

2 Trabalhos Relacionados

O problema do controle de admissão de usuários em redes sem fio possui múltiplas facetas. Enquanto que em uma rede fixa o controle de admissão se resume ao tempo de duração da chamada, na rede móvel o problema se estende devido à mobilidade do usuário que necessita trocar seu ponto de acesso à rede. A mobilidade e o tempo de retenção do canal provocam alto impacto no desempenho do sistema e devem ser considerados pelo algoritmo de controle de admissão. Desta forma, dentro da Gerência dos Recursos Rádio, o Controle de Admissão de Chamadas (CAC) é um dos fatores principais para o bom desempenho do sistema.

O número de usuários de celulares continua aumentando, fazendo com que as operadoras trabalhem com células cada vez menores para aumentar a capacidade do sistema. Por outro lado, células menores aumentam o número de *handoffs*, gerando a necessidade de algoritmos de controle de admissão cada vez melhores. Além disso, as técnicas de CAC constituem um dos fatores que permite ao operador garantir a qualidade de serviço (QoS) desejada e obter recursos adicionais através da diferenciação entre os serviços e o aumento de tráfego.

A reserva de recursos e, em particular, o controle de admissão de chamadas são temas importantes para a operação eficiente de redes sem fio. Este problema tem sido estudado por duas décadas [31], porém o tema continua relevante, conforme demonstrado em dois artigos recentes que apresentam tutoriais extensos sobre o tema [4], [20].

Com a evolução das redes móveis e a oferta de serviços diferentes dentro da mesma rede, o Controle de Admissão de Chamadas (CAC) passou a ter uma importância ainda mais elevada dentro da Gerência de Recursos Rádio.

O problema geral da gerência de recursos é tratado nas referências [77] e [78]. Estes trabalhos apresentam a escolha da forma de onda da seguinte forma: os sistemas FDMA e TDMA utilizam sinais ortogonais, possibilitando a canalização do espectro e os sistemas CDMA utilizam sinais

não ortogonais. Esta divisão dos sistemas em FDMA/TDMA e CDMA torna suas abordagens bem distintas. Os sistemas que trabalham com o espectro canalizado podem ser mais eficientes quanto à alocação dos canais.

Em [37] são apresentados diversos aspectos quanto à gerência de recursos em sistemas CDMA de terceira geração. O trabalho também trata da solução híbrida TD/CDMA onde, para cada slot de tempo, são associados vários usuários, cada um com um código. Devido à flexibilidade do sistema de divisão por tempo, o sistema pode ser empregado em vários ambientes e adicionando o código, é possível aumentar a capacidade do sistema.

Nas sessões seguintes deste capítulo são abordados trabalhos sobre as subdivisões da Gerência de Recursos Rádio. Inicialmente são abordados a Alocação de Canais, o Controle de Potência e as Células Hierarquizadas. Por último, o capítulo apresenta uma abordagem mais detalhada do Controle de Admissão de Chamadas, tema desta tese.

2.1

Alocação de Canais

Nos primeiros sistemas celulares, FDMA, a alocação de canais era fixa, FCA - *Fixed Channel Allocation*, determinados canais poderiam ser utilizados apenas em determinadas células. No entanto, esses sistemas necessitam de um planejamento inicial que não deve ser modificado. Com o aumento da demanda por comunicações móveis, houve necessidade de aumentar a capacidade dos sistemas. Os sistemas fixos eram muito rígidos para aceitar as modificações que eram constantemente requeridas.

Como uma evolução, surgiram os sistemas de alocação dinâmica de canais, DCA - *Dynamic Channel Allocation*. Nesses sistemas dinâmicos, qualquer canal pode ser utilizado em qualquer célula, necessitando apenas de um teste inicial, respeitando algum critério para a escolha do canal. Um bom resumo das técnicas de alocação de canais pode ser encontrado em [40].

Em [58] são apresentadas diversas técnicas de alocação dinâmica de canais. São técnicas válidas para sistemas FDMA/ TDMA, onde os canais são ortogonais, ou seja, utilizam frequências diferentes e os *slots* de tempo são independentes. Como nos sistemas de alocação dinâmica todas as bases podem utilizar qualquer canal, as questões que surgem são: Existe um canal melhor? Como otimizar esta alocação? Há um compromisso na escolha do canal: para maximizar a qualidade os canais em reuso devem estar longe, porém, para aumentar a eficiência espectral os canais devem estar o mais

perto possível. Em [58] são apresentadas três técnicas de alocação dinâmica que são comparadas entre si e em relação à alocação fixa.

