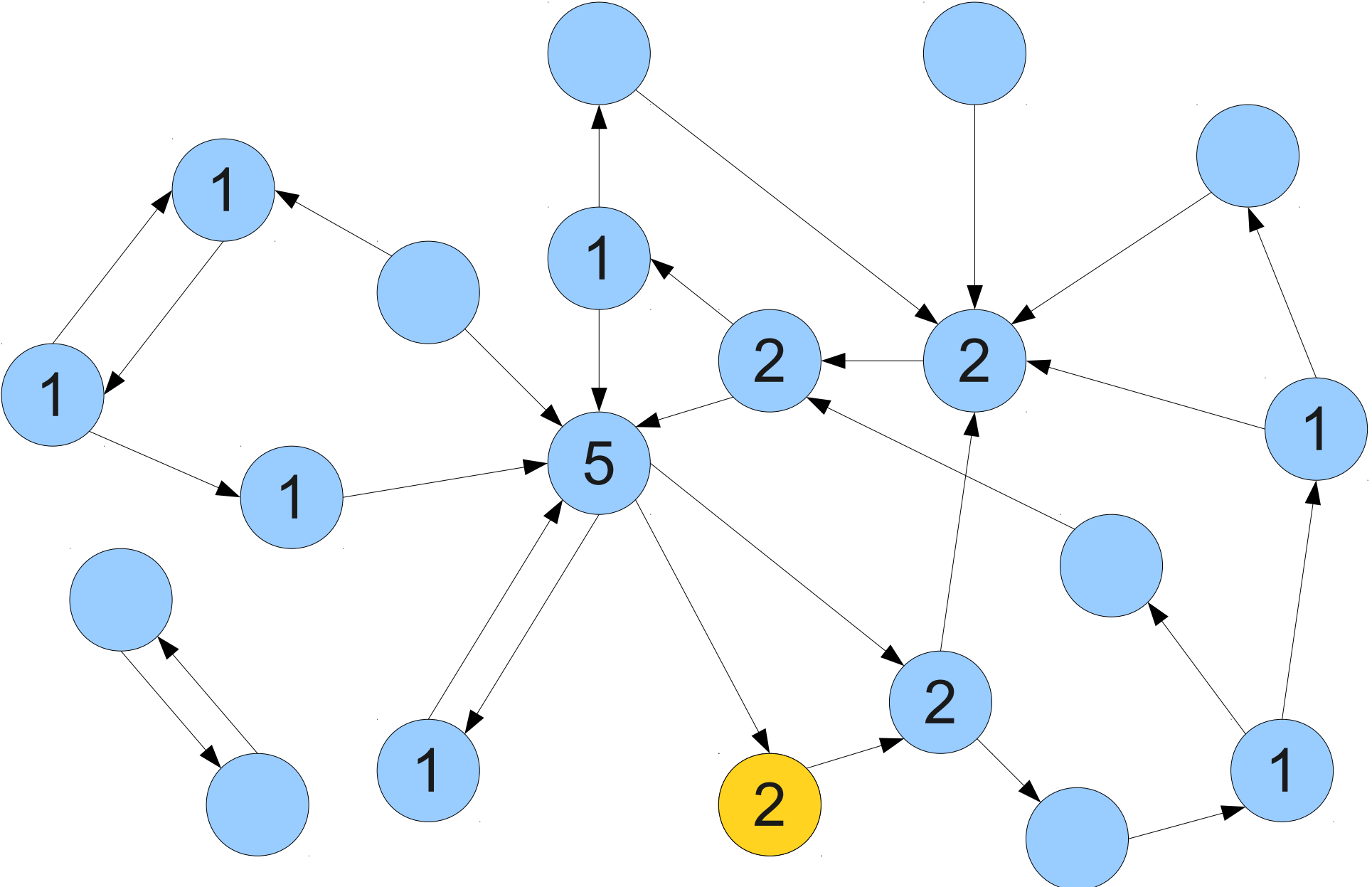




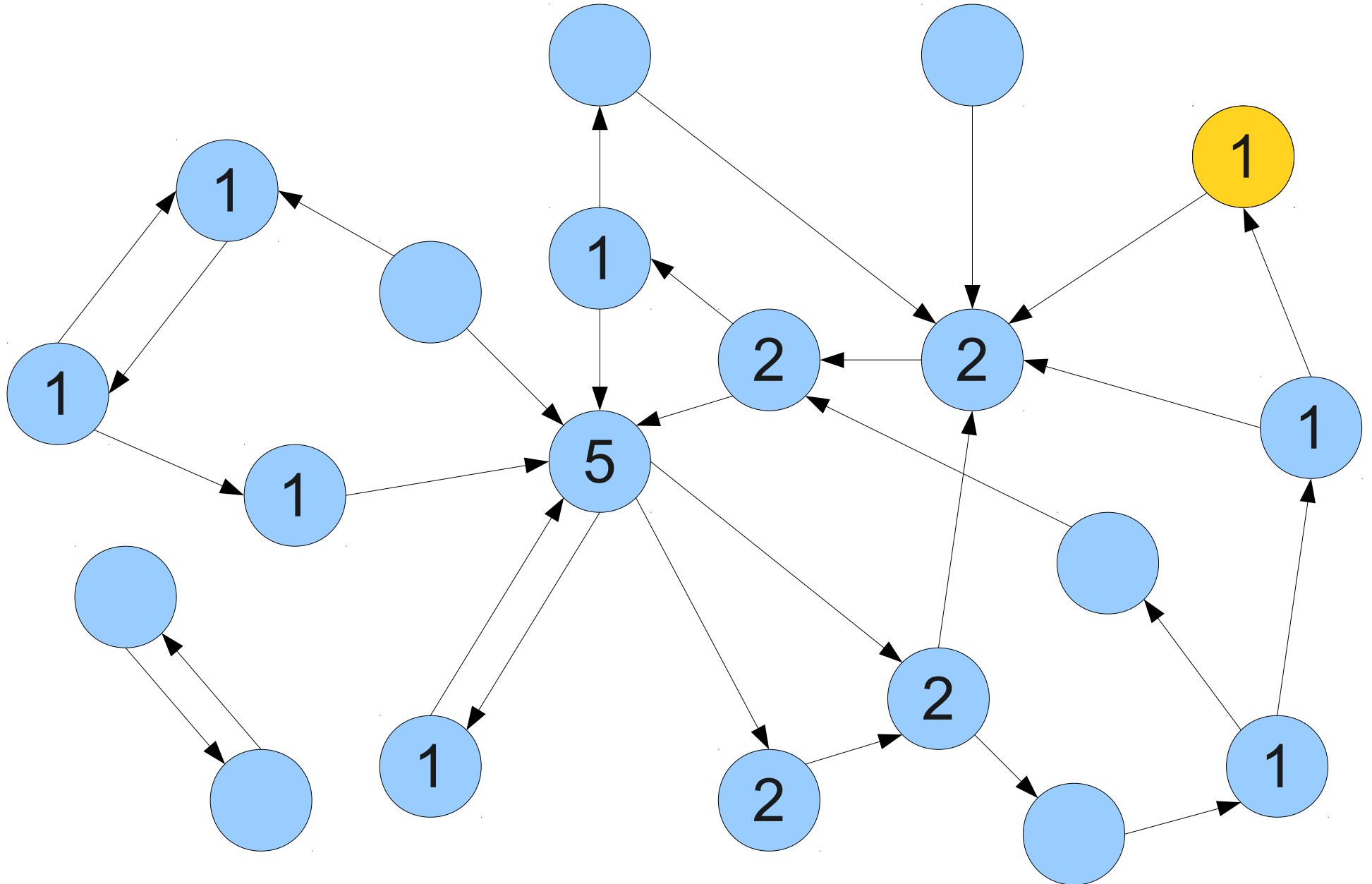
# The Random Surfer Model



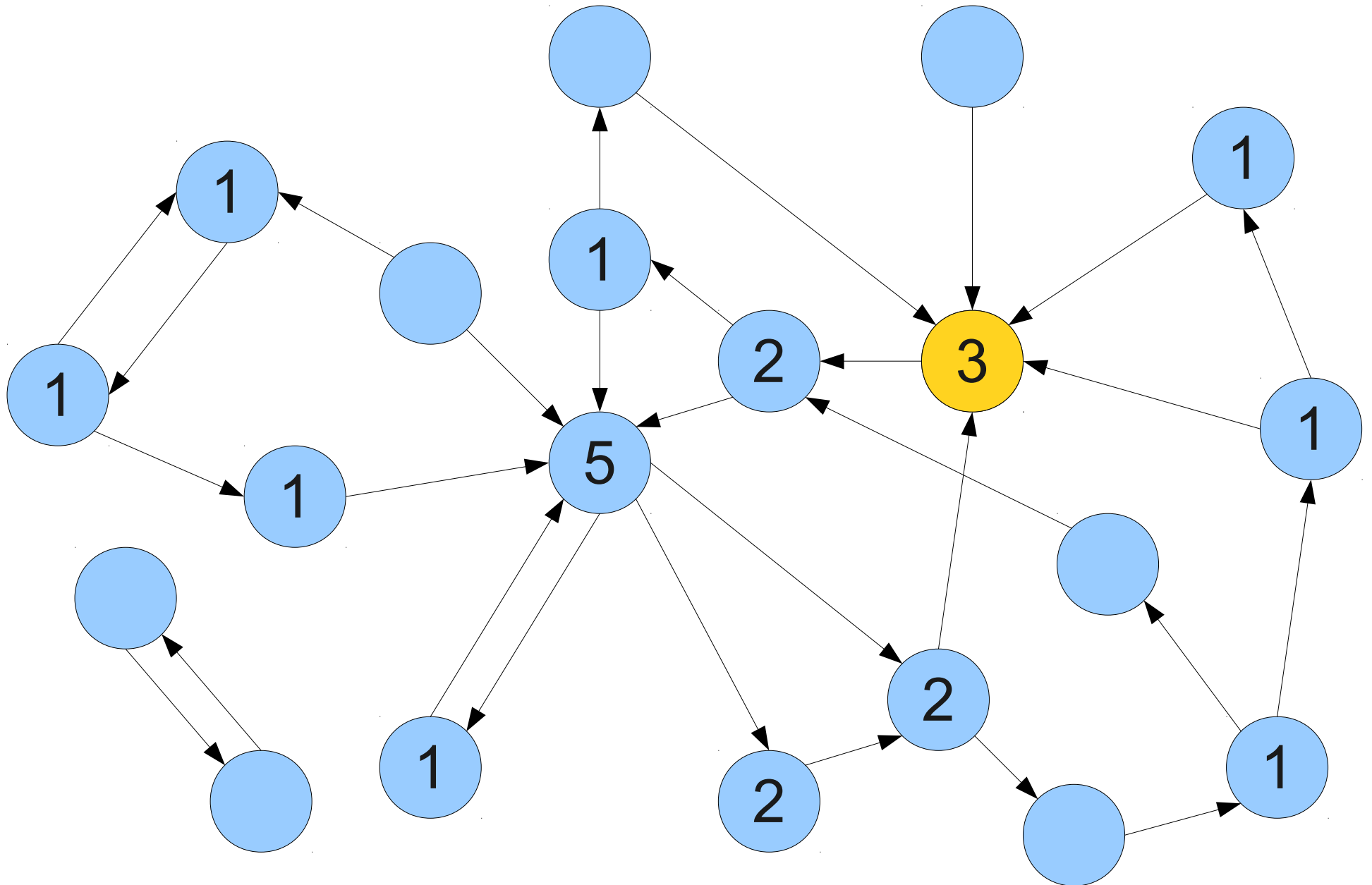
# The Random Surfer Model



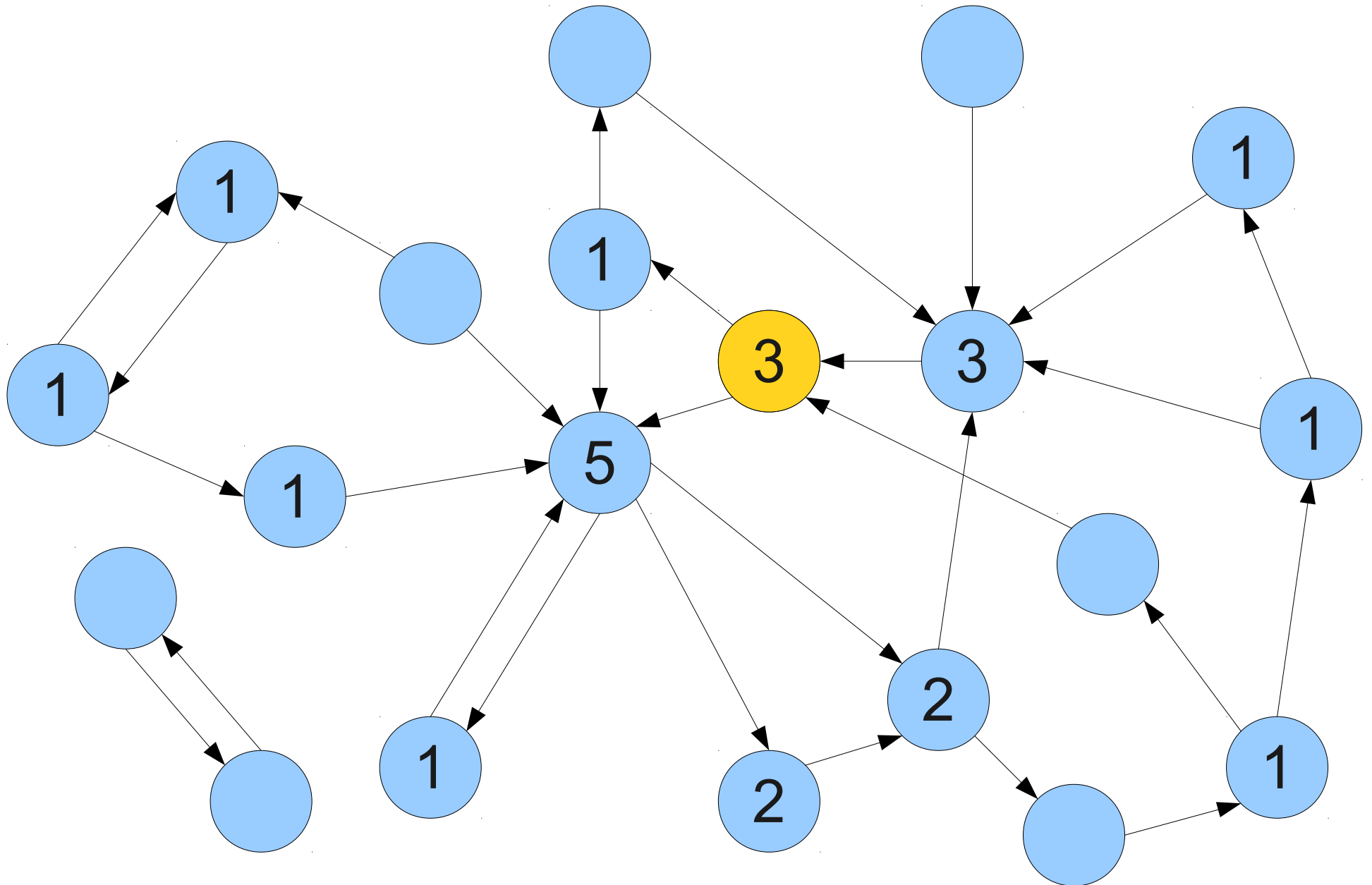
# The Random Surfer Model



# The Random Surfer Model

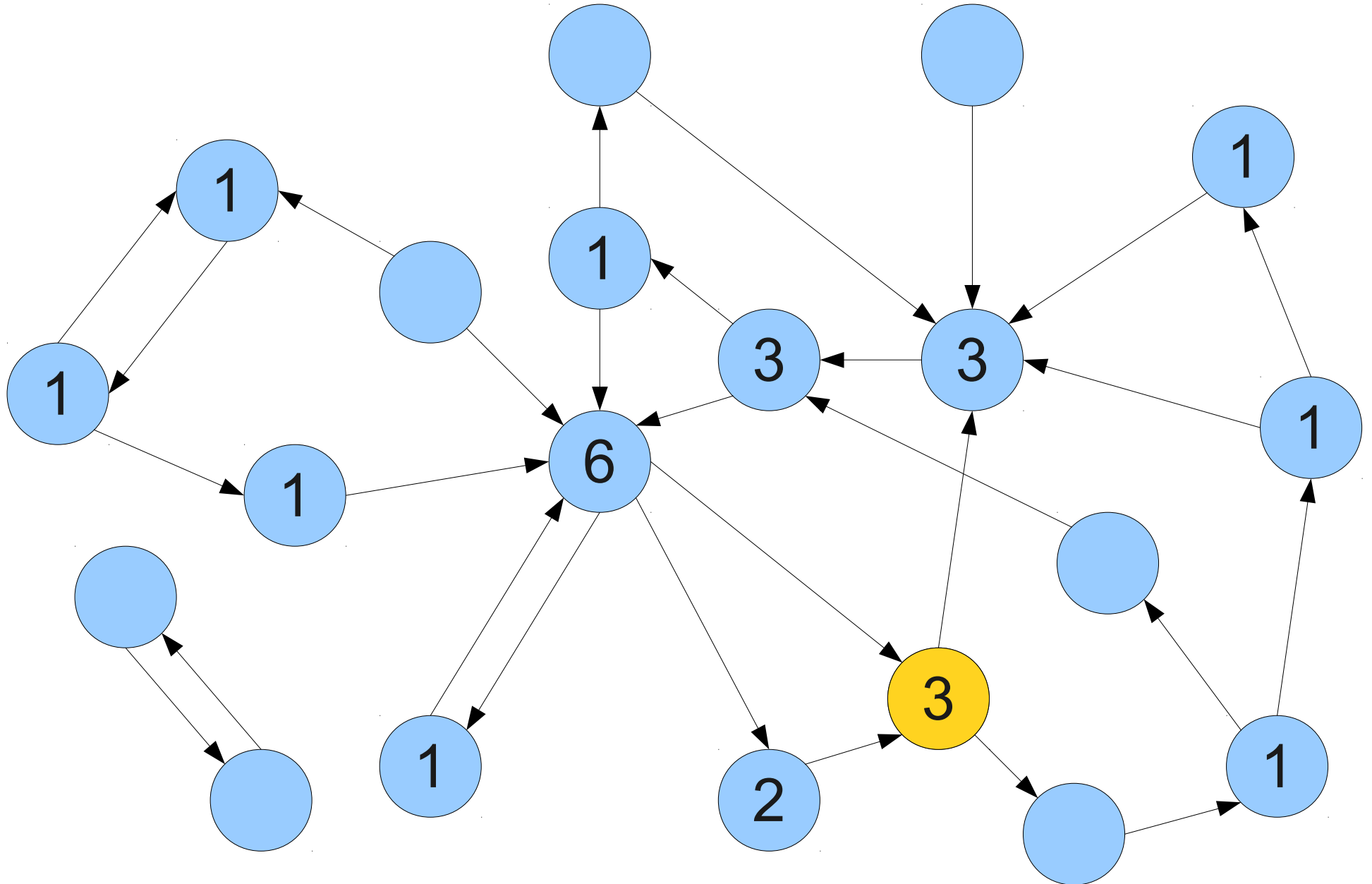


