ECE/CS 541
Computer System Analysis:
Introduction to Combinatorial Methods

Mohammad A. Noureddine
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign

Fall 2018
Learning Objectives

• Or what is this course about?

• At the start of the semester, you should have
  – Basic programming skills (C++, Python, etc.)
  – Basic understanding of probability theory (ECE313 or equivalent)

• At the end of the semester, you should be able to
  – Understand different system modeling approaches
    • Combinatorial methods, state-space methods, etc.
  – Understand different model analysis methods
    • Analytic/numeric methods, simulation
  – Understand the basics of discrete event simulation
  – Design simulation experiments and analyze their results
  – Gain hands-on experience with different modeling and analysis tools
Announcements and Reminders

• HW1 is out
  – Covers the probability review
  – Prepare you for the probability quiz
  – Due on September 18, 2018 at the start of class

• Probability quiz on September 20, 2018
  – First 30 minutes of class

• Project Proposals due near the first week of October
  – Start forming groups and thinking about your projects
  – Come to office hours for discussions
Objectives for this Module

- Introduce combinatorial (non state-space) methods of modeling
- Develop and formulate models of system reliability
- Introduce different reliability formalisms
- Combinatorial models for improved testing research at Internet scale
  - Technique generated out of UC Santa Cruz and adopted by Netflix
Lecture Outline

• Assumptions for combinatorial modeling
• Review definition of reliability
• Failure rate
• System reliability
  – Maximum
  – Minimum
Introduction to Combinatorial Methods

• Combinatorial validation methods are the simplest kind of analytical/numerical techniques and can be used for reliability and availability modeling under certain assumptions.

• **Assumption 1:**
  – The system being studied is composed of several elementary units, called *components*.

• **Assumption 2:**
  – The components of the system fail in a statistically independent manner. For availability analysis, they can be repaired independently.

• When these assumptions hold, simple formulas for reliability and availability exist.
Choosing Validation Techniques cont.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Combinatorial</th>
<th>State-Space-Based</th>
<th>Simulation</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
<td>Any</td>
<td>Any</td>
<td>Any</td>
<td>Post-prototype</td>
</tr>
<tr>
<td>Time</td>
<td>Small</td>
<td>Medium</td>
<td>Medium</td>
<td>Varies</td>
</tr>
<tr>
<td>Tools</td>
<td>Formulae, tools</td>
<td>Languages &amp; tools</td>
<td>Languages &amp; tools</td>
<td>instrumentation</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Low</td>
<td>Moderate</td>
<td>Moderate</td>
<td>high</td>
</tr>
<tr>
<td>Comparisons</td>
<td>Easy</td>
<td>Moderate</td>
<td>Moderate</td>
<td>Difficult</td>
</tr>
<tr>
<td>Cost</td>
<td>Low</td>
<td>Low/medium</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Scalability</td>
<td>High</td>
<td>Low/medium</td>
<td>Medium</td>
<td>low</td>
</tr>
</tbody>
</table>
Reliability

• One key to building highly available systems is the use of reliable components and systems.

• Reliability:
  
  – The *reliability* of a system at time $t$ ($R(t)$) is the probability that the system operation is proper throughout the interval $[0,t]$.

• Probability theory and combinatorics can be directly applied to reliability models.

• Let $X$ be a random variable representing the *time to failure* (TTF) of a component. The reliability of the component at time $t$ is given by

  $$R_X(t) = P(X > t) = 1 - P(X \leq t) = 1 - F_X(t)$$

• Similarly, we can define *unreliability* at time $t$ by

  $$U_X(t) = P(X \leq t) = F_X(t)$$
Failure Rate

What is the rate that a component fails at time $t$? This is the probability that a component that has not yet failed fails in the interval $(t, t + \Delta t)$, as $\Delta t \to 0$.

Note that we are not looking at $f_X(t)dt = P(X \in (t, t + \Delta t))$. Rather, we are seeking $P(X \in (t, t + \Delta t) \mid X > t)$

$r_X(t)$ is called the **failure rate** or **hazard rate**.
Failure Rate

What is the rate that a component fails at time \( t \)? This is the probability that a component that has not yet failed fails in the interval \((t, t + \Delta t)\), as \( \Delta t \to 0 \).

Note that we are not looking at \( f_X(t)dt = P(X \in (t, t + \Delta t)) \). Rather, we are seeking \( P(X \in (t, t + \Delta t) \mid X > t) \)

\[
P(X \in (t, t + \Delta t) \mid X > t) = \frac{P(X \in (t, t + \Delta t), X > t)}{P(X > t)}
\]

\[
= \frac{P(X \in (t, t + \Delta t))}{1 - F_X(t)}
\]

\[
= \frac{f_X(t)dt}{1 - F_X(t)} \Delta r_x(t)dt
\]

\[
r_x(t) = \frac{f_X(t)}{1 - F_X(t)} = \frac{f_X(t)}{R_X(t)}
\]

\( r_x(t) \) is called the failure rate or hazard rate.
Typical Failure Rate

- Break in
- Normal operation
- Wear out

$r_X(t)$

time
System Reliability

While $R_X$ can give the reliability of a component, how do you compute the reliability of a system?

System failure can occur when one, all, or some of the components fail. If one makes the \textit{independent failure assumption}, system failure can be computed quite simply. The independent failure assumption states that all component failures of a system are independent, i.e., the failure of one component does not cause another component to be more or less likely to fail.

Given this assumption, one can determine:

1) Minimum failure time of a set of components
2) Maximum failure time of a set of components
3) Probability that $k$ of $N$ components have failed at a particular time $t$. 
Maximum of $n$ Independent Failure Times

Let $X_1, \ldots, X_n$ be independent component failure times. Suppose the system fails at time $S$ if all the components fail.

Thus, $S = \max\{X_1, X_2, \ldots, X_n\}$

What is $F_S(t)$?
Maximum of $n$ Independent Failure Times

Let $X_1, \ldots, X_n$ be independent component failure times. Suppose the system fails at time $S$ if all the components fail.

Thus, $S = \max\{X_1, X_2, \ldots, X_n\}$

What is $F_S(t)$?

$$F_S(t) = P(S \leq t) = P(X_1 \leq t \land X_2 \leq t \land \ldots \land X_n \leq t)$$

$$= P(X_1 \leq t) \times P(X_2 \leq t) \times \ldots \times P(X_n \leq t)$$

$$= F_{X_1}(t)F_{X_2}(t)\ldots F_{X_n}(t)$$

$$= \prod_{i=1}^{n} F_{X_i}(t)$$

By independence!

By definition!
Let $X_1, \ldots, X_n$ be independent component failure times. A system fails at time $S$ if any of the components fail.

Thus, $S = \min\{X_1, \ldots, X_n\}$.

What is $F_S(t)$? Proof in Homework 1

$$F_S(t) = 1 - \Pi_{i=1}^{n} (1 - F_X(t))$$