Os sistemas híbridos se apresentam como uma solução para o aumento de capacidade do sistema TDMA, associando códigos que permitem que mais de um usuário utilize o mesmo *slot* de tempo dentro da mesma frequência. Simultaneamente apresenta ortogonalidade entre os *slots*, permitindo um tratamento de canalização dos recursos. Em [61] é apresentado um sistema dinâmico de alocação de recursos para o sistema Híbrido TDMA/CDMA.

2.2

Controle de Potência

Outro aspecto importante da gerência de recursos é o controle de potência [77][78]. Controlando as potências com que os móveis e as bases irradiam é possível melhorar a qualidade de serviço e aumentar a capacidade, suprimindo a interferência do canal adjacente (correlação cruzada) em esquemas não-ortogonais e controlando a interferência co-canal nos esquemas ortogonais. Além disso, o controle de potência minimiza o consumo, aumentando o tempo de vida da bateria.

2.3

Células Hierarquizadas

A superposição de sistemas que utilizam células de tamanhos diferentes é uma abordagem da gerência de recursos para aumentar a utilização. Os sistemas mais comuns são micro-células superpostas por macro-células.

O objetivo dessa estratégia é obter simultaneamente as vantagens dos sistemas micro e macro, dependendo da velocidade do usuário. Nos sistemas micro-celulares, como a distância de reutilização é menor, a capacidade do sistema é maior. No entanto, neste tipo de sistema micro-celular, a movimentação dos usuários através de células menores aumenta a quantidade de *handoffs* solicitados. Daí a necessidade da superposição das macro-células, para que os veículos sejam alocados nas macro-células e os pedestres nas micro-células.

Em [45] é apresentada uma revisão sobre os trabalhos na área. São apresentadas duas classes de sistemas hierárquicos: reversíveis e não-reversíveis. Os reversíveis supõem que possa ocorrer *handoff* de volta à camada mais baixa e nos não-reversíveis este retorno não é possível.

Embora o objetivo inicial seja a alocação dos usuários de alta velocidade na camada superior, esta camada também pode servir como reserva de recursos. Ou seja, se as micro-células estiverem lotadas o usuário, sendo pedestre ou veículo, será alocado na macro-célula correspondente. Mais tarde, vagando canal na micro-célula, no caso dos sistemas reversíveis, o usuário poderá retornar a camada inferior. Os sistemas reversíveis necessitam de um maior processamento para lembrar qual usuário pedestre está alocado na macro-célula e fazê-lo retornar assim que possível.

A pergunta inicial para esta estratégia de *handoffs* entre camadas é como saber se o usuário é de alta velocidade ou não. Esta é respondida pela observação do tempo de residência na célula (*dwell-time*). Todos os usuários podem ser inicialmente alocados em micro-células e se o intervalo de solicitação entre um *handoff* e outro for muito pequeno, presume-se que o usuário é de alta velocidade e deve ser repassado para a camada superior.

A referência [60] discute como distribuir o espectro entre as camadas definindo qual percentual dos canais deve estar em cada camada, de forma a obter um rendimento máximo. O artigo apresenta um desenvolvimento analítico baseado em processos de nascimento e morte multidimensionais. O cálculo assume a alocação perfeita dos móveis em suas camadas de acordo com as suas velocidades: pedestres são alocados nas micro-células e veículos são alocados direto nas macro-células. A partição do espectro é calculada baseada na probabilidade de bloqueio em cada camada.

Posteriormente, em [50] é feita uma revisão dos trabalhos publicados em arquiteturas de múltiplas camadas. Apresenta também um resumo sobre *handoff*. O trabalho comenta sistemas de três camadas com macro, micro e pico células, mas trabalha com duas camadas. O principal objetivo é obter um equilíbrio entre maximizar o número de usuários por área e diminuir o controle da rede, principalmente quanto ao número de *handoffs*. São apresentados dois algoritmos de controle de admissão da chamada: um que considera a possibilidade de transbordar usuários de uma camada inferior para a superior e outro que considera, além da mobilidade do usuário, diversas classes de usuários e suas taxas de transmissão.

Em [51], os mesmos autores voltam a abordar a alocação de usuários de classes diferentes em sistemas de duas camadas. O objetivo é fazer a alocação dos usuários, não só de acordo com sua velocidade, mas também de acordo com a prioridade de sua classe. As chamadas são enfileiradas pela prioridade e são então encaminhadas à alocação. Tal algoritmo é comparado ao algoritmo que não atribui prioridades, atende por ordem de chegada, FIFO. A comparação é feita através de resultados numéricos provenientes

de solução analítica por cadeia de Markov bi-dimensional.

