Rumor Has It

Project Module Associated with
***2nd Edition, Introduction to Computational Science: Modeling and Simulation* by**

Angela B. Shiflet and George W. Shiflet
Wofford College

Projects 1-4 can be developed with a system dynamics, agent-based, or cellular automaton model. Prerequisite for system dynamics models: Module 4.3, "Modeling the Spread of SARS— Containing Emerging Disease;" prerequisite for agent-based simulations: Module 11.2, "Agents of Interaction: Steering a Dangerous Course;" prerequisite for cellular automaton simulations: one of Module 10.3 on "Spreading of Fire," Module 10.4 on "Movement of Ants—Taking the Right Steps," or Module 10.5 on "Biofilms: United They Stand, Divided They Colonize"

Prerequisite for Project 5: Module 13.5, "The Next Flu Pandemic: Old Enemy—New Identity." Prerequisite for Project 5 parts requesting empirical modeling: Module 8.3, "Empirical Models"

Introduction

Your friend, Cathy, emails you that she just read a blog that said that using deodorant increases your risk of developing breast cancer. You are skeptical, but you then picked up a copy of *Time* magazine in your dentist office, and turned to the health section (Heid 2015). The headline read "You Asked: Can Deodorant Give You Cancer?" Now, you feel concerned. How can your deodorant give you cancer?

Cathy may have meant well, but she has alarmed her friend without verifying this information. We might term this information to be a "rumor," a word derived from the Latin, meaning clamor, hearsay, popular opinion (Online Etymology Dictionary). Today, the Oxford Dictionary defines the word to mean "a currently circulating story or report of uncertain or doubtful truth" (Oxford Dictionary 2015). If the rumor had been unverified report contained information about other people, we might term it *gossip*. Rumors pervade the internet and social media, and it is interesting that they can spread so quickly and live so long, substantiated or not.

The American Cancer Society summarizes the *inaccurate claims* made by an email rumor that has been circulated since 1999 (American Cancer Society 2014):

- Cancer-causing substances in antiperspirants are absorbed through razor nicks from underarm shaving. These substances are said to be deposited in the lymph nodes under the arm, which are not able to get rid of them by sweating because the antiperspirant keeps you from perspiring. This causes a high concentration of toxins, which leads to cells mutating into cancer.

- Most breast cancers develop in the upper outer quadrant of the breast because that area is closest to the lymph nodes exposed to antiperspirants. (Think of the breast as a circle divided by vertical and horizontal lines that cross at the nipple. Each of the 4 sectors you divide the breast into is called a quadrant. The upper outer quadrant of each breast is the part closest to the arm pit.)
- Men have a lower risk of breast cancer because they do not shave their underarms, and their underarm hair keeps chemicals in antiperspirants from being absorbed.

You can find the essence of this email at Snopes.com, which is an established website dedicated to validating or discrediting internet rumors, and the like (Snopes 2015). This rumor has little or no support from the scientific community, and some of the rumor indicates that the author had an incomplete understanding of biology. For instance, antiperspirants act by inhibiting the secretion of sweat glands, which are *not* lymph glands. Lymph glands have multiple roles in the immune system, but they do not remove toxins through sweat. The 1999 email gave life to a rumor that still floats around the web, despite the debunking that was researched and is presented at Snopes. Cathy had apparently not done any verification of this rumor, before she passed it along to you.

Projects

1. Daley-Kendall (Thompson et al 2003) considered a model for the spreading of rumors that is similar to an SIR model. In the model, the entire population (number *N*) contains the following classes:

- *X* – ignorants, who have not heard the rumor
- *Y* – spreaders, who are spreading the rumor
- *Z* – stiflers, who know the rumor but who have stopped spreading the rumor

Contact between a spreader and an ignorant can result in the latter becoming a spreader. However, contact between a spreader and someone who is not ignorant (another spreader or a stifler) causes a shift to being stiflers because the rumor is old news and not worth spreading. A stifler stays a stifler, regardless of contacts. The model assumes a closed, well-mixed society, with no births or deaths.

a. Develop a system dynamics model using these criteria. Transfer from *X* to *Y* depends on the fraction of *Y* individuals. Transfer from *Y* to *Z* varies with the proportion of people in categories *Y* and *Z*.

b. Develop an agent-based simulation using these criteria.

c. Develop a cellular automaton simulation using these criteria.

