



APLICAÇÕES COMPUTACIONAIS NO ENSINO DE PROBABILIDADE E ESTATÍSTICA EM ENGENHARIA

Júlio César Figueiredo – j.cezar@aol.com

Cibelly Araújo de Azevedo Lima – cibellyazevedo@gmail.com

Francisco Rafael Marques Lima – frmlrafael@gmail.com

Universidade Federal do Ceará

Rua Coronel Estanislau Frota, S/N

62.010-560 – Sobral – CE

Resumo: Há uma busca constante por metodologias didáticas a fim de motivar os estudantes de Engenharia da Computação, particularmente no início do curso, quando a evasão é maior. Na disciplina de Probabilidade e Estatística, são apresentadas aplicações computacionais da teoria estudada usando o software matemático/científico Scilab. Mais especificamente, é apresentado e utilizado o método de Monte Carlo, que se baseia na teoria das probabilidades, para calcular medidas de sistemas estocásticos modelados em computador. Os objetivos são despertar interesse nos estudantes, exemplificando a importância da disciplina em problemas de Engenharia que envolvem incerteza, introduzir os alunos ao software matemático/científico, ferramenta requerida por várias disciplinas no decorrer do curso e para o trabalho profissional, e mostrar como problemas podem ser solucionados por análise numérica com auxílio de computador.

Palavras-chave: Probabilidade, Estatística, Monte Carlo

1. JUSTIFICATIVA E OBJETIVO

A evasão de estudantes nos primeiros semestres é uma realidade presente na grande maioria dos cursos de engenharia do país. De acordo com a Confederação Nacional da Indústria (CNI), a média de evasão de estudantes de engenharia na última década foi de 55,5% no Brasil (MONACO, 2014). Dentre os diversos motivos para o abandono do curso por parte dos estudantes podemos citar a desmotivação devido à grande dificuldade observada em acompanhar as disciplinas de base matemática dos primeiros semestres. Além dessa dificuldade, a desmotivação é causada pela falta de associação entre a teoria ensinada em sala de aula e as aplicações práticas em engenharia.

Neste trabalho, apresentamos um exemplo de como associar os conhecimentos teóricos abordados na disciplina de Probabilidade e Estatística do segundo semestre de Engenharia da Computação da Universidade Federal do Ceará no *campus* Sobral a alguns aspectos práticos que serão vistos ao longo do curso. Essa associação foi feita através do uso constante de *softwares* matemáticos durante o ensino da teoria abordada ao longo do semestre. O *software* utilizado é o Scilab (<http://www.scilab.org>), que pode ser obtido gratuitamente.



Em especial, descreveremos neste trabalho um exemplo do uso da técnica de simulação de sistemas estocásticos chamada Monte Carlo em dois problemas encontrados em Telecomunicações que consiste em uma das vertentes do curso de Engenharia da Computação: sistemas de filas e sistemas de acesso múltiplo. Esses dois problemas possuem algo em comum: ambos são passíveis de serem resolvidos analiticamente utilizando conceitos de probabilidades e variáveis aleatórias. Portanto, não seria necessário o uso de simulações de Monte Carlo para resolvê-lo, pois esta geralmente é aplicada em sistemas muito complexos, não passíveis de soluções analíticas. Contudo, acreditamos que o confronto entre as soluções analíticas e as obtidas pela técnica de Monte Carlo consiste em uma ferramenta didática em que o aluno poderá confirmar a efetividade das simulações de Monte Carlo na resolução de problemas em engenharia. É importante ressaltar que a técnica de Monte Carlo pode ser facilmente compreendida por estudantes que possuam um nível básico de programação e que tenham noção de geração de variáveis aleatórias. Esses dois requisitos são naturalmente obedecidos pelos alunos de Probabilidade e Estatística.

O objetivo dessas aplicações é demonstrar aos estudantes da disciplina em questão, uma aplicação prática dos conceitos abordados em sala de aula. Esperamos que isso possa atenuar os problemas mencionados anteriormente, diminuindo, assim, a evasão nos primeiros semestres de engenharia.

