

10

Conclusões e Trabalhos futuros

Este capítulo apresenta as conclusões desta tese e sugere questões para trabalhos futuros.

10.1

Conclusões

Uma das principais mudanças que vivemos ao longo da última década foi o crescimento da telefonia móvel. Os aparelhos celulares estão menores e mais baratos e a sociedade incentiva a conexão permanente em qualquer lugar, a qualquer hora. No entanto, os recursos da rede estão baseados no espectro de frequências que é limitado.

Além disto, o emprego dos recursos das redes móveis para o envio de dados, além da aplicação original que era a comunicação via voz, modifica o perfil do tráfego na rede. Sendo assim, o centro de interesse desta tese foi apresentar soluções para que se possa utilizar bem os recursos da rede e manter a qualidade do serviço prestado.

O controle de admissão de chamadas é um elemento-chave da rede que é responsável por aceitar ou não uma chamada de acordo com sua demanda de recursos; sua classe de tarifação, seus requisitos de qualidade de serviço (QoS) e se o usuário é um usuário novo tentando ser aceito no sistema, ou se está em procedimento de *handoff*, proveniente de outra célula. Diante de tal importância, optou-se por desenvolver este trabalho na área de controle de admissão de chamadas.

Com o objetivo de atingir os graus de qualidade de serviço desejados para cada uma das classes de usuário existente no sistema, são analisadas técnicas que utilizam o conceito de reserva de recursos, também conhecido por uso de canais de guarda. Resultados analíticos para desempenho dos algoritmos que utilizam um número fixo de recursos reservados são apresentados empregando uma modelagem por Cadeia de Markov M-dimensional.

Os recursos a serem distribuídos entre os usuários são definidos como unidades de banda (UB) e podem representar os recursos dos vários sistemas existentes. Os recursos nos sistemas FDMA são canais ou frequências, nos sistemas TDMA são divisões de tempo, nos sistemas CDMA são códigos e níveis de potência e nos sistemas OFDMA são frequências e divisões de tempo. Portanto, o estudo apresentado neste trabalho pode ser aplicado em diversos sistemas. Os trabalhos recentes [36], [70] e [6] apresentam algoritmos de controle de admissão que podem ser utilizados nos padrões que adotam tecnologia OFDMA que utiliza *slots*, definidos por conjuntos de subportadoras e janelas de tempo em cada uma delas, como recursos a serem alocados recaindo então no problema que é considerado neste trabalho.

Nesta tese foram desenvolvidos algoritmos de Controle de Admissão de Chamadas (CAC) que são adaptativos ao tráfego em andamento no sistema. Cada Estação Rádio Base (ERB) controla a entrada de chamadas ao seu sistema de forma dinâmica, de acordo com as características do tráfego em andamento e do desempenho do sistema. Além de serem dinâmicos, os algoritmos propostos são distribuídos. Não há necessidade de troca de informação entre as Bases para a tomada de decisão da solicitação de recurso rádio, diminuindo o volume de informações de controle na rede.

Os algoritmos propostos consideram M classes de usuários que são subdivididas em $2M$ tipos diferentes de usuários. Para cada classe há uma demanda diferente de recursos da rede e cada classe está dividida entre usuários de chamadas novas e usuários de sessões transferidas de outras células. Os requisitos de qualidade de serviço considerados são os percentuais de bloqueio aceitáveis para chamadas novas e para chamadas em *handoff*. Para um total de M classes, são $2M$ requisitos de QoS. Inicialmente, foi considerado o caso $M=1$, onde uma classe de usuário é dividida em duas: a classe das chamadas novas e a classe das chamadas de usuários em *handoff*. O desenvolvimento analítico, que já havia sido proposto anteriormente, foi revisto no capítulo 3. É importante observar que, ao se considerar dois requisitos de qualidade de serviço independentes, um para o bloqueio de chamadas novas Q_{BN} e outro para a falha de chamadas em *handoff* Q_{FH} , o caso $M=1$, na realidade, equivale a duas classes de usuários.

Também é parte integrante deste trabalho o desenvolvimento de um simulador de um sistema de comunicações móveis, em linguagem C, para verificar o desempenho dos algoritmos aqui propostos. O simulador comporta uma ou mais classes de usuário com valores diferentes de demanda de recursos do sistema e requisitos de qualidade de serviço. Os detalhes do simulador estão no capítulo 4. Seus resultados são confirmados com

resultados analíticos da probabilidade de falha de chamadas em *handoff*, da probabilidade de bloqueio de chamadas novas e da utilização dos recursos rádio, para sistemas cujo controle de admissão é estático.

