Final Report: Modeling Systems of Dependent Components

New classes of stochastic models for network systems having stochastically dependent components are studied by a combination of probabilistic analysis and efficient simulation techniques. For instance, in a model in which shocks of r different types occur, with component i failing when there have been a total of n(i) type i shocks, we give a method for studying the distribution of the number of shocks needed to cause the system to fail.

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrived as an official Department of the Army position, policy or decision, unless so designated by other documentation.
ABSTRACT

New classes of stochastic models for network systems having stochastically dependent components are studied by a combination of probabilistic analysis and efficient simulation techniques. For instance, in a model in which shocks of \( r \) different types occur, with component \( i \) failing when there have been a total of \( n(i) \) type \( i \) shocks, we give a method for studying the distribution of the number of shocks needed to cause the system to fail.

Stochastic optimization problems related to stochastic assignment and stochastic knapsack models have been studied. For instance, we consider the static stochastic problem where a subset of components of known values but having random weights is initially put into a knapsack, with a reward equal to the sum of their values obtained provided that their total weight does not exceed the capacity of the knapsack. New variance reduction simulation results related to normalized importance sampling and improved estimators of tail distributions have been obtained.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TOTAL:

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TOTAL:

Number of Papers published in non peer-reviewed journals:

(c) Presentations

The Infinite Server Queue with Exponentially Distributed Setup Times, International Workshop in Applied Probability, Antalya, Turkey, June 2014
<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL:**

---

**Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):**

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL:**

---

**Peer-Reviewed Conference Proceeding publications (other than abstracts):**

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL:**

---
(d) Manuscripts

Received  Paper

08/13/2012  3.00 Sheldon Ross. Simulation Analysis of System Life when Component Lives are Determined by Marked Shocks generated by a Poisson, a Renewal, or a Hawkes Point Process, ( )


08/13/2012  7.00 Sheldon Ross. A Markov Chai Choice Problem, Probability in the Engineering and Informational Sciences (07 2012)

08/13/2012  5.00 Samim Ghamami, Sheldon Ross. Improving the Asmussen-Kroese Type Simulation Estimators, Journal of Applied Probability (03 2012)

08/13/2012  4.00 Sheldon Ross. Improving the Normalized Importance Sampling Estimator, ( )

08/18/2011  1.00 Sheldon Ross. A System in which Component Failures Increase the Random Hazards of Still Working Components, (08 2011)

08/18/2011  2.00 Sheldon Ross. Distribution of Minimal Path Lengths when Edge Lengths are Independent Heterogeneous Exponential Random Variables, (08 2011)

TOTAL:  7

Number of Manuscripts:

Books

Received  Book

TOTAL:
Patents Submitted

Patents Awarded

Awards
Sheldon Ross was named Fellow of INFORMS, October 2013

---

**Graduate Students**

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
<th>Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babak Haji</td>
<td>0.25</td>
<td></td>
</tr>
</tbody>
</table>

FTE Equivalent: 0.25  
Total Number: 1

---

**Names of Post Doctorates**

<table>
<thead>
<tr>
<th>NAME</th>
<th>FTE Equivalent</th>
<th>Total Number:</th>
</tr>
</thead>
</table>

---

**Names of Faculty Supported**

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
<th>National Academy Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheldon Ross</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

FTE Equivalent: 0.09  
Total Number: 1

---

**Names of Under Graduate students supported**

<table>
<thead>
<tr>
<th>NAME</th>
<th>FTE Equivalent</th>
<th>Total Number:</th>
</tr>
</thead>
</table>

---
**Student Metrics**
This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: ...... 0.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields: ...... 0.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields: ...... 0.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale): ...... 0.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering: ...... 0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense ...... 0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: ...... 0.00

<table>
<thead>
<tr>
<th>NAME</th>
<th>Total Number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kai Chen</td>
<td>1</td>
</tr>
</tbody>
</table>

**Names of Personnel receiving masters degrees**

**Names of personnel receiving PHDs**

**Names of other research staff**

**Sub Contractors (DD882)**

**Inventions (DD882)**
We start with two problems tackled since the last interim progress report.

In [1] we study a system where items break down according to a Poisson process at which point they go to a repair facility. Upon their arrival a repairperson is summoned (there are an unlimited number of such individuals) with it taking a random exponentially distributed time for the repairperson to be ready to service the item. The service distribution is general. When a service is completed the server (e.g., repairperson) immediately begins a new repair on a waiting item (if there is one) and the repairperson who had been summoned for that item is told that they are not needed. Thus, the model can be regarded as a queueing model where a server has to "setup" before beginning a service. We obtain the surprising result that in steady state the number of servers in setup and the number actually serving are independent, with the latter being Poisson distributed with me an lambda E[Y] where lambda is the Poisson arrival rate and E[Y] is the mean service time. We also show that the remaining service times of those servers currently serving are independent and have the equilibrium service distribution. Moreover, the distribution of the number of servers in setup does not depend on the service distribution. Finally we generalize the model by allowing servers to go idle after completing a service, even when there are waiting customers.