A referência [71] apresenta estratégias de alocação em micro e macro-células, simuladas em MATLAB[®], de sistema celular bi dimensional para usuários pedestres e carros com uma só classe de serviço.

2.4

Controle de Admissão de Chamadas

O controle de admissão de chamadas consiste em escolher qual tipo de usuário terá preferência na disputa dos recursos do sistema. Os recursos são escassos e a demanda aumenta cada vez mais com as novas aplicações para as redes móveis celulares. O controle de admissão define as prioridades que serão dadas aos usuários diante de determinados critérios.

Os algoritmos de controle de admissão podem ser classificados de acordo com seus objetivos de projeto, segundo a referência [4]. Uma das características principais dos algoritmos é a localização do comando da rede. Quando o comando da rede é local situando-se em cada base, diz-se que o CAC é **distribuído**. Se a rede possui uma central que comanda várias bases, o controle é **centralizado**.

Outra característica importante é se o algoritmo atende apenas a uma classe de usuários, como nos sistemas 1G e 2G, ou a mais classes de usuários, sistemas 3G e 4G. Estas e outras características estão apresentadas na tabela 2.1, que resume a Classificação dos Algoritmos de Controle de Admissão de Chamadas.

Tabela 2.1: Aspectos de classificação para os algoritmos de CAC.

Escolha de projeto	Opção	Vantagens	Desvantagens	Comentários
Centralização	Central	Mais eficiente	Complexo	Não realista
	Distribuído	Simples	Pouco eficiente	Normalmente usado
Escala de informação	Global	Mais eficiente	Mais Complexo	troca de informação
	Semi-local	Eficiente, menos complexo	Complexidade Moderada	menor troca de informação
	Local	Simples	Menos eficiente	Não precisa de troca de informação
Dimensão dos Serviços	Uma classe	-	-	Redes 1G e 2G
	Múltiplas classes	-	-	Redes 2.5G, 3G e 4G
Otimização	Ótima	Mais eficiente	Difícil de implementar, assume muitas simplificações	Desejável
	Sub-ótima	Mais realista	Menos eficiente	técnicas heurísticas e inteligentes
Tempo de decisão	Pro-ativo	Rápido	Mais erros, necessita de informação anterior	Baseada em parâmetros
	Reativo	Menos erros, não necessita de informação anterior	Lenta	Baseada em Medidas
Tipo de informação	Ocupação da célula	Não necessita estimar a mobilidade	Necessita modelo de ocupação da célula	-
	Mobilidade	Não necessita de modelo de ocupação da célula	Necessita estimar a mobilidade	-
Granularidade da informação	células uniformes	Simples	Granularidade simplista	-
Enlace Considerado	Uplink	-	-	Transmissão limitada
	Downlink	-	-	Necessita de retorno do móvel para a base

2.4.1

Sistemas de Uma Classe de Usuários

Em um sistema de uma única classe de usuários, o critério mais comum é o de dar prioridade aos usuários em *handoff*. Um fato amplamente reconhecido na literatura, tal como na referência [31], é o de que é preferível bloquear uma chamada nova a perder uma outra que já está em andamento e que necessita de um *handoff*.

Um dos compromissos na alocação de usuários em uma rede celular móvel é garantir requisitos de QoS diferentes para o bloqueio de chamadas novas e para as falhas de chamadas em *handoff* e, ao mesmo tempo, aumentar a eficiência da utilização da banda diante de um tráfego não uniforme no tempo e no espaço.

CAC para Controlar a Probabilidade de Falha de Handoff

Como foi dito acima, a manutenção de um serviço a um usuário com atendimento em curso é mais importante que iniciar o atendimento a um novo usuário. Sendo assim, deve-se dar atenção especial às chamadas em *handoff*. Foram desenvolvidos vários algoritmos que procuram dar prioridade às chamadas em *handoff* [31] [33] [19] [22] [47] [48].

Uma das principais linhas é a reserva ou guarda de canais exclusivos para o *handoff*, Canais de Guarda (*Guard Channel- GC*). Outros trabalhos colocam as solicitações de *handoff* em fila até que um canal seja liberado na célula-alvo, Fila de *Handoff* (*Handoff Queue- HQ*). Em [19] é apresentada uma comparação entre as estratégias de canais de guarda e filas, considerando alocação fixa e dinâmica dos canais. Os canais de guarda são compartilhados pelas células através da alocação dinâmica. Os dois métodos são comparados por solução analítica, envolvendo a teoria das filas e também por simulação.