O restante deste trabalho está organizado da seguinte forma: na seção 2 apresentamos uma descrição mais detalhada do método de Monte Carlo, assim como sua aplicação nos sistemas de filas e sistemas de acesso múltiplo. Nesta seção também confrontamos os resultados teóricos e simulados confirmando a validade dos resultados. Por fim, concluímos este trabalho na seção 3 resumando os principais resultados diretos e indiretos obtidos com esse trabalho.

2. MODELOS EMPREGADOS E RESULTADOS

2.1. Método de Monte Carlo

O método de Monte Carlo é uma técnica para resolver problemas de natureza aleatória. Realizada em computador, uma simulação de Monte Carlo é uma sequência de experimentos aleatórios. Enquanto um único experimento diz quase nada sobre uma medida de um sistema em simulação, uma quantidade relativamente grande de resultados de experimentos aleatórios pode ser bastante informativa, baseando-se na noção de probabilidade como limite da frequência relativa. Assim, quanto maior o número de experimentos numa simulação, mais tempo leva para ser finalizada e maior a precisão dos resultados.

É utilizado o *software* livre Scilab para codificar os modelos, realizar as simulações, e sintetizar os resultados. Os códigos são apresentados aos estudantes, que devem ser capazes de compreendê-los com a sintaxe que foi aprendida nas aulas regulares e com o conhecimento prévio de programação.

Na aplicação do método de Monte Carlo, um modelo matemático de um sistema real é concebido de modo que represente satisfatoriamente o seu comportamento. Um modelo computacional é então escrito para implementar o modelo matemático. Experimentos repetidos são feitos com o modelo computacional e uma fonte de dados aleatórios (como um gerador de números pseudoaleatórios). Os resultados são coligidos para análise, uma vez que cada resultado isoladamente informa quase nada. A Figura 1 ilustra esses passos.

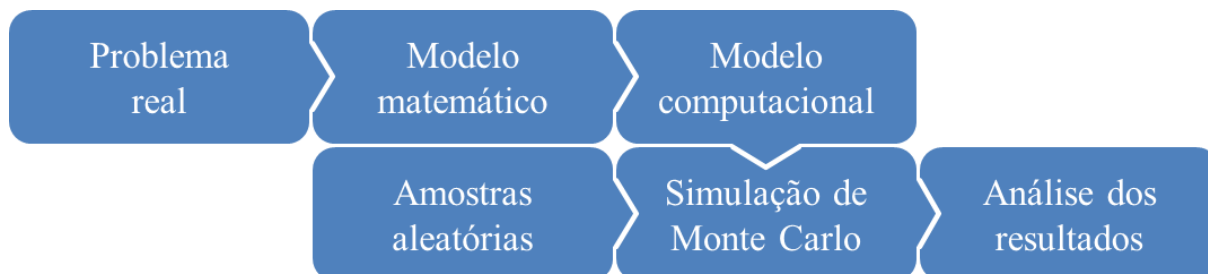


Figura 1: fluxograma de uma simulação de Monte Carlo.

2.2. Sistemas de filas

Um sistema de filas contém clientes e servidores. Clientes chegam e solicitam um serviço provido pelos servidores. Aqui, apenas um cliente pode ser atendido por vez por cada servidor. Se um cliente chega e não encontra um servidor vazio, ele pode esperar numa fila até que um servidor fique disponível, ou deixa o sistema imediatamente, caso que é chamado bloqueio. Depois de ser atendido, o cliente deixa o servidor livre para outros clientes (LAW & KELTON, 2000). A Figura 2 ilustra esse funcionamento.

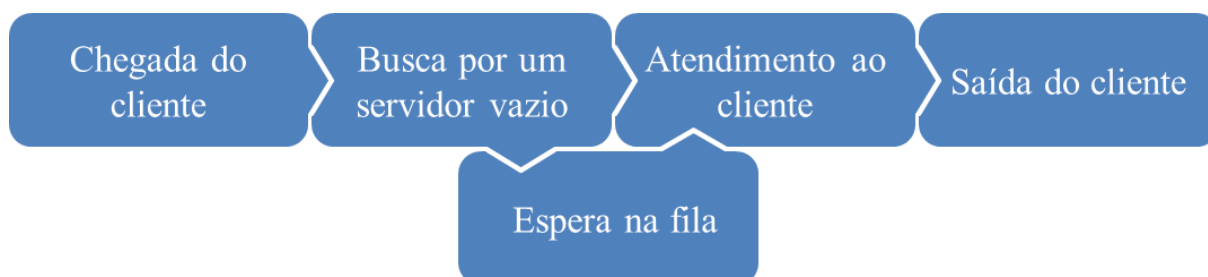


Figura 2: fluxograma de um sistema de filas.

