MIDTERM EXAMINATION
Networked Life (NETS 112)
October 23, 2018
Prof. Michael Kearns

This is a closed-book exam. You should have no material on your desk other than the exam itself and a pencil or pen. If you run out of room on a page, you may use the back, but be sure to indicate you have done so. You may also make annotations directly on any diagrams given.

Name:

Problem 1: _____/10

Problem 2: _____/10

Problem 3: _____/15

Problem 4: _____/15

Problem 5: _____/15

Problem 6: _____/10

Problem 7: _____/10

Problem 8: _____/15

TOTAL: _____/100

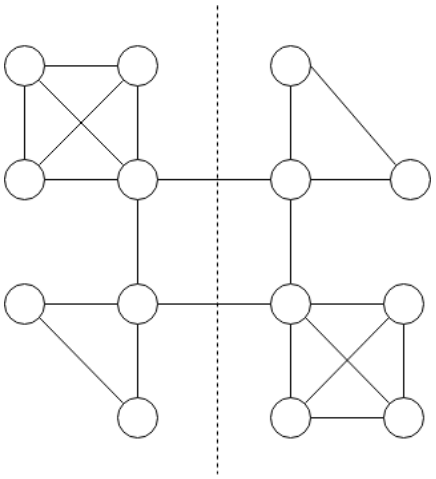
Problem 1 (10 points). Next to each lettered item on the left, write the number of the item on the right that is the best match. It might help to do the ones that are most obvious to you first.

(a) Facebook recruiting	1. Travers and Milgram
(b) few neurons	2. tennis ball
(c) cascade tree	3. C. Elegans
(d) Sharon, MA	4. no heavy tails
(e) no cycles	5. wiggly curves
(f) Erdos-Renyi	6. virality calculations
(g) audience participation	7. structural diversity
(h) overfitting	8. petitions
(i) highest virality	9. heavy tails
(j) file sizes	10. Preferential Attachment

Problem 2 (10 points). Consider the network in which there is a vertex for each word in the English language, and there is an edge between two words if they have appeared together in the same sentence in a New York Times article published in the last year. For example, there would be an edge between the words “dissident” and “journalist” due to a sentence in Saturday’s NYT, but it might be quite unlikely there is an edge between (say) “alpine” and “dissident”.

For each of the universal network properties discussed in class and readings, indicate whether you think this network will exhibit that property, and briefly but clearly justify your answer.

Problem 3 (15 points). Consider the network shown below.



(a) Ignoring the dotted line, calculate the clustering coefficient of this network.

(b) Calculate the overall edge density of this network.

(c) Calculate the edge density on just the left side of the dotted line, and also on just the right side of the dotted line. Exclude the edges crossing the dotted line, as they are neither on the left side nor the right side.

(d) Calculate the edge density between the vertices on the left and right sides of the dotted line.

(e) Based on your calculations above and class discussions, would you say this network is highly clustered or not? Briefly but precisely justify your answer.

Problem 4 (15 points). The table below is reproduced from Wikipedia, and was discussed in class. In it, ***k*** refers to (worst-case) diameter, and ***d*** refers to degree.

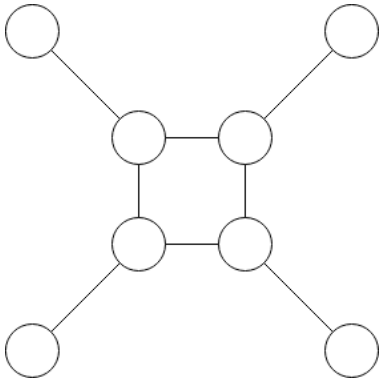
d \ *k*	2	3	4	5	6	7	8	9	10
3	10	20	38	70	132	196	336	600	1250
4	15	41	98	364	740	1 320	3 243	7 575	17 703
5	24	72	212	624	2 772	5 516	17 030	57 840	187 056
6	32	111	390	1404	7 917	19 383	76 461	331 387	1 253 615
7	50	168	672	2 756	11 988	52 768	249 660	1 223 050	6 007 230
8	57	253	1 100	5 060	39 672	131 137	734 820	4 243 100	24 897 161
9	74	585	1 550	8 268	75 893	279 616	1 697 688	12 123 288	65 866 350
10	91	650	2 286	13 140	134 690	583 083	4 293 452	27 997 191	201 038 922
11	104	715	3 200	19 500	156 864	1 001 268	7 442 328	72 933 102	600 380 000
12	133	786	4 680	29 470	359 772	1 999 500	15 924 326	158 158 875	1 506 252 500
13	162	851	6 560	40 260	531 440	3 322 080	29 927 790	249 155 760	3 077 200 700
14	183	916	8 200	57 837	816 294	6 200 460	55 913 932	600 123 780	7 041 746 081
15	187	1 215	11 712	76 518	1 417 248	8 599 986	90 001 236	1 171 998 164	10 012 349 898
16	200	1 600	14 640	132 496	1 771 560	14 882 658	140 559 416	2 025 125 476	12 951 451 931