Há ainda alguns algoritmos que monitoram o deslocamento do usuário, identificando qual a provável célula para aonde ele vai. Se a nova célula estiver com alta intensidade de tráfego, os recursos disponíveis podem não ser suficientes para garantir a qualidade do serviço, QoS, e a chamada pode ser interrompida.

As principais características dos algoritmos que controlam a probabilidade de falha de *handoff* estão apresentadas na tabela 2.2.

Tabela 2.2: Abordagens para os algoritmos de CAC que visam controlar a P_{FH} .

Abordagem	Explicação	Trabalhos	Comentários
Banda de Guarda	Recursos reservados exclusivamente para H	[68],[31]	P_{BN} alto
Carregamento na célula mãe e nas vizinhas	Chamada nova é admitida se o carregamento é inferior a um determinado valor.	[2],[55],[57]	Necessita de troca de informação
P_{FH} estimado	P_{FH} estimado é usado como critério de admissão	[22], [47], [48]	Necessita de troca de informação
CAC ótimo com restrição P_{FH}	Política de admissão determinada pela otimização de uma função de P_{FH}	[59]	Normalmente resolvida por Processo de Markov
Nível de interferência baixo para chamadas novas	Chamadas N e H são admitidas baseadas no nível de interferência	[7],[10]	Usada para controlar a qualidade do sinal em redes CDMA

Guarda de Recursos para Chamadas em Handoff

Um dos primeiros trabalhos a apresentar a idéia da reserva de recursos para o *handoff* foi feito por Hong em [31]. O trabalho tem sido um dos mais citados desde então e foi posteriormente aperfeiçoado pelos seus autores em [33]. O objetivo da guarda de canais para o *handoff* é priorizar o *handoff* de uma maneira simples e eficaz reservando recursos que podem ser usados apenas pelos usuários em *handoff*. No entanto, há uma certa tendência a aumentar a probabilidade de bloqueio das novas chamadas o que é aceitável dentro das convenções de qualidade de serviço.

Em [56] é proposto um modelo analítico para calcular os recursos de banda necessários para atender às chamadas em *handoff*. O modelo apresentado supõe tráfego uniforme de usuários de uma mesma classe, a princípio, trata-se de usuários de voz.

Inicialmente utilizava-se um número inteiro de canais de guarda, posteriormente surgem trabalhos com a preocupação de descobrir a quantidade de canais a ser reservada [59] e [18].

Margem de Interferência de Guarda

Um algoritmo baseado no conceito dos canais de guarda para sistemas TDMA/FDMA é proposto em [10], para sistemas CDMA. A idéia central consiste em reservar recursos para as chamadas em *handoff* como nos canais de guarda. Como o fator limitante da capacidade do sistema CDMA é a interferência, a reserva de recursos é feita através da reserva de uma

margem de interferência chamada de Margem de Interferência de Guarda, (*Interference Guard Margin* - IGM). O algoritmo proposto trabalha com diferentes taxas que podem ser adaptativas ou não. Dependendo do tipo de chamada, ela será aceita ou não de acordo com o nível de interferência que esta chamada causa. Esta reserva de margem de interferência é baseada no tráfego nas células vizinhas. O algoritmo proposto é simulado com os parâmetros de WCDMA em *OPNET*_®, *Optimized Network Engineering Tool*.

Disponibilidade de Recursos

Em [62] é apresentado um esquema que combina a iniciação do *handoff* com a disponibilidade de canal na célula-alvo, tal esquema proposto é testado por simulação.

Na referência [1], os autores apresentam o conceito de árvore de conexão virtual, que é um conjunto de bases e rede fixa que controla o móvel. A cada móvel é associada uma coleção de números que define o caminho de conexão entre o móvel e a parte fixa da rede. À medida que o móvel se desloca, o controle é passado de uma base à outra, através da árvore, sem que o sistema central precise interferir. O sistema central só detém o controle do *handoff* quando este é feito de uma árvore para outra. Este método permite auferir as vantagens de aumento de capacidade de micro e pico-células sem ocupar muito o processador com *handoffs*. Em [2], os mesmos autores estudam e propõem métodos de controle distribuído para redes micro celulares de alta velocidade, utilizando o conceito de árvore de conexão virtual.

CAC Ótimo com Restrições de Probabilidade de Falha de Chamadas em Handoff

No processo de admissão de chamadas, a probabilidade de falha de chamadas em *handoff* (P_{FH}) é controlada com o objetivo de mantê-la abaixo de certos valores aceitáveis e maximizar a utilização dos recursos ou uma função de custo normalmente definida em cada trabalho.