In [2] we have analyzed an infinite server tandem queues in situations where an arrival is not allowed to depart the system before any earlier arrival. For instance, one application would be to a one lane highway with a single exit location in which faster moving vehicles are not allowed to pass slower ones.

In the three years of this grant, the following 13 granted supported papers have been written, with 8 published in technical journals, 1 accepted but not yet published, 3 currently being reviewed by a journal, and 1 not yet submitted.

[1] M/G/infinity with Exponentially Distributed Setup Times


The preceding papers deal with areas of reliability, queueing, simulation, and stochastic dynamic optimization, with these different techniques often being used in tandem. Among other things we have tackled models composed of components with dependent lifetimes, stochastic optimization models concerning assignments, and a variety of stochastic variations of the knapsack model. Some of the papers, such as [9] and [10] have natural applications in communication networks; others such as [3] and [4] give improvements in applying simulation techniques which can be used in a range of different problems.

Technology Transfer
**Objectives** To study new classes of stochastic models for network systems whose components function for random, usually dependent, lengths of time and then fail. To develop new probabilistic and simulation tools to enable these and other systems to be studied. To study stochastic optimization problems of resource allocation, with particular emphasis on stochastic knapsack problems.

**Approach** A synthesis of probabilistic analysis and efficient variance reduction simulation techniques. An innovative utilization of dynamic programming to analyze the stochastic optimization models.

**Scientific Challenges and Barriers** Because closed form solutions are, for the most part, not feasible in the probability models we consider and an unsophisticated simulation is not efficient, we need to develop approaches that effectively combine probabilistic analysis with simulation. In the majority of the optimization problems not only is it very difficult to explicitly determine optimal policies but even finding numerically tractable computational approaches is hard.

**Scientific Enablers** We have developed new approaches for analyzing probabilistic models; for instance, focusing on random hazards and reverse Poissonization. We have developed new simulation variance reduction techniques when needed. We have developed innovative ways to tackle stochastic optimization problems, first determining structural results about the optimal policy and then developing techniques for utilizing this structure to obtain good heuristic policies.

**Accomplishments** We have studied models related to the life of a system composed of components having dependent lifetimes. For instance, in [1] we consider a system, composed of components that are either working or failed at any point of time, that is itself either working or failed depending on the set of failed components. We suppose that shocks occur according to a point process and that, with specified probabilities, each shock is one of r types; furthermore we suppose that component i fails when there have a total of $n(i)$
type i shocks. We are interested in efficiently using simulation to estimate the moments and distribution of the lifetime of the system. Our analysis begins with studying N, the number of shocks that are needed for the system to fail. Noting that if the times at which shocks occur were distributed according to a Poisson process then the times until the individual components failed would be independent gamma random variables, leads us to an approach that first simulates these gamma random variables and then analyzes the conditional distribution of N given the vector of component lifetimes. This results in low variance estimators of the moments and the probability distribution of N. We then improve this conditional expectation estimator of the mean of N by taking a weighted average of it and a second derived estimator. Using these results we are then able to obtain efficient simulation estimators of the mean, variance, and tail probabilities of the system life when the shock process is a Poisson process, a renewal process, or a Hawkes process. Our method for studying N involves the trick of Poissonization, which studies the discrete random variable N by moving to continuous time by assuming that shocks occur according to a Poisson process. This is done because it results in the times at which the components fail being independent random variables (whereas the numbers of shocks until they fail are dependent). After employing this well-known trick, our innovation is to then come back to the discrete random variable N by analyzing its conditional distribution given the component lifetimes.

Another model of dependent components we have studied is one in which the failure of a component increases the instantaneous hazard rates of the components that are still working. We show in [2] that the total amounts of hazard that each component experiences before failing are independent exponential random variables with common mean 1, and then use this result to obtain efficient ways to simulate the system lifetime.