2. Develop a system dynamics, agent-based, or cellular automaton model of the spread of a rumor using the same categories as in Project 1 and the following criteria: Contact between a spreader and an ignorant can result in the latter becoming a spreader. Due to the rumor becoming less and less interesting, an increasing proportion of spreaders become stiflers as time advances. A stifler stays a stifler, regardless of contacts. The model assumes a closed, well-mixed society, with no births or deaths.

3. (Thompson et al 2003) refined the Daley-Kendall model (Project 1) to include the following classes:

- S_p – passive susceptibles, who have not heard the rumor, have fewer contacts, and generally do not wish to spread a rumor
- S_a – active susceptibles, who have not heard the rumor, have more contacts, and generally want to spread a rumor
- G_p – passive gossipers, who have heard the rumor but are less likely to spread the rumor than active gossipers. Passive gossipers that spread the rumor do so at a smaller rate than active gossipers.
- G_a – active gossipers, who have heard the rumor and are more likely to spread the rumor than passive gossipers. Active gossipers that spread the rumor do so at a higher rate than passive gossipers.
- Z – stiflers, who know the rumor but who have stopped spreading the rumor

Upon hearing a rumor, an active susceptible will have a higher likelihood of becoming an active gossiper than becoming a passive gossiper. However, a passive gossiper could become a stifler as well as a passive or active gossiper. A gossiper spreads a rumor to a susceptible person, who changes categories as indicated above. Contact between two gossipers causes them both to become stiflers. For example, they may interpret the gossip as "old news." Similarly, contact between a gossiper and a stifler results in the gossiper becoming a stifler. A stifler stays a stifler, regardless of contacts. Assume a closed society with no births or deaths.

a. Develop a system dynamics, agent-based, or cellular automaton model using these criteria.

b. Adjust parameters to consider the results of a frivolous rumor, where active and passive susceptibles are likely to remain in their same activity group (active or passive, respectively). Discuss the results.

c. Adjust parameters to consider the results of an interesting rumor where both active and passive susceptibles are likely to spread the rumor at a higher level of activity. Discuss the results.

d. Adjust parameters to consider the results of a boring rumor, where most of the susceptibles spread the rumor at a lower level of activity. Discuss the results.

e. Adjust parameters to consider the results of an unbelievable rumor, in which active susceptibles are more likely to go to the passive gossiper class, and passive susceptibles are more likely to transfer to the stifler class. Discuss the results.

f. Compare the results of Parts b-e.

4. (Thompson et al 2003) presented a variation of the model in Project 3 that allowed births and deaths. In this case, birth/death means the addition/deletion of an account on an internet site, where information can be shared. Incorporate such births and deaths in completing all parts of Project 3.

5. Download *facebook.tar.gz* and *readme-Ego.txt* from http://snap.stanford.edu/data/egonets-Facebook.html. This project will be using the files that end in *.edges*, such as *107.edges*. Project 3 of Module 14.15, "Social Networks: Value in Being Well-Connected," describes the files.

(Doerr et al 2012) examined "Why Rumors Spread Fast in Social Networks." They concluded, "a good explanation for this phenomenon is that small-degree nodes quickly learn a rumor once one of their neighbors knows it, and then again quickly forward the news to all their neighbors. This phenomenon in particular facilitates sending a rumor from one large-degree node to another."

a. Develop a simulation for the spread of a rumor in a social network. Start the rumor with a random person (node). At each time step, each person contacts a random neighbor and exchanges information, so that any node with a rumor transmits the gossip to a random friend. In picking a random neighbor, exclude the friend contacted in the previous step, unless the node only has one friend. Run the simulation until a fixed percentage, say 50%, of the people acquires the rumor, and display the number of nodes (*n*), the number of edges (*numEdges*), and the **density** of the network (*density*, the number of edges divided by the maximum number of possible edges, $n(n - 1)/2$) and the number of time steps for the simulation (*numSteps*).

Some social networks are connected, so that it is impossible to spread the rumor to every person. Moreover, the random initial person could be a member of a small connected subgraph, so that the rumor cannot even spread to 50% of the people. Thus, abort one the run, select a different random initial person, and run the simulation again should the state of the network not change (no one new hears the rumor) in a fixed number of time steps, say 10, and 50% of the people are not yet spreaders.