Normalmente, a política de admissão ótima é modelada por um Processo de Decisão Markoviano (MDP). Então, utiliza-se a programação linear para encontrar a melhor política de admissão em cada estado, conforme exposto nas referências [11],[42],[67],[44],[30],[72].

Em [30] foi demonstrado que a política de otimização ótima para o CAC que maximiza a função de custo e atende as restrições de QoS relacionadas a Probabilidade de Falha de Chamadas em *Handoff* (P_{FH}) é a utilização de Canais de Guarda Fracionários. No entanto, se o número de células e usuários é muito grande o uso de MDP torna-se impraticável. Então, algoritmos genéticos foram propostos como uma alternativa para encontrar políticas de CAC sub-ótimas, porém com menor esforço computacional. Um algoritmo genético foi usado em [75] para encontrar o CAC quase-ótimo baseado em algoritmos genéticos, minimizando uma combinação linear de P_{BN} e P_{FH} . Similarmente, um algoritmo genético é utilizado em [73] para maximizar a utilização dos recursos, enquanto mantém a restrição de P_{FH} . Embora a utilização dos recursos, obtida pelo algoritmo, seja menor que o valor ótimo de MDP (menos de 10%), o autor afirma que esta redução é satisfatória devido a redução do esforço computacional.

2.4.2

Sistemas de M Classes de Usuários

A evolução natural da utilização das redes de comunicações móveis de serviços de voz para incluir serviços de dados, tais como transferência de arquivos, mensagens curtas e vídeo, forçou os sistemas a se adaptarem às novas demandas.

Para oferecer serviços diversos, com requisitos de qualidade diferentes, uma solução natural seria a reserva de recursos. Inspirados na reserva de recursos para os usuários em *handoff*, já bem conhecida e aplicada, passou-se a reservar recursos para os usuários de diferentes classes. No entanto, como o espectro é limitado, não se admite a possibilidade de reservar em excesso, uma vez que a reserva de recursos diminui a sua utilização.

O principal objetivo dos últimos trabalhos na área é a flexibilidade das redes para se adaptar às novas demandas e o aumento da utilização da banda mantendo os requisitos de QoS.

O problema do controle de admissão de usuários de múltiplas classes em redes móveis pode ser visto como uma extensão de múltiplas classes em redes fixas, tais como [34], [38], [39] ou como redes sem fio com algoritmos de uma classe [40], [32] para múltiplas classes de usuários.

Esquemas com Qualidade Adaptativa devido aos Recursos que variam durante a Chamada

Em [69] é apresentado um modelo de alocação de canal com prioridade chamado de *Priority Based Channel Allocation* (PCA). O objetivo é atender a duas classes de usuários com diferentes níveis de qualidade. Os canais são divididos em canais de alta prioridade e canais de baixa prioridade. O esquema é baseado em dois conjuntos de informação que são armazenados em dois vetores. O primeiro vetor é NCS - *Next cell State*, que é um vetor de M componentes, onde cada componente armazena uma possível célula-alvo e o segundo vetor é chamado de TP - *Transition Probability*, um vetor de N elementos com probabilidade de transição para cada uma das N células vizinhas. Existem dois vetores TP, um para alta prioridade e outro para baixa.

Na referência [80] é descrito um esquema de prioridade de *handoff* com QoS para sistemas micro e pico-celulares. O esquema proposto é adaptativo e trabalha com dois tipos de classes de usuários: classe I, banda larga, é suposta como adaptativa em QoS, aceitando canais com taxas baixas; e a classe II, banda estreita, não aceita degradação de QoS. As novas chamadas de ambas as classes são aceitas, se for possível assegurar seus requisitos de banda. Quanto ao *handoff*, a chamada será aceita na nova célula, se esta for capaz de proporcionar a banda total requerida ou uma banda menor, diminuindo a qualidade do serviço prestado, desde que respeite o limite de QoS. A nova célula pode obter a banda necessária para a chamada em *handoff* diminuindo os níveis de QoS de uma ou mais chamadas adaptativas em andamento. Se a chamada em *handoff* é não adaptativa, então o *handoff* só terá êxito se a nova célula puder prover a banda requisitada. Neste mesmo trabalho, é proposta uma outra versão do algoritmo que é mais flexível. A modificação consiste na definição de frações para a aceitação das chamadas em *handoff* e das novas chamadas, o restante é perdido. Esta flexibilidade alivia o sistema do *handoff* banda larga e faz com que as chamadas em andamento recebam mais do que seus requisitos mínimos de banda. Os esquemas propostos são comparados por simulação com os esquemas tradicionais de canais de guarda, (*Guard Channel - GC*), e enfileiramento de *handoff* (HQ) apresentando bons resultados.