Os seis primeiros algoritmos propostos são apresentados no capítulo 5. Estes algoritmos possuem objetivos diferentes entre si e consideram uma classe de usuários dividida entre usuários novos e em *handoff*. Seus requisitos de qualidade de serviço são diferentes e são denominados de vetor de QoS $Q\% = (Q_{BN}; Q_{FH})$. Dentre estes algoritmos, apenas o CAC5 não utiliza os canais de guarda de uma forma explícita. No entanto, ao admitir o usuário solicitante de acordo com um critério probabilístico, o algoritmo está reservando recursos para outros usuários que venham a solicitá-los mais tarde. Foi provado no capítulo 5 que, de acordo com as condições de mobilidade, o algoritmo CAC5 se porta de forma equivalente a utilização de canais de guarda para usuários em *handoff* ou como canais de reserva para usuários novos.

No capítulo 6 foi desenvolvido um método que permite calcular a intensidade de tráfego máxima à qual o sistema pode ser submetido, atendendo aos valores de qualidade de serviço estabelecidos, no caso $M=1$ e tempo de retenção do recurso rádio modelado por uma chamada com distribuição exponencial. Além da intensidade de tráfego máxima, o método também calcula a quantidade de recursos a ser reservada para diferentes perfis de mobilidade, assumindo que o objetivo é maximizar a utilização do sistema. Foi proposto um algoritmo simples, dinâmico e distribuído, baseado em medidas em tempo real, cuja meta é acompanhar a curva ótima de número de recursos reservados. Observou-se que o algoritmo apresenta um excelente desempenho diante do modelo de tráfego em que o tempo de retenção do recurso do sistema é dado pela distribuição exponencial.

O capítulo 7 compara o desempenho dos sete algoritmos através de três métricas M, F e I definidas nesta tese. A comparação tem o objetivo de ampliar estudo sobre as possibilidades de se atender aos requisitos de qualidade de serviço nos diversos cenários que envolvem a mobilidade dos usuários e a intensidade de tráfego de entrada, considerando diferentes distribuições para o tempo de permanência do usuário na célula que define o tempo de retenção do recurso da rede.

A maioria dos trabalhos encontrados na bibliografia utiliza a distribuição exponencial para o tempo de permanência do usuário na célula. Neste trabalho, além da distribuição exponencial, a permanência do usuário na célula também é modelada pelas distribuições uniforme, constante e Weibull, apresentando o desenvolvimento analítico para o cálculo do tempo

de retenção e o comportamento simulado da rede. Observou-se no capítulo 7 que quase não há diferença entre os resultados obtidos com a utilização da distribuição exponencial e os resultados obtidos através das distribuições uniforme, constante e Weibull para o tempo de permanência do usuário na célula, principalmente nos casos em que a mobilidade é baixa.

Os cálculos referentes ao tempo de retenção do canal quando a distribuição do tempo de permanência do usuário é dada pelas distribuições constante, uniforme, exponencial e Weibull estão no apêndice A.

Na etapa seguinte são considerados usuários de múltiplas classes, ou seja, os usuários podem entrar no sistema solicitando quantidades diferentes de recursos ou com diferentes requisitos de qualidade de serviço. Os recursos limitados da rede sem fio são alocados pelo CAC aos usuários que solicitam recursos do sistema considerando: sua classe de usuário; se o solicitante é um usuário em *handoff* ou se é um usuário novo; e o desempenho instantâneo do sistema para cada tipo de usuário. O objetivo é atender, de uma forma eficiente, aos diferentes tipos de tráfego, respeitando seus requisitos de QoS.

Através dos resultados apresentados no capítulo 8, observou-se que uma alteração em cada uma das restrições de uma classe altera o valor dos percentuais de bloqueio de todas as classes. Isso se deve ao fato de que as taxas de mensagens admitidas em cada classe são influenciadas pelo valor das restrições estabelecido e qualquer mudança nas taxas altera o valor das probabilidades estacionárias dos estados da cadeia de Markov. Observou-se também que, conforme esperado, o percentual de bloqueio cujo valor é mais afetado é exatamente o correspondente ao requisito alterado. Concluímos que o bloqueio é uma combinação do Vetor Restrição R , da requisição de banda B_γ e do percentual de usuários de cada classe, além das condições do tráfego. O modelo analítico apresentado no capítulo 8 para 2M classes generaliza o modelo do capítulo 3 para o caso $M=1$.