Esse simples modelo está presente, por exemplo, em sistemas de comunicações celulares em que o servidor consiste em uma estação rádio-base e os clientes são pedidos de chamadas telefônicas que são originadas em tempos aleatórios e que, para serem executadas, necessitam do uso de recursos de transmissão da estação rádio-base. Portanto, o estudo desse problema fornece informações importantes para o projeto de sistemas de comunicações.

A chegada de clientes é modelada como um processo estocástico de Poisson. Isso significa que a chegada de um cliente é independente da chegada dos clientes anteriores, e que a probabilidade de mais do que um cliente chegar num intervalo diferencial de tempo é zero. Ainda, isso implica que a duração entre as chegadas dos usuários é uma variável aleatória com distribuição de probabilidade exponencial com média I . A duração do atendimento também é modelada com distribuição exponencial com média H . A fila é modelada com capacidade infinita, sem limite de comprimento. Portanto, a aleatoriedade nesse modelo está nos instantes de tempo de chegada dos clientes e na duração dos atendimentos nos servidores.

A partir das hipóteses acima, pode-se chegar analiticamente às fórmulas para a probabilidade de um cliente ter que esperar na fila, que a partir daqui será chamada probabilidade de espera, e para o tempo médio de espera na fila. Nas equações a seguir, C é o

número de servidores, $A = \frac{H}{CI}$ é a intensidade de tráfego, P é a probabilidade de espera, e D é o tempo médio de espera (LAW & KELTON, 2000):

$$P = \frac{A^C}{A^C + C! \left(1 - \frac{A}{C}\right) \sum_{k=0}^{C-1} \frac{A^k}{k!}} \quad (1)$$

$$D = \frac{A^C}{A^C + C! \left(1 - \frac{A}{C}\right) \sum_{k=0}^{C-1} \frac{A^k}{k!}} \cdot \frac{H}{C - A} \quad (2)$$

Os gráficos para essas duas medidas de qualidade de serviço em função da taxa de chegada de usuários $\frac{1}{I}$ são mostrados na Figura 3 e na Figura 4. Para as curvas teóricas exatas foram usadas a Equação (1) e a Equação (2). Para cada curva simulada aproximada, foi realizada, no Scilab, uma simulação de Monte Carlo com 50000 iterações com o modelo descrito anteriormente.