We have considered other simulation questions that relate to analyzing systems. In [3] we considered the problem of using simulation to estimate the probability that the sum of n independent random variables having a common distribution exceeds some value c, when this probability is small and the distribution is heavy tailed. Because of the heavy tailed assumption, the standard trick of importance sampling can not be used. This problem had previously been considered in the literature by Asmussen and Kroese, who came up with a clever idea of writing the probability that the random sum exceeds c as n multiplied by the joint probability that the sum exceeds c and that the last summand is the largest, and then taking their estimator.
to be the conditional expectation of this latter quantity given all but the final summand. In [3] we show how to improve upon this estimator by a second conditional expectation estimator that conditions on a smaller set of random variables (and thus always has a smaller variance). In addition we present very efficient estimators in the case where the number of terms in the sum is not fixed but rather is random. Our estimator in this case combines conditional expectation, stratification and a control variable, and makes use of “single simulation run stratification” which obtains estimators for all strata in each run.

Another grant supported paper dealing with simulation is the paper [4] which shows how to improve the normalized importance sampling estimator. The normalized importance sampling estimator allows one to estimate the expected value of a function of a random vector when the density of the vector is only specified up to a multiplicative constant. It works by using a different density function to generate random vectors and then uses as the estimator a weighted average of values of the function evaluated at the generated vectors with weights proportional to the ratios of the densities of the generated vectors. In [4] we show how this estimator can be derived by a delta method based approximation of a conditional expectation acceptance-rejection estimator. Using additional terms in the delta method then allows for a new estimator which, as indicated by numerical examples, usually outperforms the normalized importance sampling estimator in terms of mean square error.

The knapsack problem is a widely studied problem in combinatorial optimization. In its original form there are multiple items, with each item having a weight and a value, and the problem is to choose a subset of items to maximize the sum of the values of these items subject to a given capacity constraint on the sum of their weights. It is called the knapsack problem because one can imagine that the items are to be put in a knapsack that can only hold up to a certain weight before falling apart. The problem often arises in resource allocation where there are financial or time constraints. For instance, one might want to extract the maximum value for a set of candidate projects, each of which yields a certain profit but requires a given budget, when there is a constraint on the sum of the budgets of chosen projects, or, if the weight refers to time, the constraint could be on the time it entails to se-
quentially finish all projects. The knapsack problem has been applied in such areas as combinatorics, computer science, complexity theory, cryptography and applied mathematics.

To further expand the applicability of these models we have studied variations in which there is uncertainty connected to either the item values or weights. The grant supported paper [5] considered a stochastic knapsack model where the knapsack has a given capacity, and where there are n types of items with an unlimited number of each type being available. The value of an item depends on its type and is specified; however, the weight of an item is random with a distribution that depends on the item’s type. The objective is to decide how many of each type to put in the knapsack to maximize, for a specified return function, the expected return. In [5], we consider two sub-problems depending on the return function assumed. The broken knapsack problem supposes that there is no reward earned if the sum of the weights of the items put in the knapsack exceeds its capacity, while the return is the sum of the values of the inserted items if their total weight does not exceed the capacity. The recourse and penalty cost problem supposes that the sum of the values of the items put in the knapsack is always retained but a fixed penalty plus a linear overcapacity cost is incurred if the knapsack’s capacity is exceeded. A policy for either of these stochastic knapsack models is a vector \( k(1), \ldots, k(n) \) with the interpretation that \( k(i) \) type i items are to be placed in the knapsack for each i. In the broken knapsack model we show that if all but one of the values of the vector are fixed, then the expected return as a function of the remaining variable is a unimodal function of that variable provided that the weight distributions all have a decreasing reverse hazard rate. (This result is not true without this assumption.) Using this enables us to determine a search algorithm for finding the optimal strategy. We also show that the unimodality property only holds in the recourse and penalty model when the penalty value is 0 and the distributions have decreasing reverse hazard rates. However, it does hold for positive penalty costs when the weight distributions are all exponential. We also show that our previous results all remain true when, rather than being a fixed constant, the capacity of the knapsack is exponentially distributed with a specified mean.

