Course Syllabus

18-687: Analytical Performance Modeling & Design of Computer Systems

Spring, 2015

Instructor: Osman Yagan
Office Location: SV 23-221
Email Address: oyagan@ece.cmu.edu
Office Hours: Fri 2-3pm PST, or appt

Teaching Assistant(s): Serim Park
Office Location: HH C120
Email Address: serimp@ece.cmu.edu
Office Hours: Fri 4-5pm EST, or by appt

Course Management Assistant: Stephanie Scott
Email: scottsm@andrew.cmu.edu
Office Location: B19 Room 1052

Course Description: In designing computer systems one is usually constrained by certain performance requirements and limitations. For example, one might need to guarantee a response time SLA or certain throughput requirement, while at the same time staying within a power budget or cost budget. On the other hand, one often has many choices: One fast disk, or two slow ones? More memory, or a faster processor? A fair scheduler or one that minimizes mean response time? For multi-server systems, one can choose from a wide array of load balancing policies, a wide array of migration policies, capacity provisioning schemes, power management policies ... The possibilities are endless. The best choices are often counter-intuitive. Ideally, one would like to have answers to these questions before investing the time and money to build a system. This class will introduce students to analytic stochastic modeling with the aim of answering the above questions. Throughout, the theory developed will be applied to a wide array of computer systems design problems including the design of efficient data centers, web servers, DBMS, disks, call centers, routers, and supercomputer centers.

Number of Units: 12

Pre-requisites: 36-217

Pre-requisite for: -

Breadth Area: Algorithms/Complexity/Programming Languages
Undergraduate Course Designation: -

Undergraduate Course Area: -

Class Schedule:
• Lecture: 4hrs/week. MW 10:30am-12:20pm PST.
• Labs/Recitation: 2hrs/week. F 10:30am-12:20pm PST.

Required Textbook:

Suggested Reading:


Other Supplemental Materials:

Brief List of Topics Covered:
• Operational Laws: Little's Law, response-time law, asymptotic bounds, modification analysis, performance metrics;
• Markov Chain Theory: discrete-time Markov chains, continuous-time Markov chains, renewal theory, time-reversibility; Poisson Process: memorylessness, Bernoulli splitting, uniformity, PASTA;
• Queueing Theory: open networks, closed networks, time-reversibility, Renewal-Reward, M/M/1, M/M/k, M/M/k/k, Burke's theorem, Jackson networks, classed networks, load-dependent servers, BCMP result and proof, M/G/1 full analysis, M/G/k, G/G/1, transform analysis (Laplace and z-transforms);
• Simulations: time averages versus ensemble averages, generating random variables for simulation, Inspection Paradox;
• Modeling Empirical Workloads: heavy-tailed property, Pareto distributions, heavy-tailed distributions, understanding variability and tail behavior, Matrix-analytic methods;
• Management of Server Farms: capacity provisioning, dynamic power management, routing policies;
• Analysis of Scheduling: FCFS, non-preemptive priorities, preemptive priorities, PS, LCFS, FB, SJF, PSJF, SRPT, etc.

Course Blackboard: To access the course blackboard from an Andrew Machine, go to the login page at: http://www.cmu.edu/blackboard. You should check the course blackboard daily for announcements and handouts.

Course Wiki:
Students are encouraged to use the ECE wiki to provide feedback about the course at: http://wiki.ece.cmu.edu/index.php.

Grading Algorithm:

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>45%</td>
<td>Homework</td>
</tr>
<tr>
<td>25%</td>
<td>Midterm</td>
</tr>
<tr>
<td>30%</td>
<td>Final</td>
</tr>
</tbody>
</table>

HW Policy: You will receive regular homework problems. These will be difficult. Start immediately so that you can take full advantage of office hours. You will find office hours very helpful! Some of these homework problems may be repeated from previous years. Do not ask people who took this course in previous years to help you with the homeworks. This is considered cheating and will be reported to the dean. On the other hand, I strongly encourage you to collaborate with your current classmates to solve the homework problems after you have tried solving them by yourself. Each person must turn in a separate write-up. You should note on your homework specifically which problems were a collaborative effort and with whom.