# The Random Surfer Model

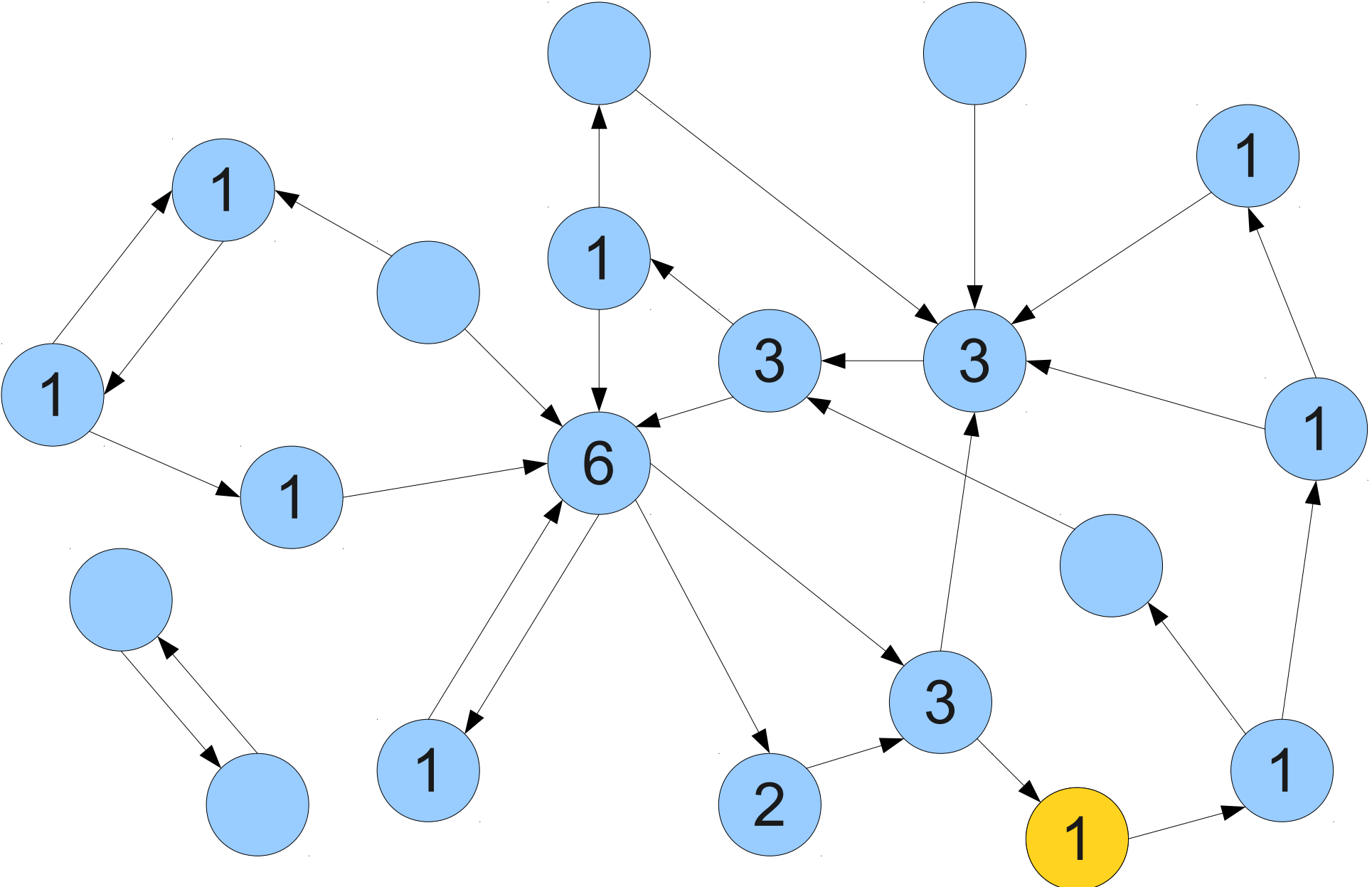




# The Random Surfer Model

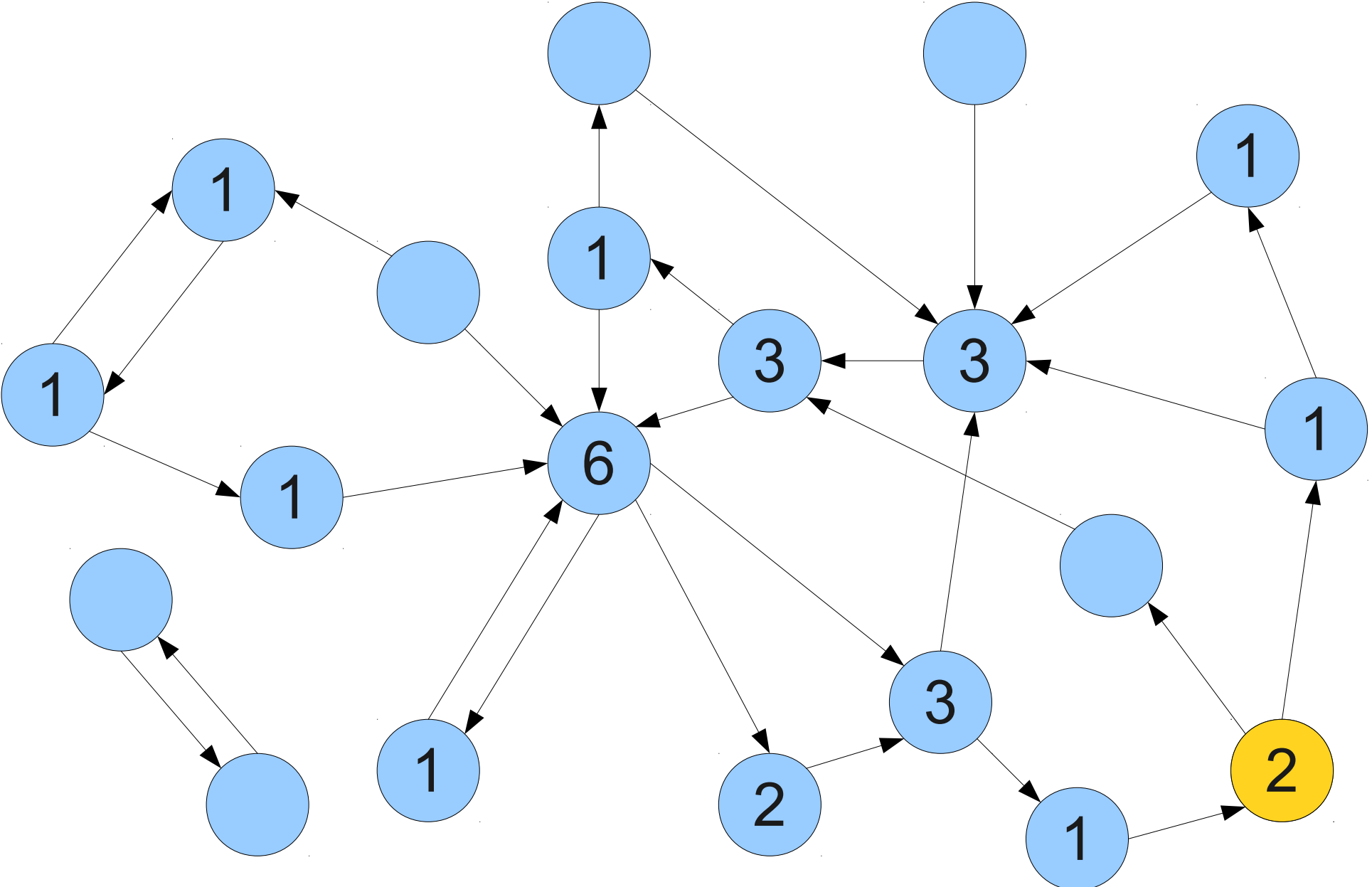


# The Random Surfer Model

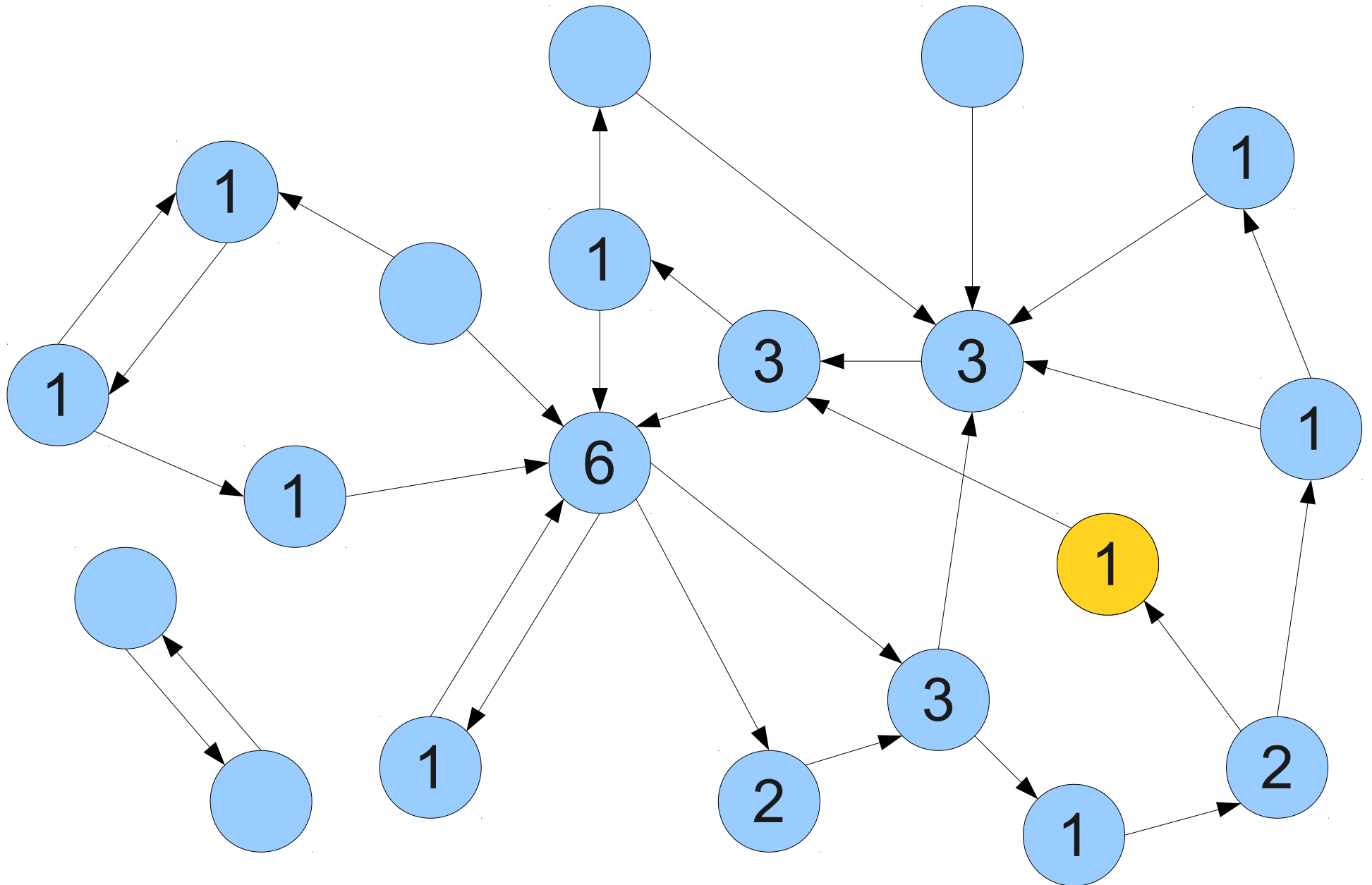




# The Random Surfer Model

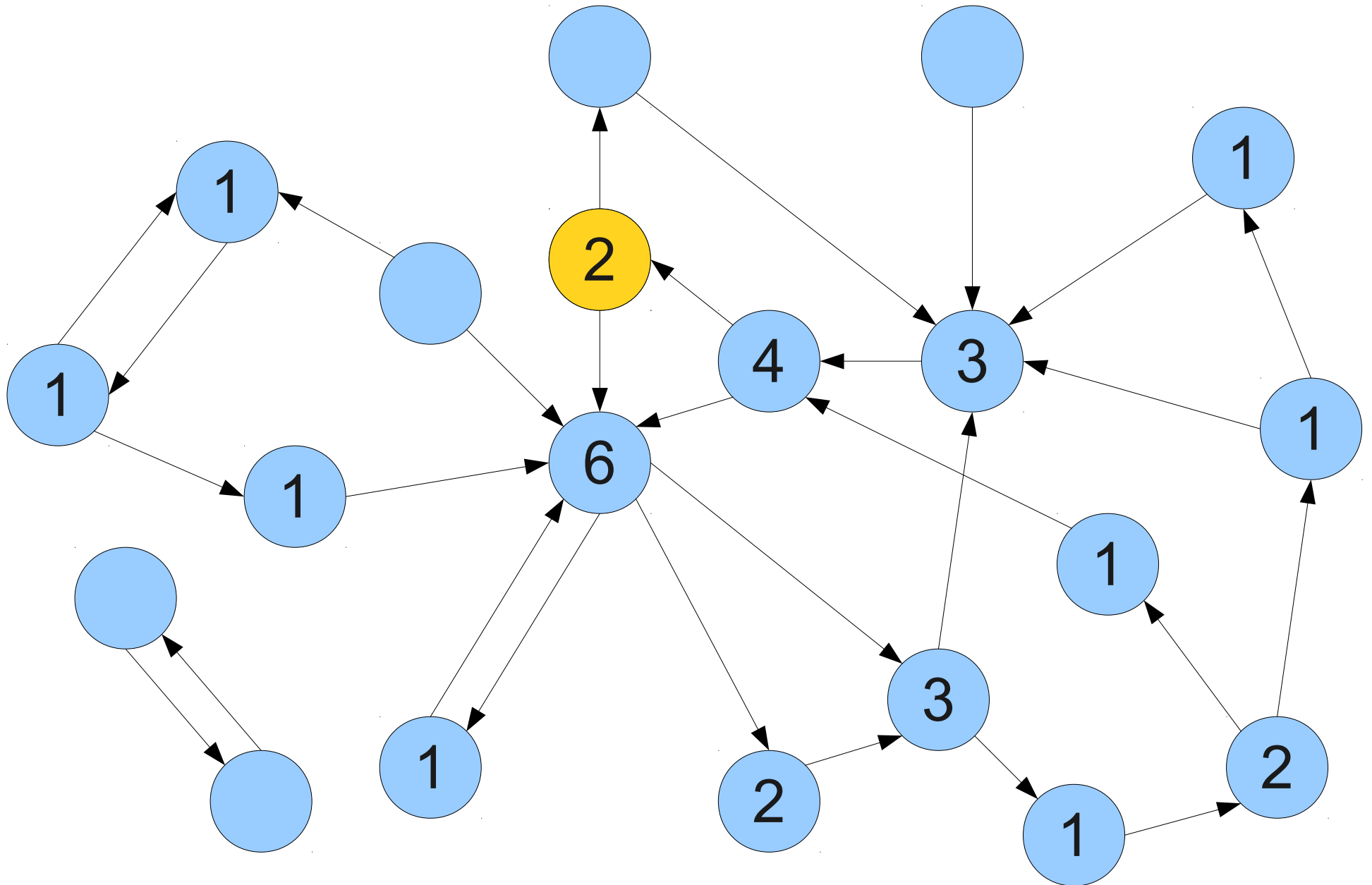