Whereas the paper [5] dealt with a static decision problem in that all decisions are made at once, we have also studied a variety of dynamic stochastic knapsack problems where some information is gained after each item is put in the knapsack. In all of these we suppose that the problem ends with 0 return if the sum of the weights of items in the knapsack exceeds its capacity.
and that you are allowed to stop at any earlier time with a return equal to the sum of the knapsack’s current values. We chose this “zero reward if knapsack is broken” criterion as it is the most conservative and thus serves as a bound in similar problems which allows some return even when the knapsack is broken. In the paper [6] we suppose that there is an unlimited number of each of \( n \) types of items, and that the weight of an item is exponentially distributed with a known mean depending on its type. In addition we suppose that the value of an item is proportional to its weight with a proportionality factor that depends on the type of the item. (For instance, one might be loading minerals of known type onto a spaceship, with value per weight of each type mineral being specified.) It is supposed that an item’s weight is learned after it is placed in the knapsack. By analyzing this as a dynamic programming problem we are able to explicitly give the optimal policy when there are only 2 types of items and to come up with a heuristic approach when there are more than 2 types. To establish our results we first consider the case where there are only type \( i \) items available, and thus the decision is whether to stop or not. In this case it is easy to show that the optimal policy is one that stops whenever the return from immediate stopping is larger than the expected return if one puts in one additional type \( i \) item and then stops. Using this we show that there is a function \( v(i, r) \) such that if \( r \) is the remaining capacity of the knapsack and only type \( i \) items are available then it is optimal to stop if and only if the total value of items currently in the knapsack is at least \( v(i, r) \). We show how to compute these functions and then show that even when all types are available one should never put a type \( i \) item in the knapsack if the remaining capacity is \( r \) and the value of the knapsack is at least \( v(i, r) \). Using the derived formula for the optimal expected return function when only a single type is available, we are able, after much analysis, to characterize and then use this characterization to obtain an algorithm for determining the optimal policy when there are 2 types of items. Because the optimal expected return function in this case of 2 types is not analytically tractable we are not able to then find the optimal policy when there are more than 2 types. However, using intuition garnered from our studies we are able to suggest a heuristic policy in that case which, as indicated by numerical work, yields very good results.

The paper [7], a variation of [6], assumes that the knapsack capacity is not fixed but rather is random with an exponential distribution. An algorithm that can be used to efficiently find the optimal policy is given. Other variations of this random exponentially distributed capacity model are also
considered. For instance, we show if an item’s reward is a deterministic function of its type and its weight is exponentially distributed then the problem reduces to one of the static stochastic knapsack models considered in [5].

In the preceding models it is assumed that the items are all initially present and one must continually decide either to stop or to choose the type of item to put in the knapsack. In [8], however, we suppose that items of differing types arrive sequentially, with the successive types of arriving items constituting a Markov chain with specified transition probability matrix. Upon an item’s arrival, its type is learned and a decision must be made as to whether to put the item in the knapsack or to reject it. If we reject it, then the problem ends with a return equal to the sum of the values currently in the knapsack. If the item is accepted and so put in the knapsack, we learn its reward and whether the knapsack is broken. If the knapsack is broken the the problem ends with 0 return. If not, then we can either pay a fixed cost to observe the type of the next item or we can stop with a reward equal to the sum of the values currently in the knapsack. Assuming that the knapsack has an exponentially distributed capacity, we prove that the structure of the optimal policy is that there are values r(i) such that the optimal policy is to reject a type i item if and only if the total reward of items currently in the knapsack is at least r(i). We first show how to identify and determine the largest of these critical values, and then show how to use that to determine a search algorithm to find the optimal values r(i). This latter is done first when the cost to observe the next type item is 0 and then, using results from the preceding case, when it is positive.

Significance The component failure models we have introduced are of importance because they can reasonably be used to model a variety of situations that result in component lifetimes being dependent. The simulation methods we have developed are quite general and so can be used in a multitude of possible applications. Our study of stochastic knapsack models expands the applicability of this well known problem to situations where there is randomness in an item’s weight or in its value. Our analysis is novel in that it uses a combination of dynamic programming, efficient simulation, and probabilistic analysis to obtain algorithms for either finding, or when that is not possible, approximating optimal policies. Our results should be helpful to the US Army in analyzing existing systems and in making allocation decisions.
Future Plans We plan to continue our study of system models where component failures are caused by randomly occurring shocks of various types. We will consider both more general models where a shock can effect more than just a single component and less general ones where the structure of the system is given. For instance, in the latter case we will be interested in seeing if we can derive new analytic results for parallel systems that fail when there have been at least \( n(i) \) type \( i \) shocks for all \( i \). We plan to consider other types of optimization problems that should be of interest to the Army such as making optimal decisions in problems where the “state of the system” is only probabilistically known. Another problem of interest is one in which one must assign randomly appearing assets (say organs needed for transplants) among a group of individuals when the payoff of an assignment depends on both the asset and individual and where unassigned individuals can be lost.
Grant Supported Papers


2. A System in which Component Failures Increase the Random Hazards of Still Working Components, S. Ross, technical report


5. Static Stochastic Knapsack Problems, K. Chen and S. Ross, accepted for publication *Prob. in Eng. and Inf. Sci.*