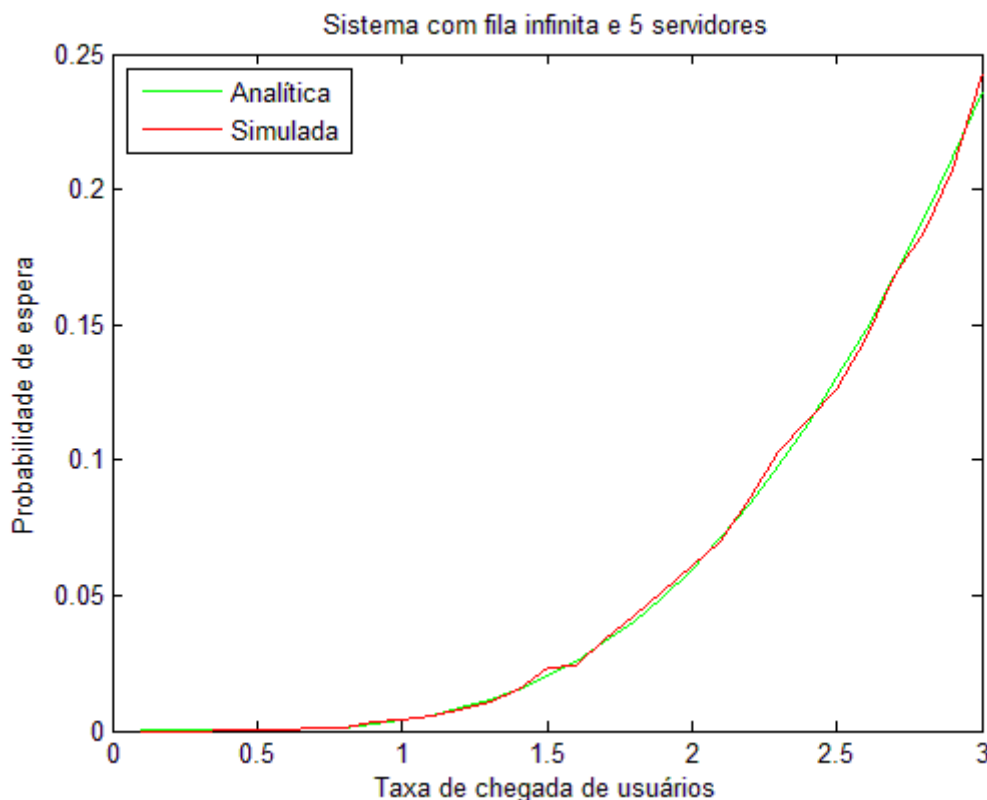


Figura 3: gráfico com curvas teórica e simulada da probabilidade de espera em função da intensidade de tráfego.

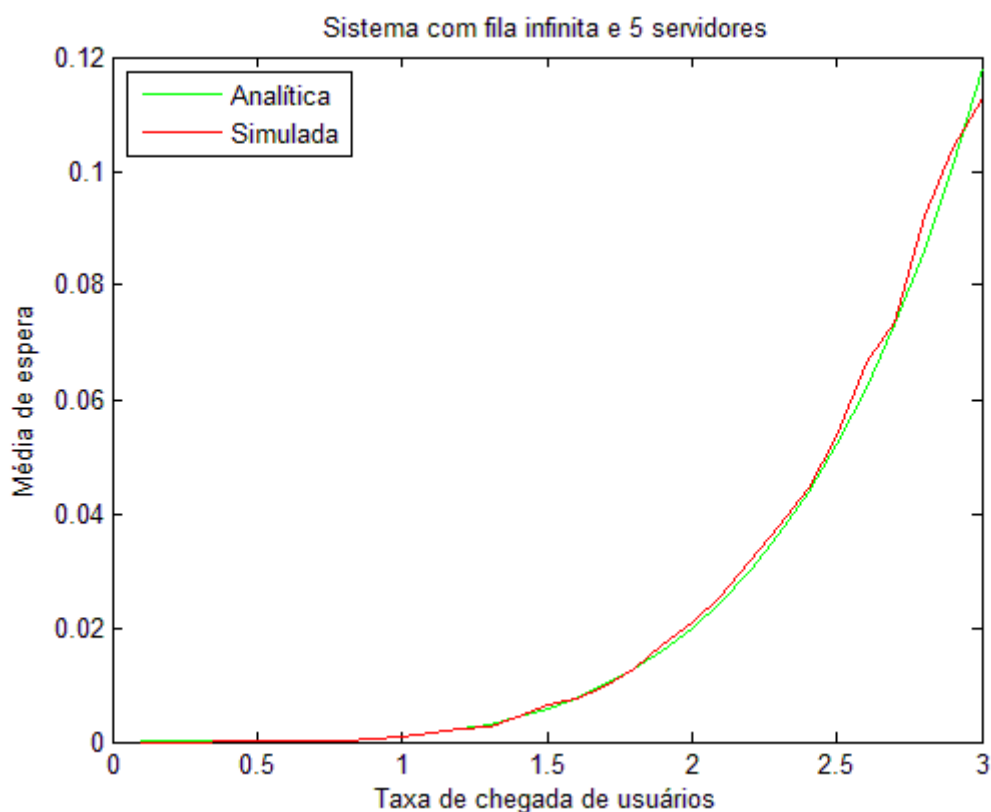


Figura 4: gráfico com curvas teórica e simulada do tempo médio de espera em função da intensidade de tráfego.

Podemos observar nestes resultados que as curvas teóricas e simuladas coincidem dentro de um limite de tolerância aceitável. É importante ressaltar que as curvas teóricas e simuladas podem ficar tão próximas quanto queiramos, bastando para isso aumentar o número de iterações na simulação de Monte Carlo.