Em [27] são estudadas estratégias de alocação flexíveis (*Flexible or Adaptive Resource Allocation - FRA*), que podem ou não ter esquemas que usam degradação/compensação de sessões em progresso. A degradação de uma sessão (chamada de voz ou transferência de dados) consiste na redução gradual dos recursos alocados para uma sessão em progresso. A

compensação é o processo reverso, há o aumento dos recursos para a sessão. No caso de compensação ou de degradação, ocorre a variação da banda alocada ao usuário ao longo do tempo da sessão. O artigo também analisa o desempenho de Estratégias de Alocação de Recursos Adaptativas e as compara com estratégias de alocação anteriormente propostas, [35], [46] e [18]. Em [27] é também proposto um algoritmo novo para analisar o desempenho das estratégias de alocação flexíveis com QoS uniforme. A análise é baseada no fato de que nas estratégias flexíveis que promovem QoS uniforme, todas as sessões ativas de cada tipo de serviço compartilham igualmente seus recursos e tendem a utilizar uma banda média. A referência [35] propõe uma estratégia chamada Alocação de Recursos Igualmente Compartilhados (ERSA), onde este compartilhamento é feito de forma uniforme dentre as chamadas novas, baseado no número de chamadas em andamento. Em [46] é apresentado um esquema de alocação de recursos adaptativo (AREAS) que reduz a realocação de recursos através da introdução de um limiar para decidir quando a realocação é necessária. Esta estratégia é generalizada em [18] para múltiplos serviços.

No artigo [74], os autores propõem uma abordagem integrada de gerência de recursos que combinam o uso de monitoramento por posição (*position-assisted*) e predição da mobilidade para fazer reserva de banda com um controle de admissão e estratégia de reconfiguração de banda para gerenciar uma QoS flexível. Introduz uma estrutura baseada em agentes móveis que pode ser usada para executar as funções de geolocalização e gerência dos recursos da rede. Também é desenvolvido um modelo para obter o intervalo ótimo de atualização de informação, visando minimizar o custo de operação do sistema. A comparação dos resultados do esquema proposto com os resultados correspondentes de um sistema convencional, que garante apenas a reserva de banda como meio de se manter os requisitos de QoS, demonstra que o sistema integrado proposto pode minorar o problema de reservar em excesso e aumentar significativamente a capacidade do sistema.

Reserva de Recursos por Classe de Usuário

Em [21] são comparados métodos de alocação de recursos e suas características de desempenho no que se refere a sistemas que partilham e sistemas que reservam recursos. Nesta referência os recursos reservados são estáticos, não variam com o tráfego.

Posteriormente, em [22] os mesmos autores descrevem dois algoritmos distribuídos e preditivos para o controle de admissão de chamadas. Os

algoritmos fornecem probabilidades de bloqueio separadas para cada tipo de classe de usuário do sistema, baseadas no tráfego das bases vizinhas e diferem na admissão de chamadas em *handoff*: o primeiro compartilha a banda e é chamado CS(*completely shares*); e o segundo faz a reserva de recursos, chamado RES.

Em [23] são apresentados algoritmos de controle de admissão para usuários de múltiplas classes, incluindo uma família de algoritmos preditivos que fazem a reserva de recursos de acordo com o tráfego nas células vizinhas. Os recursos rádio são definidos como unidades de banda que são alocadas aos usuários de acordo com as suas necessidades e apresenta a comparação da reserva estática, reserva dinâmica e compartilhamento de recursos. A comparação das estratégias de gerência de recursos é feita analiticamente e por simulação de um sistema celular para uma rede unidimensional, como uma estrada com duas classes de usuários: voz e banda larga. A simulação é feita com alocação fixa de unidades de banda e sem a movimentação dos usuários. A necessidade de *handoff* é definida por uma variável aleatória que calcula o tempo de permanência do usuário na célula.

A referência [49] tem como objetivo a comparação entre duas técnicas para lidar com a banda: Compartilhamento Total (CS - *Complete Sharing*) e Partição Dinâmica (DP - *Dynamic Partition*). A comparação entre as duas técnicas é feita através de análise e simulação. O artigo demonstra que ambos os esquemas apresentam desempenho semelhante através da manipulação apropriada dos parâmetros de controle. O desafio é que a Partição Dinâmica (DP) pode atingir os requisitos de QoS mais facilmente, ao custo de reservar demais e diminuir a eficiência de utilização da banda, quando comparada com a política de Compartilhamento Total (CS).