a) As clearly and precisely as you can, explain what the numbers in the table are measuring. Do the numbers in the table represent empirical measurements, or are they from mathematical theories?

b) What can you observe about the growth rate of the numbers? Explain why it makes broad sense.

c) Consider the entry for 7 041 746 081 , which is approximately the size of the world population. Are the values for which this entry is achieved consistent with the readings from class? Justify your answer.

Problem 5 (15 points). Consider the network shown below.



(a) What is its worst-case diameter?

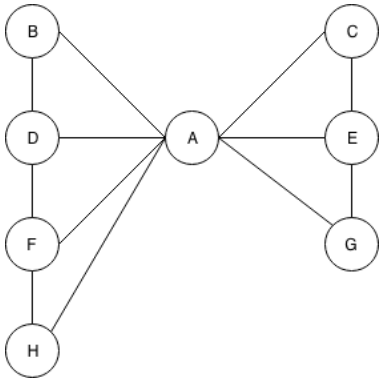
(b) Clearly add two edges so that the worst-case diameter decreases by exactly 1.

(c) Draw a network of 5 vertices and 6 edges such that the worst-case diameter is 3, and when you remove 2 edges of your choice (clearly marked on your drawing), the worst-case diameter becomes 4.

Problem 6 (10 points). Fill in the following table by writing either "yes" or "no" in each entry.

	Erdos-Renyi	**Preferential Attachment**
Giant Component		
Small Diameter		
Sparsity		
Heavy Tails		
High Clustering		

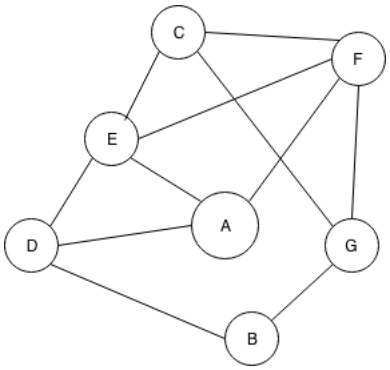
Problem 7 (10 points).



Time	Nodes who shared
1	A
2	B, C
3	D
4	H
5	E
6	F
7	G

(a) Draw the cascade tree for network above and the given timestamps of when each vertex shared. Break ties by attributing the cause of resharing to the neighbor who shared **most recently.**

(b) Do the same, except break ties by attributing the cause of resharing to the neighbor who shared it **first**.



Time	Nodes who shared
1	C
2	E
3	D
4	F
5	A
6	G

(c) Draw the cascade tree for network above and the given timestamps of when each vertex shared. Break ties by attributing the cause of resharing to the neighbor who shared **most recently.**

(d) Compute the virality of the cascade tree for part (c).

Problem 8 (15 points). One of the assigned readings begins with the following passage:

> *"A crucial task in the analysis of on-line social-networking systems is to identify important people — those linked by strong social ties — within an individual's network neighborhood. Here we investigate this question for a particular category of strong ties, those involving spouses or romantic partners."*

(a) Briefly but clearly describe what the source of data for the study was. Be sure to indicate exactly what information was used, and what information was not used.

(b) Briefly but clearly describe the question or hypothesis the authors were investigating.

(c) Briefly but clearly describe the findings or results of the study.