2.3. Sistemas de acesso múltiplo

Num sistema de acesso múltiplo, vários terminais de acesso compartilham um recurso. Aqui, o recurso é um canal de comunicação que os terminais utilizam para transmitir dados na forma de pacotes. Quando mais do que um terminal tenta acessar o canal comum ao mesmo tempo, acontece um conflito de acesso e nenhum deles consegue transmitir seu pacote de dados (RAPPAPORT, 2002). Daí a necessidade de um protocolo de acesso múltiplo para resolver (e talvez também evitar) conflitos de acesso e buscar uma utilização eficiente do canal com distribuição justa entre os terminais.

Num protocolo sem competição, o recurso é alocado para um terminal por vez. Assim, os terminais sabem quando podem e quando não podem transmitir. Num protocolo com competição, um terminal nunca tem certeza sobre o resultado de uma transmissão antes de tentá-la. O protocolo sem competição mais simples é o *pure ALOHA*, que tem uma variante chamada *slotted ALOHA* (HARADA & PRASAD, 2002).

No protocolo *pure* ALOHA, os terminais simplesmente podem tentar acessar o canal a qualquer momento. No protocolo *slotted* ALOHA, os terminais só podem tentar acessar o canal no início de um *slot* de tempo. Um *slot* é um intervalo com duração igual ao tempo que os terminais levam para transmitir um pacote. Isso reduz a frequência de conflitos de acesso para metade da frequência do *pure* ALOHA (HARADA & PRASAD, 2002). Em ambos os protocolos, no caso de conflito de acesso, cada terminal adia sua transmissão por um intervalo de tempo aleatório. A Figura 5 e a Figura 6 ilustram a transmissão de pacotes de três terminais para um ponto de acesso.

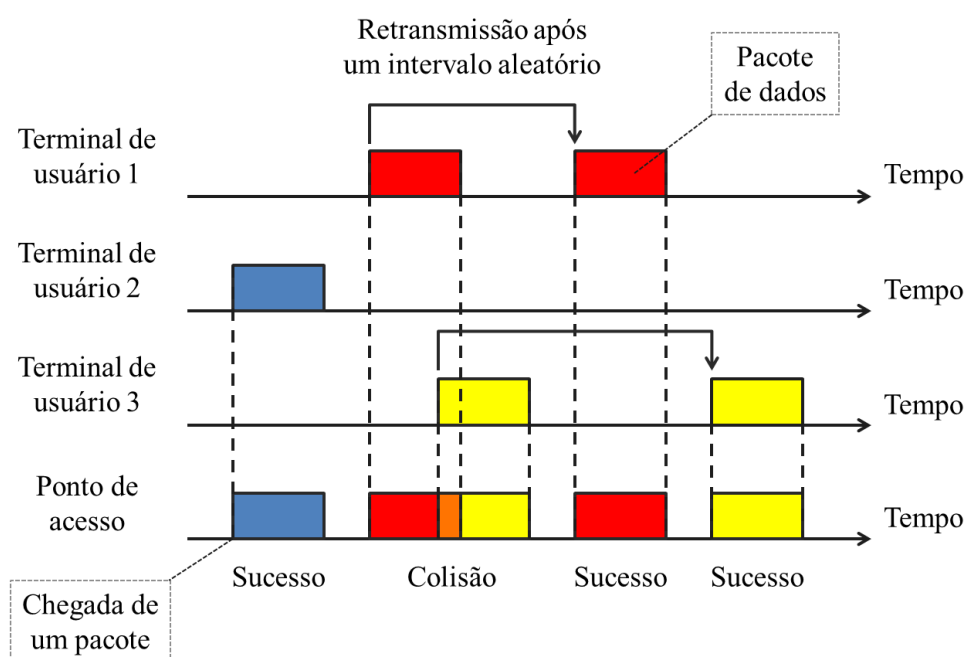


Figura 5: Transmissão de pacotes de dados sob o protocolo *pure* ALOHA.