No capítulo 9 foram apresentados quatro algoritmos propostos para M classes de usuários. Alguns dos algoritmos definidos para o caso de $M=1$ são estendidos para o caso de $M=2$. Os algoritmos foram testados no caso em que $M = 2$, portanto lidam com os usuários novos e em *handoff* da classe 1, $(N1, H1)$, e com os usuários novos e em *handoff* da classe 2, $(N2, H2)$. Seu desempenho também é avaliado, utilizando o simulador desenvolvido neste trabalho. O método para calcular a intensidade máxima de recursos que o sistema comporta, sem violar os requisitos de qualidade de serviço, é estendido para o caso de duas classes de usuários, considerando quatro requisitos de qualidades de serviço independentes, $Q\% = (Q_{BN1}, Q_{FH1}, Q_{BN2}, Q_{FH2})$. Da mesma forma que os algoritmos

propostos para uma classe de usuários $M = 1$, apresentados no capítulo 5, todos algoritmos propostos para $M = 2$, no capítulo 9, são dinâmicos e distribuídos. Cada setor registra a entrada, ou não, de determinado tipo de chamada através de janelas de registro deslizantes descritas no capítulo 4.

Finalmente, ainda no capítulo 9, a definição dos parâmetros F e I é estendida para o caso de M classes permitindo comparar o desempenho dos algoritmos com $M > 1$, considerando uma distribuição genérica para o tempo de permanência do usuário na célula.

Um traço específico dos algoritmos propostos nesta tese é a preocupação simultânea com $2M$ requisitos de qualidade de serviço, supondo M classes de usuários (novos e em *handoff*). O desempenho de cada classe de usuário é avaliado segundo as medidas de bloqueio obtidas durante a execução do algoritmo, e os respectivos requisitos de qualidade.

Os algoritmos também tem o seu desempenho avaliado de acordo com a taxa de mobilidade, θ , que obriga o sistema a se adaptar às características do perfil de tráfego em andamento.

Ao aplicar a técnica de reserva de recursos com o objetivo de atender aos requisitos de qualidade de serviço, observou-se que é possível determinar o valor dos recursos a serem reservados de modo a se obter uma utilização máxima da quantidade total de recursos. No entanto, a solução é possível somente até certo valor de intensidade de tráfego de entrada determinado pela taxa de mobilidade, pelas necessidades de banda de cada classe e pelo conjunto de requisitos de QoS. Para valores de intensidade de tráfego superiores, não é mais possível atender aos requisitos de qualidade de serviço. Além disto, existem certos valores atribuídos aos elementos do vetor de requisitos de qualidade $Q\%$, que não permitem solução para o problema da reserva de recursos para cada classe de modo a obter utilização máxima.

Outra conclusão interessante dos resultados apresentados é que o comportamento dos algoritmos não muda muito quando o tempo de permanência do usuário na célula é modelado por distribuições diferentes da exponencial, principalmente no caso em que a taxa de mobilidade é alta. No caso em que a taxa de mobilidade é baixa a distribuição constante para o tempo de permanência provoca bloqueios maiores dos usuários.

10.2

Trabalhos Futuros

Apresentamos como sugestões para trabalhos futuros em extensão a este trabalho:

Trabalhar com uma quantidade maior de vetores de QoS para identificar se há alguma lei de formação para a reserva de recursos que torna a utilização máxima no caso de retenção do recurso dada pela distribuição exponencial.

Considerar outras distribuições no tempo de duração da chamada. Tal procedimento pode considerar que as classes de usuários possuem distribuições diferentes nos seus intervalos de duração da chamada. Por exemplo, uma classe pode representar usuários de voz através da distribuição exponencial e outra classe pode representar usuários de dados modelados por uma distribuição de Pareto para o tempo de duração da chamada.

Considerar uma rede com mais de uma célula no simulador de forma a incluir outros elementos da gerência de recursos rádio tais como o controle de potência e a alocação dinâmica.

Aperfeiçoar o simulador de forma que seja possível variar um número maior de parâmetros do sistema.

10.3

Publicações

Alguns dos resultados apresentados neste trabalho de tese originaram publicações que constam das referências [63], [64], [65], [66].