Social search

Matthew J. Salganik

Social Network (Soc 204)
Spring 2017
Princeton University

February 27, 2017

Logistics:

- Blackboard update
- Remember to save your answer and log out
- Story behind the essay this week

Questions?

Fides: A Responsible Data Science Platform

Julia Stoyanovich

Tuesday, 12:30 - 1:30, Sherrerd Hall 306

Recent attention in FAT research has been focused primarily on analyzing learning algorithms and their outputs. Yet, issues of fairness, accountability and transparency begin further upstream in the data science lifecycle: bias in source data goes unnoticed, spurious correlations lead to reproducibility problems, and pre-processing steps strongly influence analysis results. As machine learning methods continue to be applied broadly by non-experts, the potential for misuse increases. In this talk reasons to advocate for systems support for responsible data science will be presented. . . .

`https://citp.princeton.edu/event/stoyanovich/`

Vote:

1. Watts, Chapter 5.
2. Lee, N.H. (1969). The Search for an Abortionist: Preface, Chapter 1, and Chapter 5.

- sometimes the edges that don't exist are as important as the edges that do exist

- sometimes the edges that don't exist are as important as the edges that do exist
- affiliation networks (people and groups) help us understand patterns in personal network structure

- sometimes the edges that don't exist are as important as the edges that do exist
- affiliation networks (people and groups) help us understand patterns in personal network structure
- compare and contrast psychological vs sociological explanations for network structure

- sometimes the edges that don't exist are as important as the edges that do exist
- affiliation networks (people and groups) help us understand patterns in personal network structure
- compare and contrast psychological vs sociological explanations for network structure
- sociological principles can shape the design of technical systems

Travers and Milgram showed two things

1. short paths exist
2. people can find them using only local information

Travers and Milgram showed two things

1. short paths exist
2. people can find them using only local information

Watts distinguishes between

- Broadcast search (what the computer did your wikipedia assignment)
- Directed search (what you did in your wikipedia assignment)

Travers and Milgram showed two things

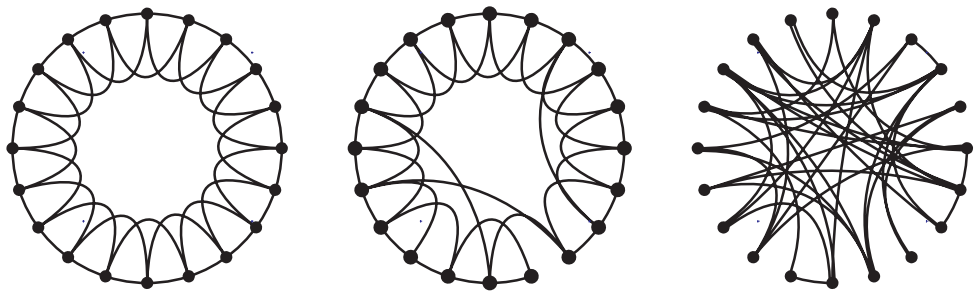
1. short paths exist
2. people can find them using only local information

Watts distinguishes between

- Broadcast search (what the computer did your wikipedia assignment)
- Directed search (what you did in your wikipedia assignment)

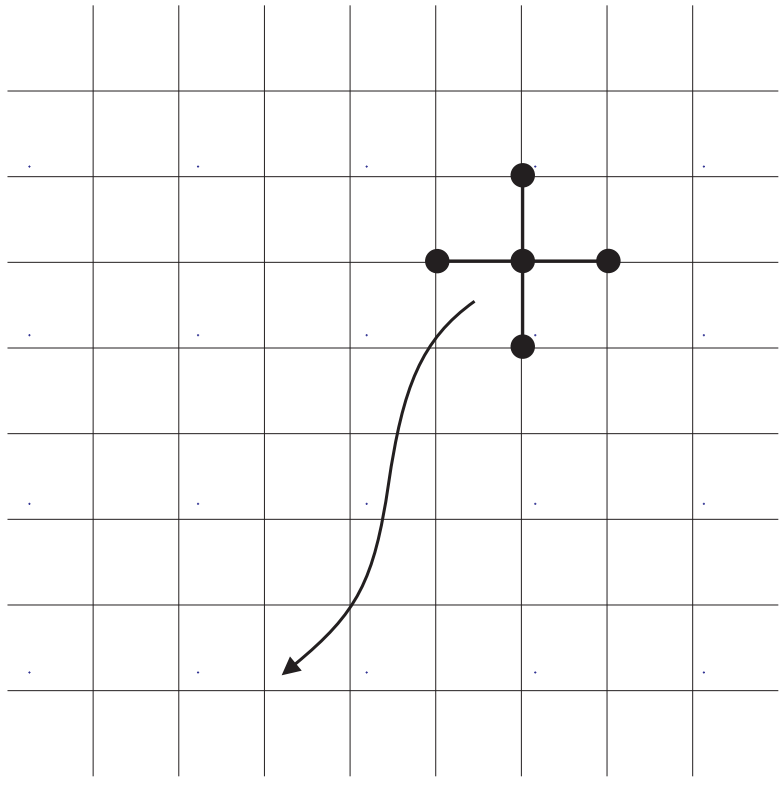
How is it that directed search ever works?