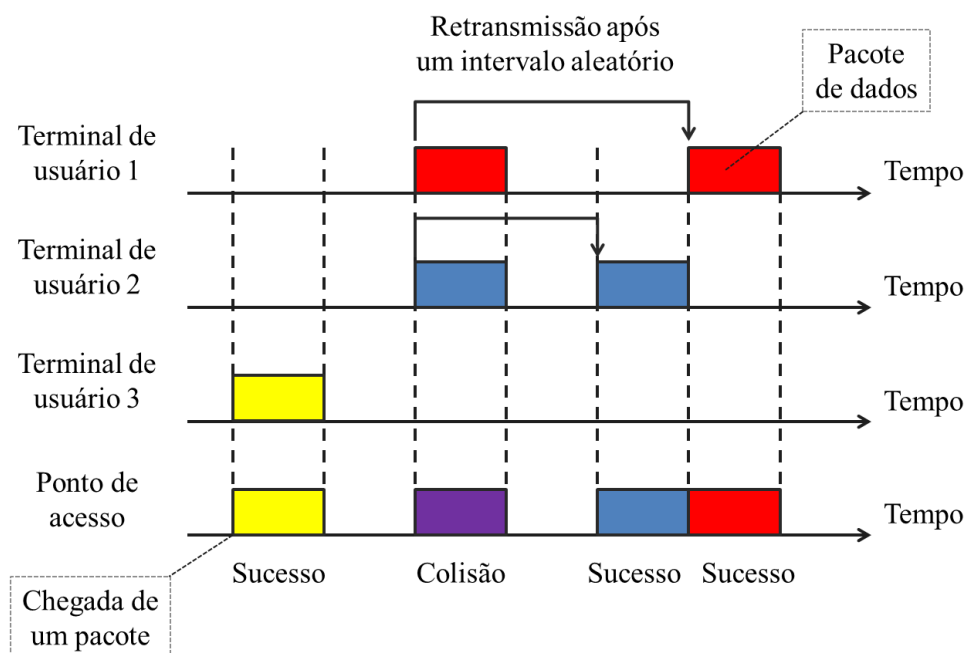


Figura 6: Transmissão de pacotes de dados sob o protocolo *slotted* ALOHA.

O tráfego oferecido G num intervalo de tempo t é a razão entre a quantidade de pacotes que os terminais de acesso tentaram transmitir T_t durante t e a quantidade máxima de pacotes que pode ser transmitida durante t , que é o produto da taxa de transmissão de dados R do canal compartilhado pelo intervalo t :

$$G = \frac{T_t}{Rt} \quad (3)$$

O tráfego oferecido pode ser maior do que 1.

O *throughput* é a principal medida da eficiência de um protocolo de acesso múltiplo. O *throughput* S num dado intervalo de tempo t é definido pela seguinte fórmula, em que S_t é a quantidade de pacotes que foram transmitidos com sucesso durante t e R é a taxa de transmissão de dados do canal compartilhado:

$$S = \frac{S_t}{Rt} \quad (4)$$

O *throughput* é não maior do que 1.

Num protocolo ideal:

$$S = \begin{cases} G & (G < 1) \\ 1 & (G \geq 1) \end{cases} \quad (5)$$

Mas a frequência de colisões aumenta com o aumento de G , e o *throughput* passa a decrescer a partir de um limiar de *throughput* máximo (HARADA & PRASAD, 2002). Esse

comportamento não só é previsto, mas também é verificado pelas fórmulas obtidas analiticamente.

O *throughput* num sistema de acesso múltiplo com protocolo *pure* ALOHA é:

$$S = Ge^{-2G} \tag{6}$$

O valor máximo dessa função é $S = 0,184$ em $G = 0,5$.

A Figura 7 exhibe a curva da Equação (6) e uma outra curva resultante de uma simulação de Monte Carlo com um modelo de sistema de acesso múltiplo com protocolo *pure* ALOHA.