For ease of writing later a procedure to perform the simulation multiple times, break the simulation into two parts, *startRumor* and *endRumor*. The procedure *startRumor* has an edges matrix as a parameter and returns *n*, *numEdges*, *density*, and two other matrices, *deg* and *g*. The vector *deg* contains the degrees of the nodes. For example, suppose an edges file for Figure 1 is as follows:

```
37 28
28 19
19 28
28 37
```

The IDs for the vertices are not sequential from 1 but can appear in the following order:

Index	**Vertex**
1	19
2	28
3	37

Using this ordering, the *deg* vector with indicated indexing would appear as follows, with the second node, 28, having degree 2:

Index	**Degree (*deg*)**

1	1
2	2
3	1

We could have a connection matrix as follows:

	1	2	3
1	0	1	0
2	1	0	1
3	0	1	0

However, for a large number of vertices, the connection matrix would be very large, usually with many zeros. An alternative representation would be to store the indices of the adjacent nodes in a matrix, *g*, as follows with 0 as a place holder:

Index	**Indices of Adjacent Nodes (*g*)**
1	2 0
2	1 3
3	2 0

Node 19, which has index 1, is adjacent to Node 28, with index 2. The node with index 2 (Node 28) is adjacent to the nodes with indices 1 (Node 19) and 3 (Node 37). Such an arrangement would also make it easy for us to keep track of a neighbor just picked by swapping that index to be at the end of the list. For example, if the node with index 2 selected index 1 to exchange information, at the next time step *g* would be as follows, with 1 at the end of the second row:

Index	**Indices of Adjacent Nodes (*g*)**
1	2 0
2	**3 1**
3	2 0

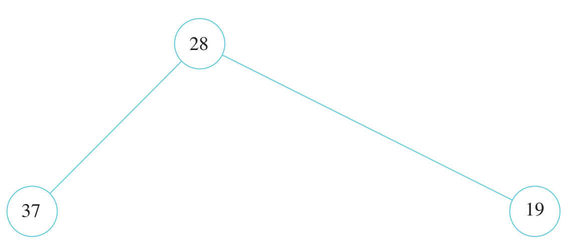

Figure 1 Example network

b. Develop a function with input parameters of an edges matrix and the number of simulations to run and returns the number of edges, density, and average number of simulation time steps.

c. As well as the edges in the files, develop a procedure to include edges with the ego as an endpoint. For example, *0.edges* contains "236 186." Thus, vertex 0 is adjacent to vertices 236 and 186, so include (0, 236), (236, 0), (0, 186), and (186, 0). By including edges to the ego, we ensure that the graph is connected.

d. Run the simulation at least 10 times for each of the different augmented *.edges* Facebook files from Part c. Plot the number of time steps to spread the rumor versus the number of nodes and the number of time steps versus the number of edges. Using empirical modeling, discover a function that captures the trend of the data.

e. Develop a procedure that for a given number of nodes generates an edges matrix for a complete graph, or a graph in which each node is connected to every other node. The matrices should have the same format as those for Facebook.

f. Repeat Part d for complete graphs with the same number of vertices as the *.edges* Facebook files from Part c. Because the graphs are complete, we only need to run the simulation once for each graph. Compare the results with those of Part d.

References

American Cancer Society. 2014. "Antiperspirants and Breast Cancer Risk." http://www.cancer.org/cancer/cancercauses/othercarcinogens/athome/antiperspirants-and-breast-cancer-risk (accessed November 19, 2015)

Doerr, Benjamin, Mahmoud Fouz, and Tobias Friedrich. 2012. "Why Rumors Spread Fast in Social Networks," *Communications of the ACM* 55.6 (2012): 70-75.

Heid, Markham. 2015. "You Asked: Can Deodorant Give You Cancer?" *Time* June 17, 2015

Online Etymology Dictionary. 2015. "Rumor." http://www.etymonline.com/index.php?term=rumor (accessed November 5, 2015)

Oxford Dictionary. 2015. "Rumour." http://www.oxforddictionaries.com/definition/english/rumour (accessed November 5, 2015)

Snopes.com. 2015. "Breast Defense." http://www.snopes.com/medical/toxins/antiperspirant.asp (accessed November 6, 2015)

Thompson, K., R. Castro Estradat, D. Daugherty, and A. Cintr6n-Arias. 2003. "A Deterministic Approach to the Spread of Rumors." http://www.researchgate.net/profile/Ariel_Cintron-Arias/publication/228925590_A_deterministic_approach_to_the_spread_of_rumors/links/0c96052c19a8b4f872000000.pdf. (accessed Oct. 29, 2015)