# AI607: GRAPH MINING AND SOCIAL NETWORK ANALYSIS (FALL 2020)

# Homework 5: Take-home Final Exam

Release: Dec 04, 2020,
Due: Dec 18, 2020, 11:59pm
(Late submissions will not be accepted. Late passes cannot be used for this homework.)

## 1 Linear Threshold Model (25 points)

Recall the *linear threshold model*, which can be described below:

- Consider a graph $G$ where every node starts with behavior $B$.
- A small set $S$ of early adopters adopt $A$ and never changes their behavior.
- Each of the other nodes switches from $B$ to $A$ if and only if at least a fraction $q$ of their neighbors have adopted $A$.
- The process terminates if no node changes their behavior.

**Definition 1.** *A* ***cluster of density*** $p$ *is defined as a set of nodes such that each node in the set has at least a* $p$ *fraction of its neighbors in the set.*

**Definition 2.** *We say that a* ***complete cascade*** *happens if eventually every node switches from* $B$ *to* $A$*.*

1. **(15 points)** Prove or disprove the following claim: *"If there exists a cluster of density greater than* $1 - q$ *such that no early adopter belongs to the cluster, then the set of initial adopters will not cause a complete cascade."*

2. **(10 points)** Prove or disprove the following claim: *"If a set of initial adopters does not cause a complete cascade, there must exists a cluster of density greater than* $1 - q$ *such that no early adopter belongs to the cluster."*

## 2 Triangle Counting (25 points)

Consider a graph $G$. $G$ contains $\tau$ triangles (i.e., 3-cliques), and $k$ (unordered) pairs of the triangles share an edge. We delete each edge in $G$ independently with probability $1-p$. Let $G'$ be the remaining graph, and let $X$ be the number of triangles in $G'$.

1. **(15 points)** Prove or disprove the following claim: $E[X] = p^3\tau$.
2. **(10 points)** Prove or disprove the following claim: $Var(X) = \tau(p^3 - p^6) + 2k(p^5 - p^6)$

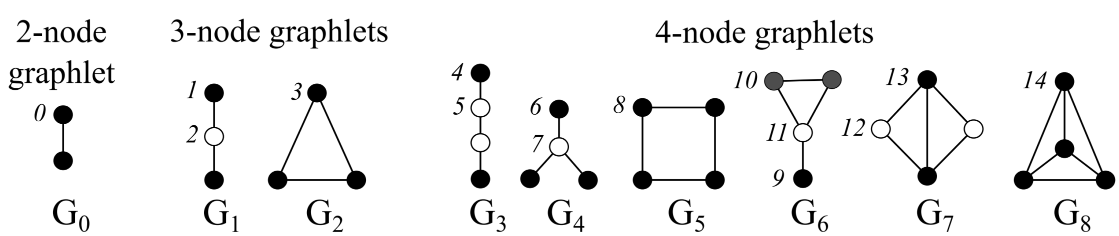
## 3 Kronecker Model (30 points)

Consider an initiator graph be $G_1$ without self-loops. In the (deterministic) Kronecker model, by recursion, we produce successively larger graphs $G_2, G_3, \cdots$.

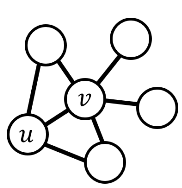
1. **(15 points)** Prove or disprove the following claim: *"If we let $d_k^{max}$ be the maximum degree of nodes in $G_k$, then $d_k^{max} = (d_1^{max})^k$."*
2. **(15 points)** Prove or disprove the following claim: *"If we let $D_k$ be the diameter of $G_k$, then $D_k = (D_1)^k$."*

## 4 Graphlet (20 points)

Below, we show all 2-, 3-, and 4-node graphlets, from which we can calculate a Graphlet Degree Vectors of size 14 (GDV-14).




Consider the following graph $G$:




1. **(10 points)** What is the GDV-14 of node $u$?
2. **(10 points)** What is the GDV-14 of node $v$?

# 5 How to submit your assignment

1. Create hw5-[your student id].tar.gz, which should contain the following files:
   - **report.pdf**: this should contain your answers.
2. Make sure that no other files are included in the tar.gz file.
3. Submit the tar.gz file at KLMS (`http://klms.kaist.ac.kr`).
4. If you have any questions, please post them on CLASSUM. Any ideas can be taken into consideration when grading if they are written in the *readme* file.