Em [26] há a proposta de uma fronteira móvel para a partição dos canais. Isso facilita prover os diferentes requisitos de QoS e é adaptável ao tráfego não uniforme, variável no tempo. Suas limitações são que, tanto voz quanto dados requerem a mesma banda e não há diferença de tratamento para as chamadas de voz novas e para as chamadas em *handoff*.

No desenvolvimento analítico, o esquema de Partição Dinâmica (DP) e o esquema de Reserva de Banda por Limiar Duplo (DTBR - *Dual Threshold Bandwidth Reservation*) são modelados como cadeias de Markov tri e bidimensionais, respectivamente. Tanto a chegada do tráfego de voz quanto do tráfego de dados são assumidas como sendo distribuições de Poisson, num sistema homogêneo. São escritas as equações de balanceamento que podem ser resolvidas pela técnica desenvolvida por [29], que é baseada no desenvolvimento típico das equações de Champan-Kolmogoroff, onde

existe um subconjunto de probabilidades de estado chamado fronteira e todos os estados podem ser expressos como uma função dos estados de fronteira. A idéia básica dessa técnica é escolher as fronteiras primeiro e depois derivar as expressões para todas as probabilidades de estado remanescentes como funções dos valores das fronteiras e então, resolver um sistema de equações para essas fronteiras. Depois disso, determinar todas as probabilidades de estado através das fronteiras. Os dois esquemas são comparados e apresentam as seguintes características: 1) o esquema DP sempre tem a menor probabilidade de falha por *handoff* de voz; 2) o esquema DTBR com um limiar pode atingir a maior utilização de canal e a menor probabilidade de bloqueio de dados; e 3) o esquema DTBR com dois limiares pode atingir a menor probabilidade de bloqueio de novas chamadas de voz.

Em [3], os autores propõem e comparam esquemas de alocação dinâmicos para múltiplas classes. O controle de admissão proposto é a aceitação da chamada se houver banda suficiente para tal classe acrescida de um limiar de margem, que varia por classe e também considera se a chamada é nova ou *handoff*. O cálculo do número de canais de guarda é feito segundo o método proposto, que espera o maior caso, *ExpectedMax*, segundo a referência [76], para fins de comparação. O objetivo do algoritmo proposto em [3] é evitar o bloqueio em massa das classes que demandam maior quantidade de recursos.

Os algoritmos são comparados através de simulação em uma rede de 18 células, sem movimentação. O tempo de residência na célula é sorteado simulando uma distribuição exponencial. A taxa de chegada de novas chamadas varia ao longo da simulação: aumenta na primeira metade da simulação e diminui na segunda metade. A Probabilidade de *handoff* de uma célula para outra é variável, com o intuito de simular o tráfego em uma cidade nos horários de tráfego intenso.

Na referência [9] os autores propõem um controle de admissão cujo limiar de acesso é baseado no número de solicitações das diferentes classes de QoS. A política é semelhante ao esquema de canais de guarda dinâmico, onde a estimativa dos recursos necessários para atender as chamadas de alta prioridade é baseada na razão sinal/ruído do móvel (*SNR*) e na informação do perfil de tráfego dos usuários das células vizinhas. O algoritmo proposto é simulado em *OPNET*®.

Controle de Admissão de Chamadas com Qualidade Adaptativa na Entrada

Em [52] é apresentado o esquema de alocação de chamadas em um sistema integrado de voz e dados chamado de Ajuste Gradual Preditivo (*Predictive Gradual Adjustment - PGA*), que é comparado com dois outros esquemas anteriores para sistemas *High Speed Circuit Switched Data* (GSM-HSCSD). O primeiro esquema de alocação é o de Capacidade Máxima, onde o usuário é alocado com a taxa desejada ou é bloqueado, ou terminado no caso de solicitação de *handoff*. O segundo esquema para comparação é o de Alocação Flexível, onde o usuário define sua taxa ideal e também sua taxa mínima de aceitação. Porém, uma vez alocado, não há mudança depois de iniciada a chamada.

No esquema proposto em [52], de Ajuste Gradual Preditivo, o usuário define sua taxa ideal e a taxa mínima aceitável. Considerando o número de canais disponíveis, o sistema definirá a taxa com a qual o usuário poderá entrar. Os três esquemas são comparados através de uma ferramenta de simulação com mobilidade dos usuários em um sistema que utiliza células em uma região urbana. O esquema proposto apresenta bons resultados.