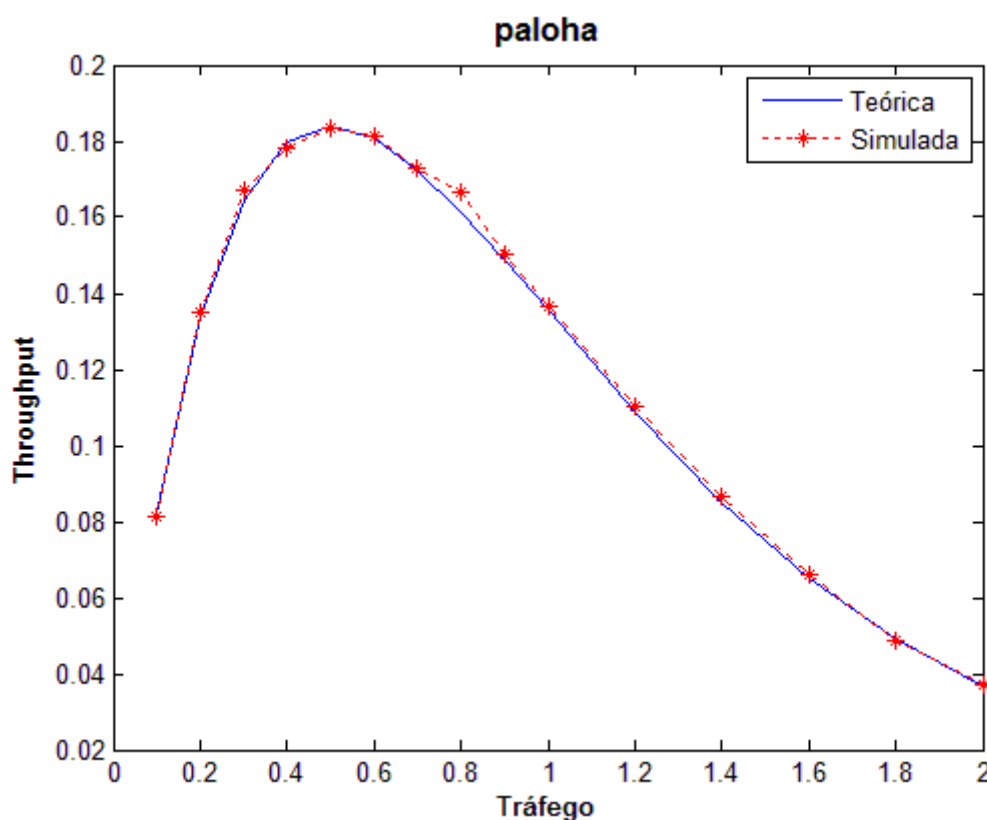


Figura 7: gráfico com curvas teórica e simulada de *throughput* do protocolo *pure* ALOHA em função do tráfego.

O *throughput* num sistema de acesso múltiplo com protocolo *slotted* ALOHA é:

$$S = Ge^{-G} \tag{7}$$

O valor máximo dessa função é $S = 0,368$ em $G = 1$.

A Figura 8 exhibe a curva da Equação (7) e uma outra curva resultante de uma simulação de Monte Carlo com um modelo de sistema de acesso múltiplo com protocolo *slotted* ALOHA.