3.6

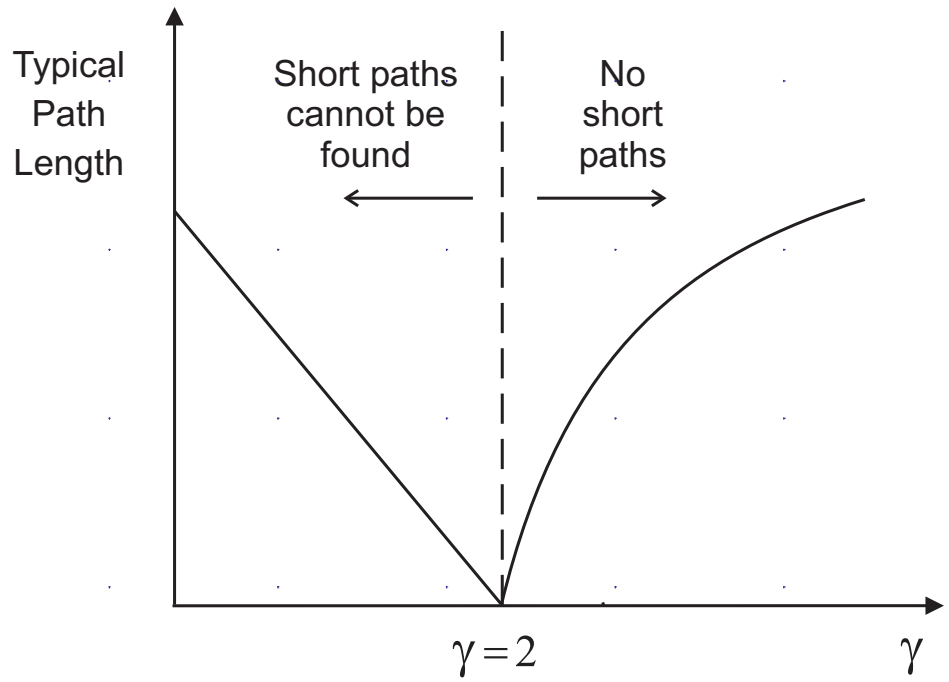


Increasing randomness

5.1



5.3



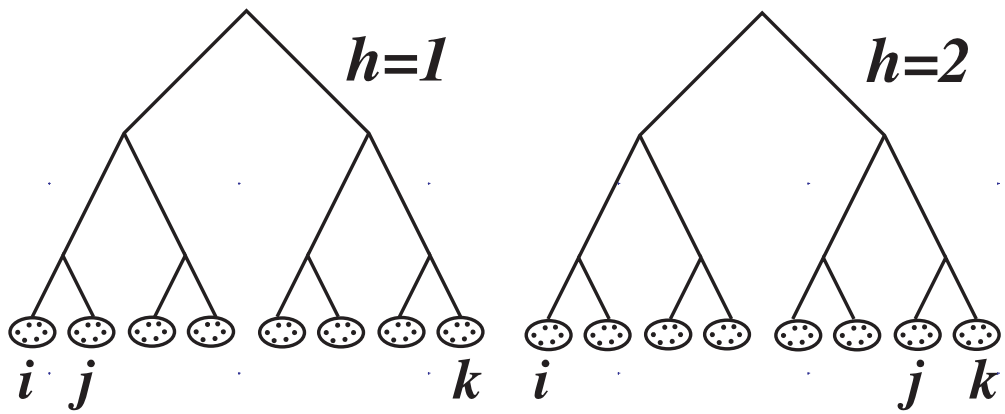
higher γ fewer long distance edges

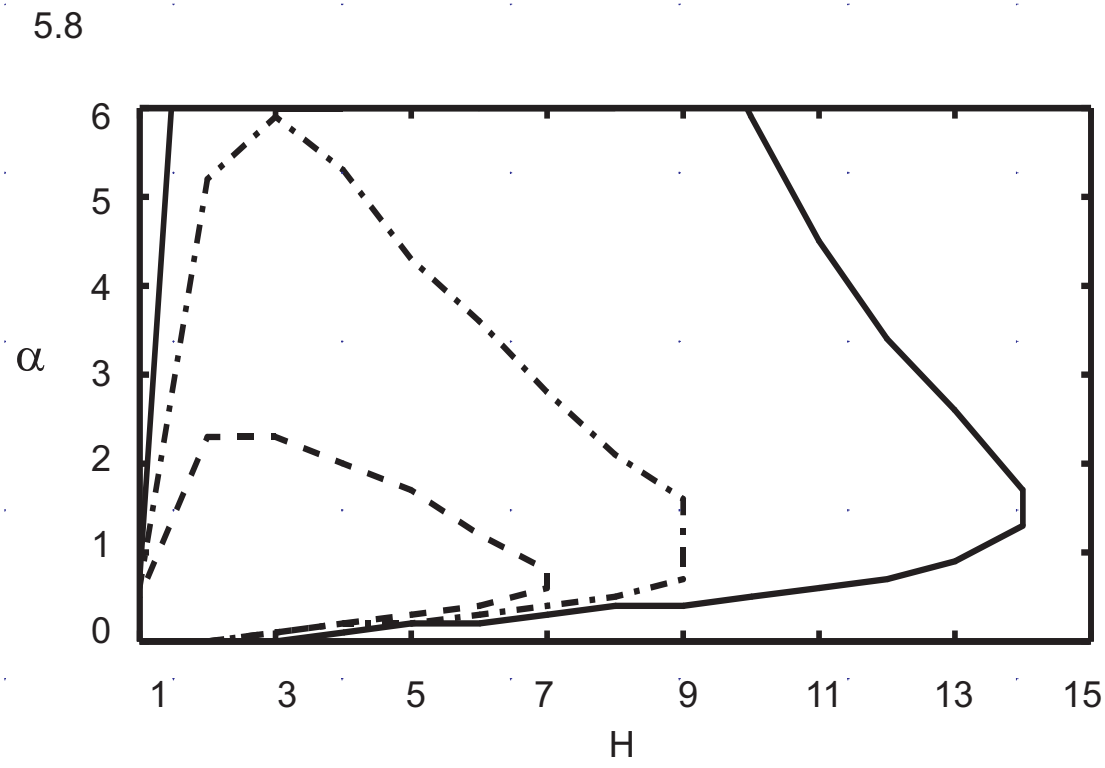
Navigation in a small world