# The Random Surfer Model



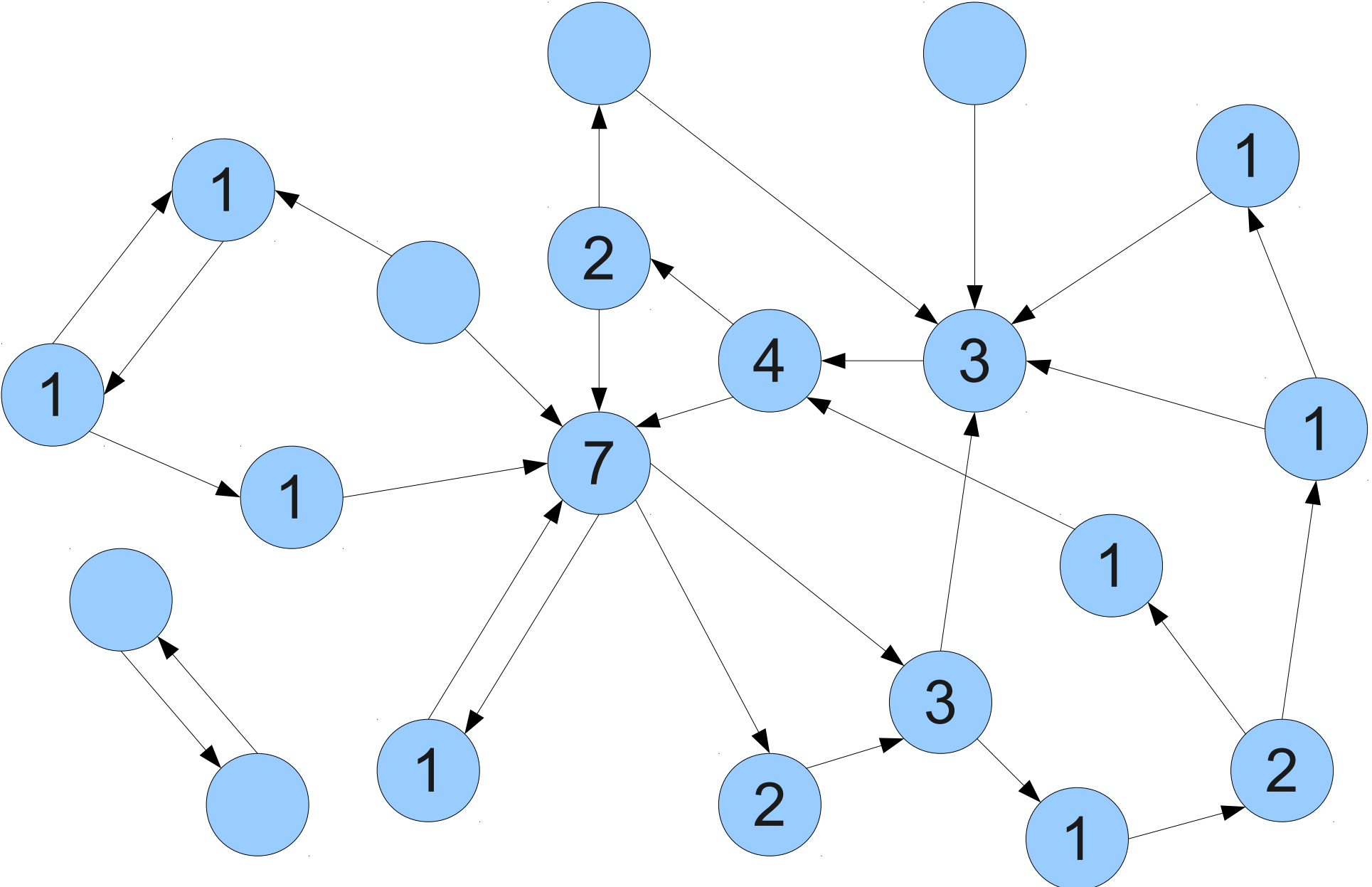


# The Random Surfer Model

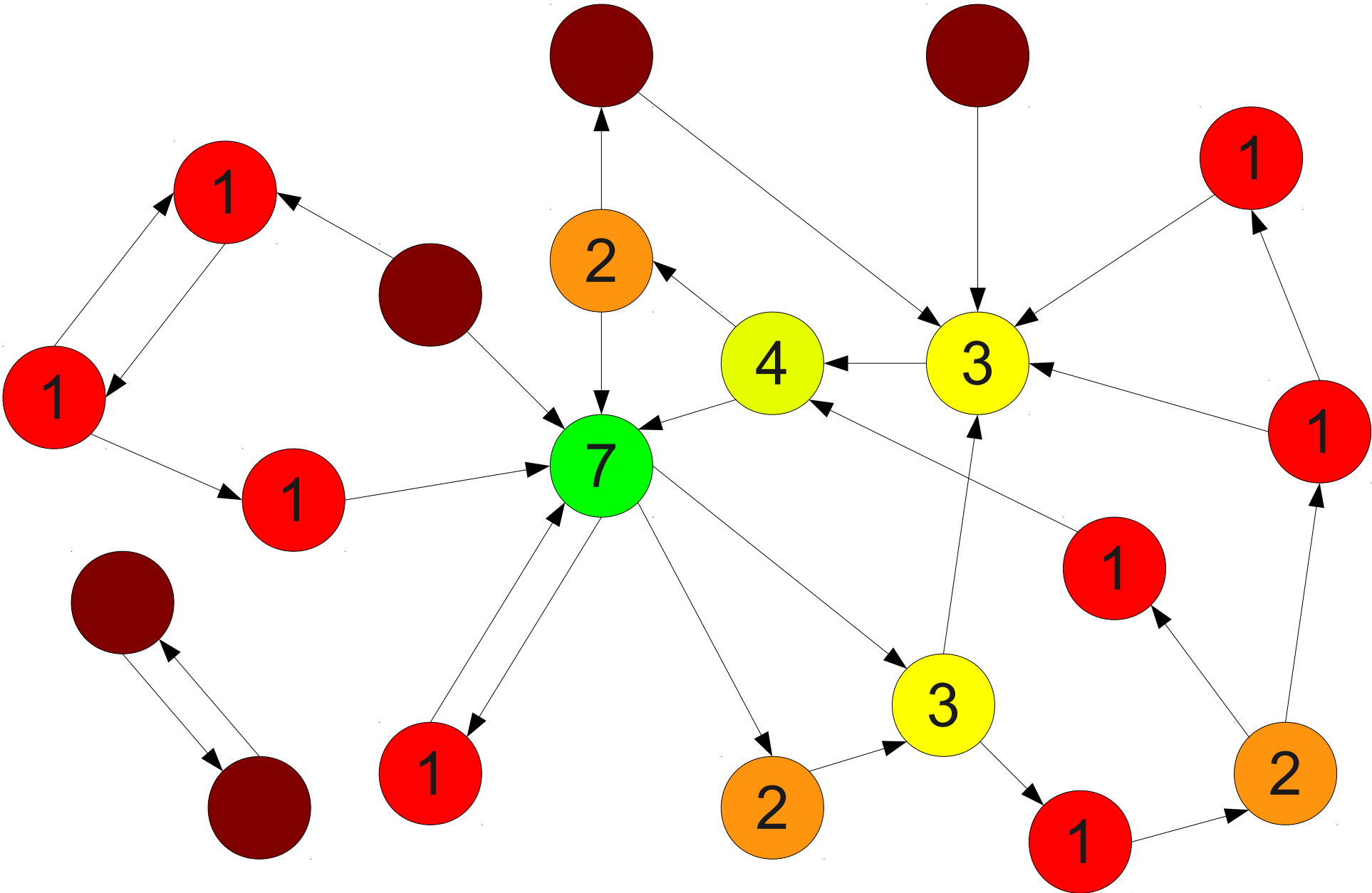




# The Random Surfer Model



# The Random Surfer Model



# Ranking Articles with the RSM

- Randomly walk through the graph.
- At each step, either
  - Jump to a totally random article, or
  - Follow a random link.
- Record how many times each article was visited.
- The most-visited articles are, in some sense, the most important.



# Other Applications of the RSM

- Ecosystem Stability:
  - Each node represents a species.
  - Edges represent one species that eats another.
  - High-value species are those that are important to the stability of the ecosystem.
- Learn more:
  - <http://news.bbc.co.uk/2/hi/8238462.stm>

Who invented this?

[Our approach] can be thought of as a model of user behavior. We assume there is a "random surfer" who is given a web page at random and keeps clicking on links, never hitting "back" but eventually gets bored and starts on another random page.

[Our approach] can be thought of as a model of user behavior. We assume there is a "random surfer" who is given a web page at random and keeps clicking on links, never hitting "back" but eventually gets bored and starts on another random page. The probability that the random surfer visits a page is its **PageRank**.

# **The Anatomy of a Large-Scale Hypertextual Web Search Engine**

Sergey Brin and Lawrence Page  
Computer Science Department,  
Stanford University, Stanford, CA 94305, USA  
sergey@cs.stanford.edu and page@cs.stanford.edu



Great things are possible in computing.

Great things are possible in computing.

You just need to do a little random surfing.