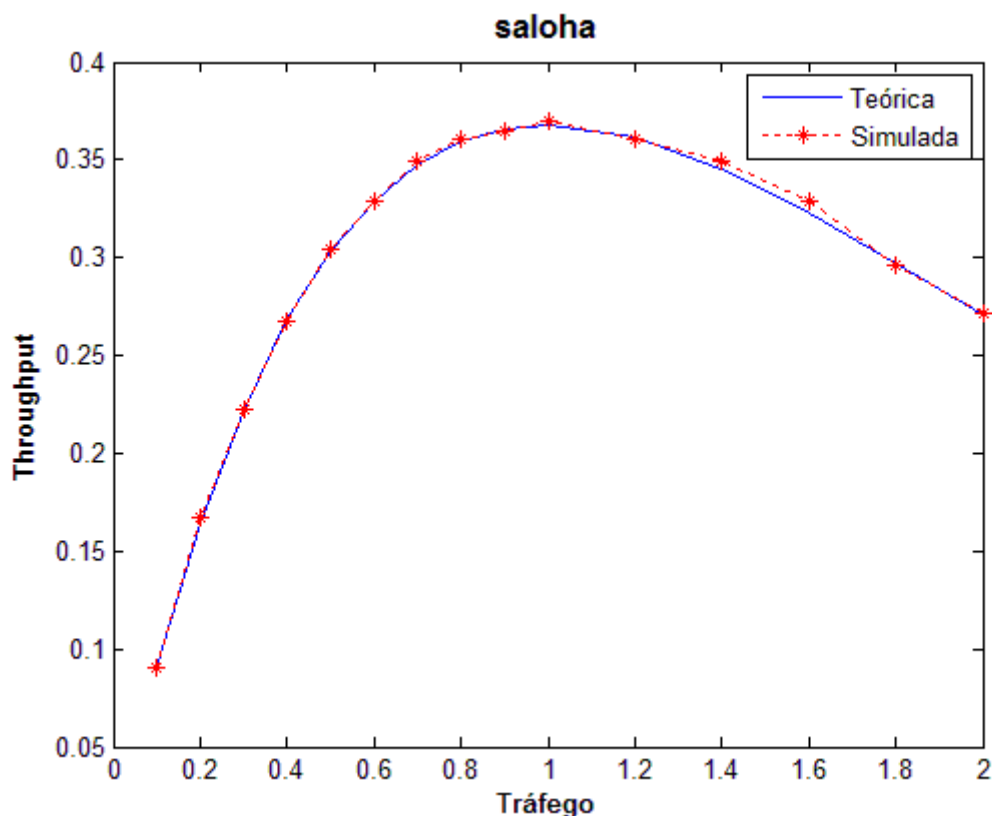


Figura 8: gráfico com curvas teórica e simulada de *throughput* do protocolo *slotted* ALOHA em função do tráfego.

Com uma pequena tolerância, coincidem as curvas teórica e simulada mostradas na Figura 7 e na Figura 8, confirmando a validade do modelo. E podem ficar tão próximas quanto se queira, bastando aumentar o número de iterações na simulação.

3. RESULTADO E CONCLUSÕES

Neste trabalho buscamos formas de motivar os alunos da disciplina de Probabilidade e Estatística através de aplicações computacionais da teoria vista em sala de aula. Especificamente, aplicamos os conceitos vistos na disciplina para aplicação do método de Monte Carlo, que consiste em uma ferramenta de simulação estocástica utilizada para obter respostas de problemas em sistemas muito complexos que não são passíveis de serem resolvidos de forma analítica. Contudo, para fins didáticos, escolhemos dois problemas simples na área de telecomunicações que possuem soluções analíticas: sistemas de filas e sistemas de acesso múltiplo. Os resultados presentes neste artigo demonstram a eficácia da técnica de Monte Carlo que proporcionou resultados muito próximos da solução analítica.

Além da maior motivação por parte dos estudantes da disciplina, que tiveram contato com ferramentas sofisticadas de engenharia logo no segundo semestre do curso, tivemos



também benefícios quanto a formação do bolsista autor deste trabalho. O trabalho direto com o docente responsável pela disciplina propiciou uma maior capacitação técnica do bolsista assim como um primeiro contato com práticas docentes, contribuindo assim para formação de recursos humanos na área.

REFERÊNCIAS BIBLIOGRÁFICAS

HARADA, H.; PRASAD, R. Simulation and Software Radio for Mobile Communications. Artech House, 2002.

LAW, A. M.; KELTON, W. D. Simulation Modeling and Analysis. McGraw Hill, 2000.

MONACO, R. Reportagem: Mais da Metade dos Estudantes Abandona Cursos de Engenharia. Disponível em: <http://www.portaldaindustria.com.br/cni/imprensa/2013/07/1,19276/mais-da-metade-dos-estudantes-abandona-cursos-de-engenharia.html>. Acessado em 20/05/2014.

RAPPAPORT, T. S. Wireless Communications: Principles and Practice. Prentice Hall, 2002.

COMPUTER APPLICATIONS IN TEACHING OF PROBABILITY AND STATISTICS

Abstract: *There is a constant seek for didactic methods aiming to motivate students in Computer Engineering, particularly in the beginning of the course, when most of the evasion of students occurs. In the course of Probability and Statistics, some computer applications of the theory subject are presented to the students by using the scientific software Scilab. More precisely, the Monte Carlo method, a method based on probability theory, is introduced and used to calculate measures about computer models of stochastic systems. The goals are to arise interest of the students in engineering by presenting practical problems that involve uncertainty, to introduce them to scientific software as an essential tool through the entire course and for the professional work, and to show them how problems can be solved through numerical analysis with the aid of computers.*

Key-words: *Probability, Statistics, Monte Carlo*