It is easier to find short chains between points in some networks than others.

http://dx.doi.org/10.1038/35022643

5.7



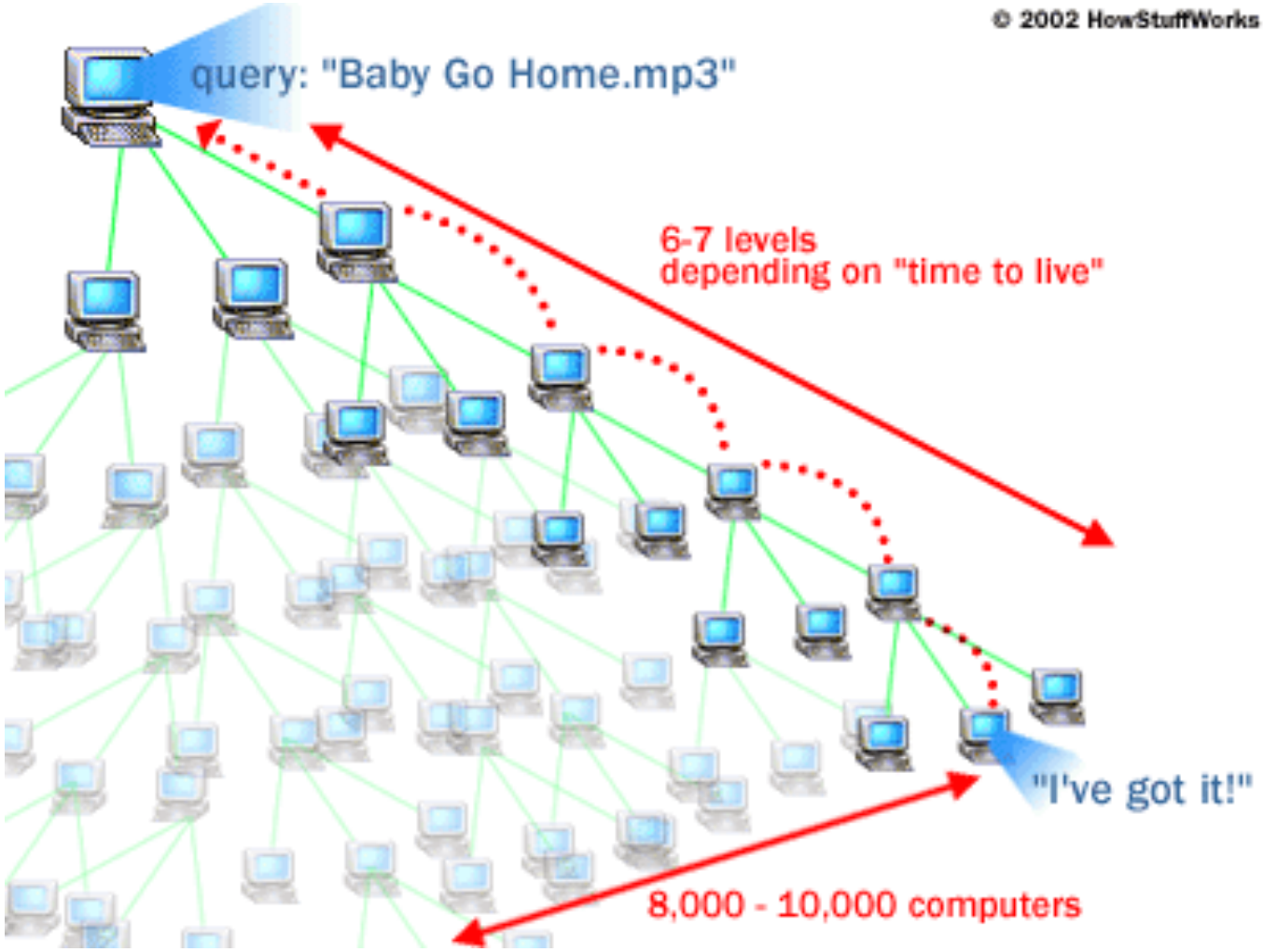


Details of this model are not as important as the other models we have learned about