Tentative Course Calendar

<table>
<thead>
<tr>
<th>Date</th>
<th>Day</th>
<th>Class Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Mon</td>
<td>Introduction to class. Queuing theory basics.</td>
</tr>
<tr>
<td>14</td>
<td>Wed</td>
<td>Probability Review</td>
</tr>
<tr>
<td>19</td>
<td>Mon</td>
<td>MARTIN LUTHER KING</td>
</tr>
<tr>
<td>21</td>
<td>Wed</td>
<td>Operational Laws: Little's Law, response-time law, asymptotic bounds</td>
</tr>
<tr>
<td>26</td>
<td>Mon</td>
<td>Modification analysis, performance metrics</td>
</tr>
<tr>
<td>28</td>
<td>Wed</td>
<td>Markov Chain Theory: discrete-time Markov chains</td>
</tr>
<tr>
<td>February</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Mon</td>
<td>Discrete-time MC cont d. Real world examples: PageRank, Aloha.</td>
</tr>
<tr>
<td>4</td>
<td>Wed</td>
<td>Continuous-time Markov chains. Renewal theory, time-reversibility.</td>
</tr>
<tr>
<td>9</td>
<td>Mon</td>
<td>Poisson Process: memorylessness, Bernoulli splitting, uniformity, PASTA</td>
</tr>
<tr>
<td>11</td>
<td>Wed</td>
<td>Queuing Theory: open networks, closed networks, time-reversibility</td>
</tr>
<tr>
<td>16</td>
<td>Mon</td>
<td>Renewal-Reward, M/M/1</td>
</tr>
<tr>
<td>18</td>
<td>Wed</td>
<td>M/M/k, M/M/k/k, Burke's theorem</td>
</tr>
<tr>
<td>23</td>
<td>Mon</td>
<td>Jackson networks, classed networks, load-dependent servers, BCMP result and proof,</td>
</tr>
<tr>
<td>25</td>
<td>Wed</td>
<td>Transform analysis (Laplace and z-transforms), M/G/1 full analysis.</td>
</tr>
<tr>
<td>March</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Mon</td>
<td>M/G/k, G/G/1</td>
</tr>
<tr>
<td>4</td>
<td>Wed</td>
<td>Test 1</td>
</tr>
</tbody>
</table>
Education Objectives (Relationship of Course to Program Outcomes)

(a) an ability to apply knowledge of mathematics, science, and engineering:
<<description>>

(b) an ability to design and conduct experiments, as well as to analyze and interpret data: <<description>>

(c) an ability to design a system, component, or process to meet desired needs within realistic constraints such as economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability: <<description>>

(e) an ability to identify, formulate, and solve engineering problems:
<<description>>

(h) the broad education necessary to understand the impact of engineering solutions in a global, economic, environmental, and societal context:
<<description>>

(i) a recognition of the need for, and an ability to engage in life-long learning

(j) a knowledge of contemporary issues: <<description>>
(k) an ability to use the techniques, skills, and modern engineering tools necessary for engineering practice: <<description>>

Academic Integrity Policy (http://www.ece.cmu.edu/student/integrity.html):

The Department of Electrical and Computer Engineering adheres to the academic integrity policies set forth by Carnegie Mellon University and by the College of Engineering. ECE students should review fully and carefully Carnegie Mellon University’s policies regarding Cheating and Plagiarism; Undergraduate Academic Discipline; and Graduate Academic Discipline. ECE graduate student should further review the Penalties for Graduate Student Academic Integrity Violations in CIT outlined in the CIT Policy on Graduate Student Academic Integrity Violations. In addition to the above university and college-level policies, it is ECE’s policy that an ECE graduate student may not drop a course in which a disciplinary action is assessed or pending without the course instructor's explicit approval. Further, an ECE course instructor may set his/her own course-specific academic integrity policies that do not conflict with university and college-level policies; course-specific policies should be made available to the students in writing in the first week of class.

This policy applies, in all respects, to this course.

Carnegie Mellon University’s Policy on Cheating and Plagiarism (http://www.cmu.edu/policies/documents/Cheating.html) states the following,

Students at Carnegie Mellon are engaged in preparation for professional activity of the highest standards. Each profession constrains its members with both ethical responsibilities and disciplinary limits. To assure the validity of the learning experience a university establishes clear standards for student work.

In any presentation, creative, artistic, or research, it is the ethical responsibility of each student to identify the conceptual sources of the work submitted. Failure to do so is dishonest and is the basis for a charge of cheating or plagiarism, which is subject to disciplinary action.

Cheating includes but is not necessarily limited to:
1. Plagiarism, explained below.
2. Submission of work that is not the student's own for papers, assignments or exams.
3. Submission or use of falsified data.
4. Theft of or unauthorized access to an exam.
5. Use of an alternate, stand-in or proxy during an examination.
6. Use of unauthorized material including textbooks, notes or computer programs in the preparation of an assignment or during an examination.
7. Supplying or communicating in any way unauthorized information to another student for the preparation of an assignment or during an examination.
8. Collaboration in the preparation of an assignment. Unless specifically permitted or required by the instructor, collaboration will usually be viewed by the university as cheating. Each student, therefore, is responsible for understanding the policies of the department offering any course as they refer to the amount of help and collaboration permitted in preparation of assignments.

9. Submission of the same work for credit in two courses without obtaining the permission of the instructors beforehand.

Plagiarism includes, but is not limited to, failure to indicate the source with quotation marks or footnotes where appropriate if any of the following are reproduced in the work submitted by a student:
1. A phrase, written or musical.
2. A graphic element.
3. A proof.
4. Specific language.
5. An idea derived from the work, published or unpublished, of another person.

This policy applies, in all respects, to 18-687