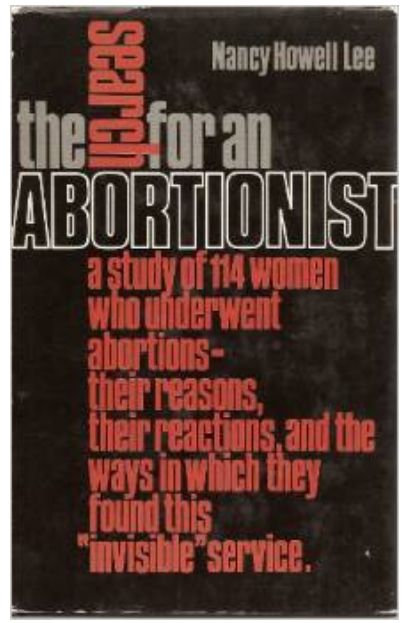
Identity and Search in Social Networks

Duncan J. Watts,[1,2,3]* Peter Sheridan Dodds,[2] M. E. J. Newman[3]

http://dx.doi.org/10.1126/science.1070120



http://computer.howstuffworks.com/file-sharing3.htm


Nancy Howell Lee
the search for an
ABORTIONIST
a study of 114 women who underwent abortions - their reasons, their reactions, and the ways in which they found this "invisible" service.

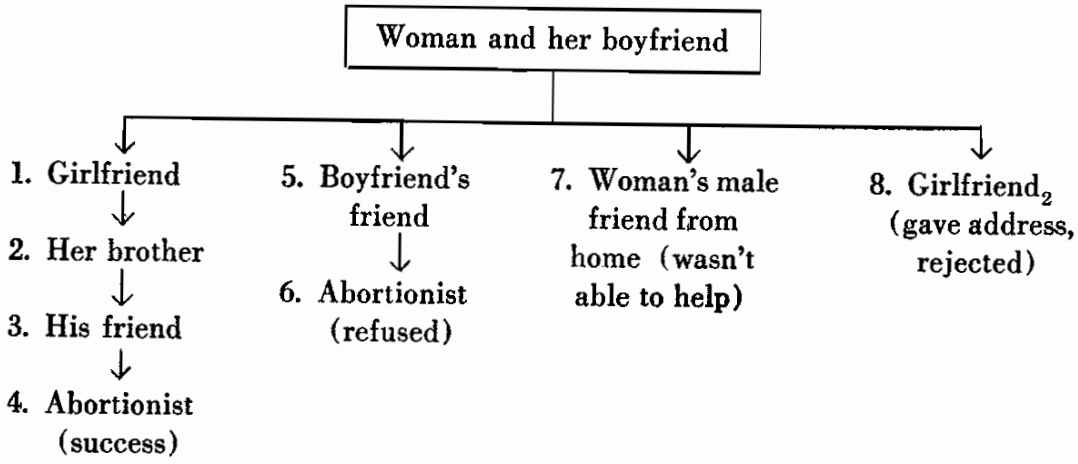


FIG. 2. Diagram of a search process—Example 1

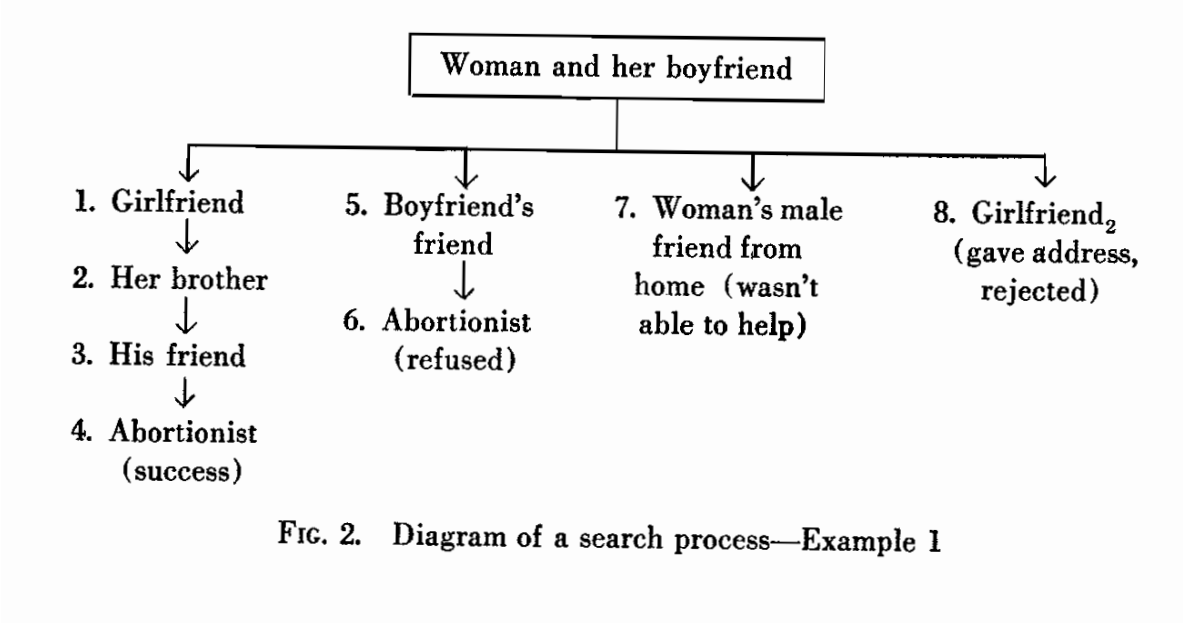


FIG. 2. Diagram of a search process—Example 1

Summary statistics:

- 8 people involved
- 4 fresh starts
- 4 links to abortionist (3 intermediaries)

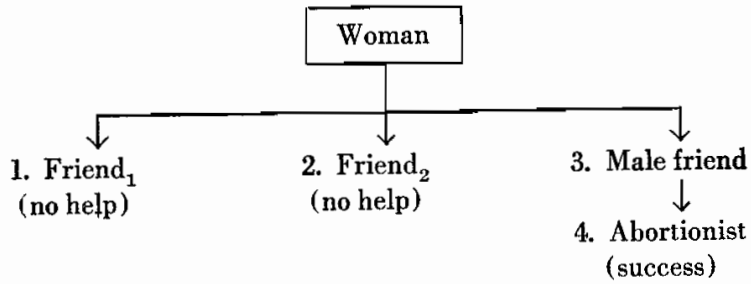


FIG. 3. Diagram of a search process—Example 2

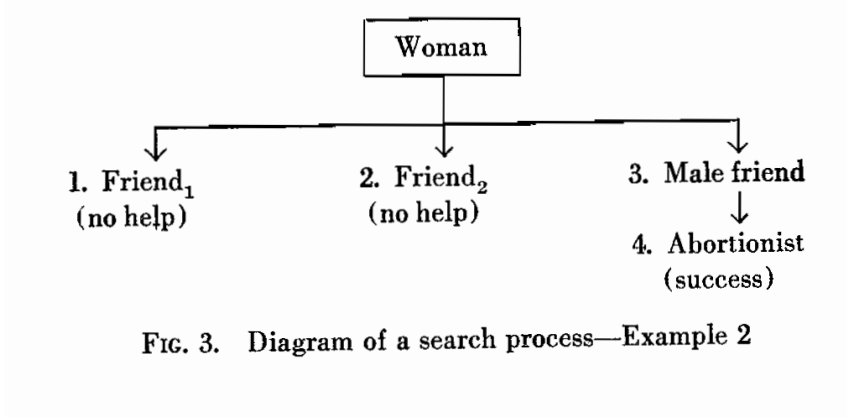


FIG. 3. Diagram of a search process—Example 2

Summary statistics:

- people involved:
- fresh starts:
- links to abortionist:

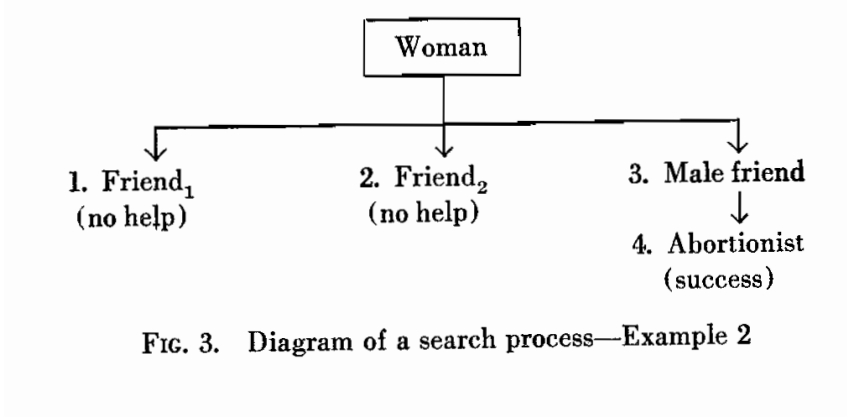


FIG. 3. Diagram of a search process—Example 2

Summary statistics:

- people involved: 4
- fresh starts:
- links to abortionist:

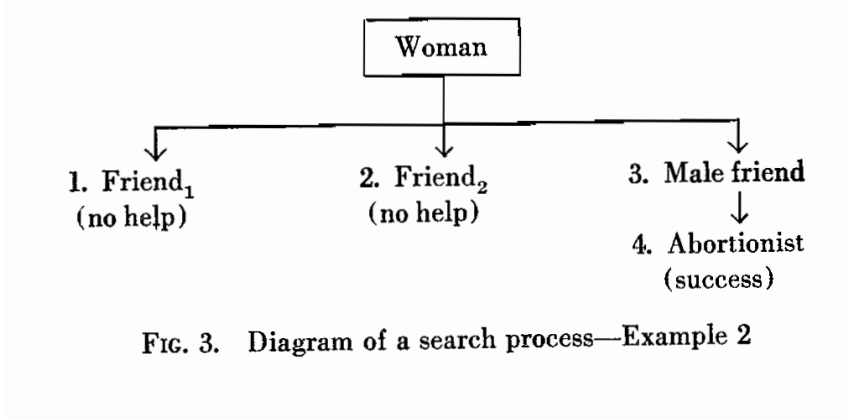


FIG. 3. Diagram of a search process—Example 2

Summary statistics:

- people involved: 4
- fresh starts: 3
- links to abortionist:

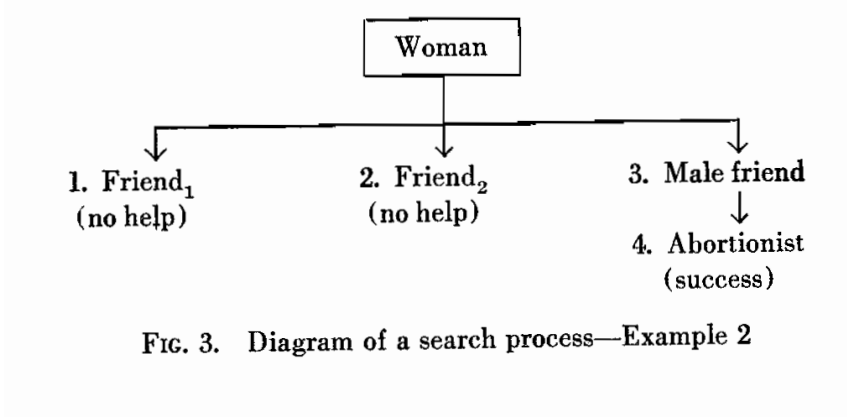


FIG. 3. Diagram of a search process—Example 2

Summary statistics:

- people involved: 4
- fresh starts: 3
- links to abortionist: 2

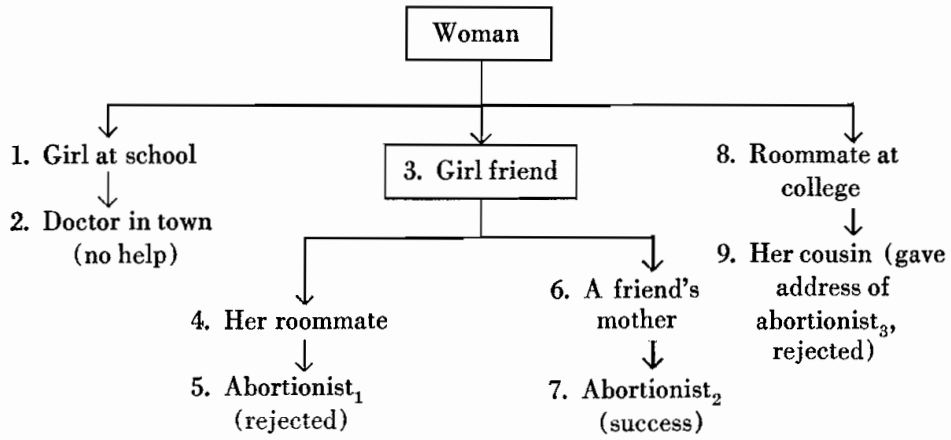


FIG. 4. Diagram of a search process—Example 3

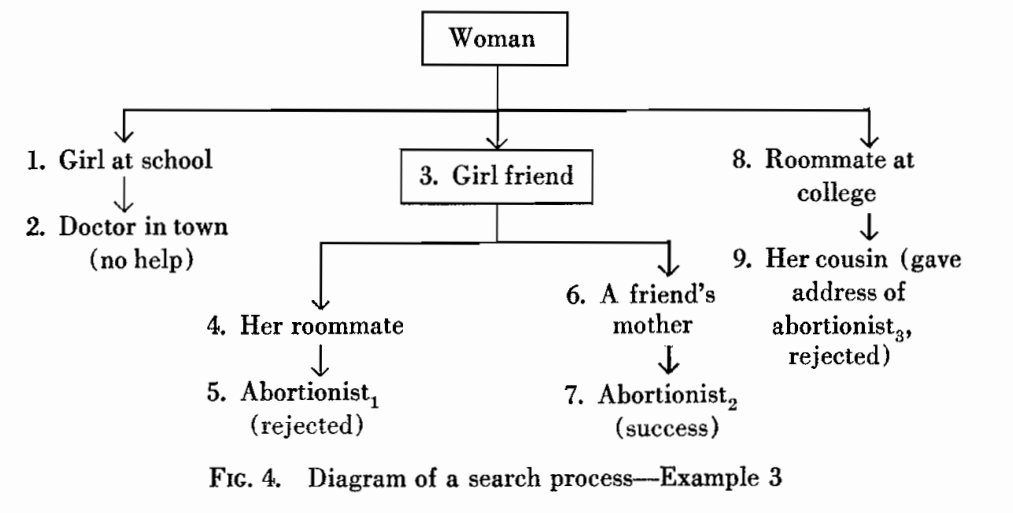


FIG. 4. Diagram of a search process—Example 3

Summary statistics:

- people involved:
- fresh starts:
- links to abortionist:

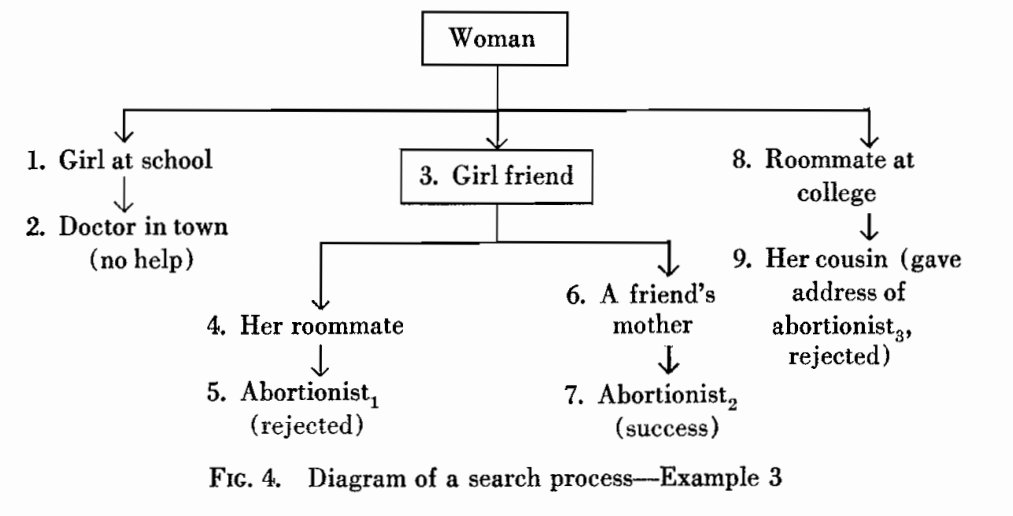


FIG. 4. Diagram of a search process—Example 3

Summary statistics:

- people involved: 9
- fresh starts:
- links to abortionist:

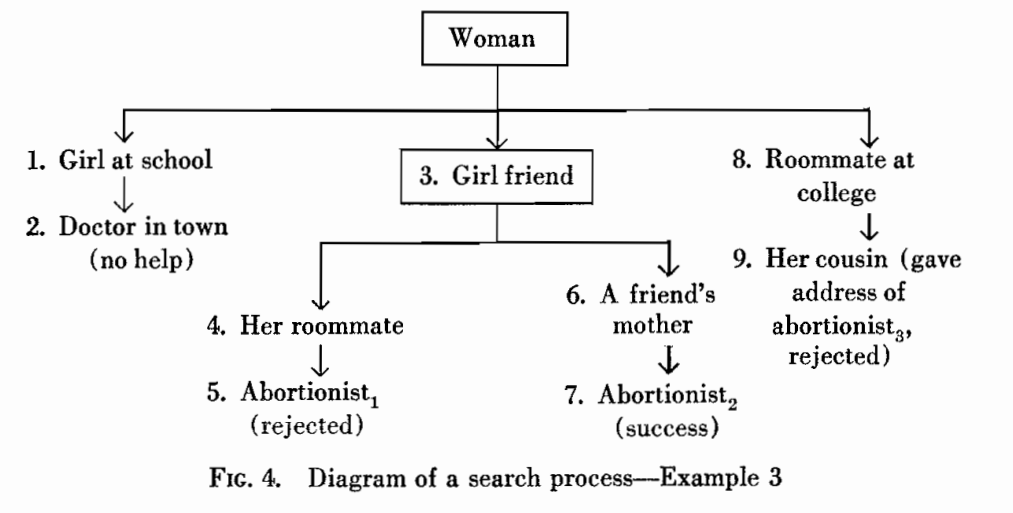


FIG. 4. Diagram of a search process—Example 3

Summary statistics:

- people involved: 9
- fresh starts: 3
- links to abortionist:

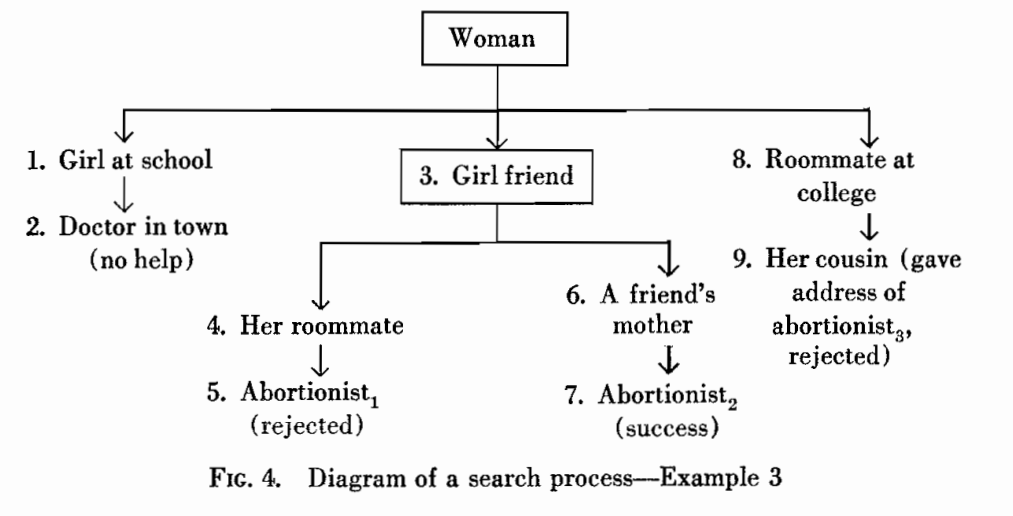


FIG. 4. Diagram of a search process—Example 3

Summary statistics:

- people involved: 9
- fresh starts: 3
- links to abortionist: 3

TABLE 15

TOTAL NUMBER CONSULTED DURING THE SEARCH,
BY NUMBER OF FRESH STARTS PER SEARCH

Number of Fresh Starts	Number of Persons Consulted									
	1	2	3	4	5	6	7	8	9+	Total
1	1	19	11	5	5	1	0	1	1	44
2		—	2	6	1	3	0	2	1	15
3			—	3	4	4	2	3	0	16
4				1	4	5	3	2	1	16
5					—	4	4	1	2	11
6						—	0	1	3	4
7							—	2	3	5
8								—	3	3
Total	1	19	13	15	14	17	9	12	14	114

TABLE 17

LENGTH OF CHAIN THAT LED TO ABORTIONIST USED

Length of Chain (x)	N	Number of Persons in Successful Chains (x · N)	Cumulative Percent Who Reached Abortionist by Chain of x or Less
1	2	2	2
2	52	104	47
3	34	102	77
4	11	44	87
5	10	50	96
6	1	6	97
7	1	7	98
Total	111*	315	

Note: Median = 2.0; Mean = 2.83.

*Three women are omitted in this table: two induced their own abortions, and one successful chain was not adequately described.

TABLE 17

LENGTH OF CHAIN THAT LED TO ABORTIONIST USED

Length of Chain (x)	N	Number of Persons in Successful Chains (x · N)	Cumulative Percent Who Reached Abortionist by Chain of x or Less
1	2	2	2
2	52	104	47
3	34	102	77
4	11	44	87
5	10	50	96
6	1	6	97
7	1	7	98
Total	111*	315	

Note: Median = 2.0; Mean = 2.83.

*Three women are omitted in this table: two induced their own abortions, and one successful chain was not adequately described.

chain length of 2 was most common

TABLE 18

SOURCES OF SUCCESSFUL FRESH STARTS

Kinds of People	Number of Fresh Starts	Successful Fresh Starts	Percent Successful
Relatives			
Mother or parents	14	8	57
Sister, brother, or cousin	6	4	66
Other relatives	4	0	0
Friends			
Girl friends	102	23	23
Male friends	22	11	50
Older friends	8	4	50
Other friends	27	8	29
Relative of a friend, friend of a friend	22	1	5
Man involved in pregnancy	45	24	53
Doctors			
Personal physician	26	7	27
Psychiatrist	4	1	25
Doctor recommended by others	5	3	60
Doctor selected by chance	12	4	33
Other			
Abortionists	2	2	100
"Abortion specialist"	7	2	28
Acquaintances, met during search	16	7	44
Other	2	2	100
Total	324	111	

TABLE 18

SOURCES OF SUCCESSFUL FRESH STARTS

Kinds of People	Number of Fresh Starts	Successful Fresh Starts	Percent Successful
Relatives			
Mother or parents	14	8	57
Sister, brother, or cousin	6	4	66
Other relatives	4	0	0
Friends			
Girl friends	102	23	23
Male friends	22	11	50
Older friends	8	4	50
Other friends	27	8	29
Relative of a friend, friend of a friend	22	1	5
Man involved in pregnancy	45	24	53
Doctors			
Personal physician	26	7	27
Psychiatrist	4	1	25
Doctor recommended by others	5	3	60
Doctor selected by chance	12	4	33
Other			
Abortionists	2	2	100
"Abortion specialist"	7	2	28
Acquaintances, met during search	16	7	44
Other	2	2	100
Total	324	111	

Weak ties not always very effective, probably because of effort

Title:	Getting a Job: An Analysis of Networking Techniques used by Princeton Students in the Employment Market
Authors:	Ott, Gregory
Advisors:	Salganik, Matthew
Department:	Sociology
Class Year:	2013
Abstract:	Building off the work of Mark Granovetter and Nancy Howell Lee, this thesis observes the employment search networks of Princeton University seniors. I interviewed thirty-five students who were actively or had recently completed searches for employment and constructed visual representations of each student's search network. I found that students searching for employment in locations geographically close to home conduct more focused searches and students searching for employment in industries requiring specific technological or scientific knowledge or skill conducted shorter, more efficient searches. I also found evidence supporting Granovetter's 1973 hypothesis regarding the strength of weak ties, as Princeton students use far more weak ties in searches for employment. All students used impersonal methods of communication most frequently, but also indicated the power of personal methods of communication.

- small world experiments show both 1) short paths exist and 2) people can find them with local information

- small world experiments show both 1) short paths exist and 2) people can find them with local information
- new models explain why social networks might be searchable

- small world experiments show both 1) short paths exist and 2) people can find them with local information
- new models explain why social networks might be searchable
- examples of directed searches in networks with real consequences

Turning point in the class:
Network structure $\rightarrow$ dynamics on networks

Small world TV show

http://bit.ly/socnet204

http://bit.ly/socnet204

Next class

- Watts, Chapter 6. (not posted on Blackboard)
- Bearman, P.S., Moody, J.M., and Stovel, K. (2004). Chains of affection: The structure of adolescent romantic and sexual networks. American Journal of Sociology.