Em [16], [17] e [28], os autores apresentam uma seqüência de artigos sobre alocação flexível de recursos em sistemas integrados de voz e dados.

2.4.3

Controle de Admissão Considerando o Tráfego em Andamento

Em [56], é apresentado um algoritmo de controle de admissão distribuído que considera a própria célula e o número de chamadas das células vizinhas. Com o objetivo de atender prontamente o tráfego em tempo real, a condição de aceitação da chamada é simples: basta que a probabilidade de queda seja menor que um determinado limiar. É apresentado um modelo analítico para um sistema unidimensional com extensão para o bi dimensional.

Em [13], é proposto um esquema preditivo e adaptativo que reserva frações da banda, dependendo do tráfego das células vizinhas. A localização do móvel é feita contando com as informações do sistema *GPS*. Os autores analisam o controle da janela de tempo para o processamento dos dados dos diferentes tipos de tráfego.

2.4.4

Reserva de Recursos Baseada no Comportamento Estatístico da Mobilidade do Usuário

Em [8], é apresentado um esquema de reserva de recursos com agrupamento dinâmico para prover os requisitos de QoS nas redes sem fio da próxima geração. O esquema proposto é baseado na estimativa estatística de recursos para prover QoS para o tráfego multimídia nas redes celulares. São estabelecidas várias reservas de divisões no tempo de acordo com a informação de mobilidade de cada base. A quantidade de banda reservada para cada base é ajustada dinamicamente para cada grupo de reserva.

Reserva de banda melhora o QoS das chamadas em andamento, porém aumenta a taxa de bloqueio de novas chamadas e reduz a utilização do espectro. Então, tem-se um compromisso entre a utilização e a reserva de banda.

O esquema proposto em [8] é baseado na estimativa probabilística de recursos. É assumido que quando o móvel solicita uma nova conexão ou se move para uma nova célula ele fornece as informações: (1) tempo estimado do móvel na célula; (2) tempo de residência na célula vizinha; (3) probabilidades de *handoff* para as células vizinhas e (4) a banda requerida pelo móvel, incluindo a soma de todos os fluxos desse móvel. Os tempos (1) e (2) são estimados pela velocidade atual do móvel e pelo histórico de tráfego gravado na célula. Neste artigo, as estimativas de tempo acima são usadas no método proposto de estabelecimento de grupos de reservas, onde a quantidade de banda reservada é dinamicamente ajustada para cada grupo. O esquema proposto é simulado e o tempo de duração de cada chamada é uma distribuição exponencial.

Em [48] é apresentado um algoritmo que utiliza o conceito de *shadow cluster* aonde a aceitação da chamada depende do tráfego nas células vizinhas. As bases compartilham a informação sobre seus móveis. O algoritmo de admissão da chamada é baseado no comportamento estatístico de cada usuário. A probabilidade de movimentação do móvel é calculada de acordo com a função densidade de probabilidade do tempo de residência e do vetor de probabilidade de *handoff*. O modelo é equacionado para sistemas uni e bidimensionais e simulado para um sistema unidimensional.

Na referência [12] é sugerido um algoritmo de reserva de banda adaptativo de acordo com o perfil de movimentação do usuário. Para cada usuário são monitorados: índice da célula atual; índice da célula anterior; índice da célula seguinte (de acordo com movimentações anteriores) e o número de vezes que o usuário se moveu desta para cada uma das vizinhas.

De acordo com o percurso que o usuário está acostumado a fazer, a reserva de recursos é feita para esperar o *handoff* na provável célula seguinte. Esta reserva é baseada na probabilidade de movimentação do usuário que é calculada através dos parâmetros monitorados. Com o objetivo de não reservar continuamente recursos para usuários que estejam parados, as células são divididas em duas regiões: central e próxima à borda. Se o usuário está próximo à borda e é do tipo sensível a atraso em tempo real, serão reservados recursos. Outra medida proposta para evitar desperdício de recursos é que a reserva seja calculada individualmente, mas os recursos reservados são compartilhados. Assim, se for feita a previsão de reserva para um usuário que acaba não executando o *handoff*, os recursos reservados, e não utilizados por este, poderão ser utilizados por outro.

2.4.5

Outras Técnicas

Em [24] é proposto um algoritmo para garantir a probabilidade de queda mínima para serviços múltiplos e maximizar a aceitação de novas chamadas, sem fazer reserva de recursos.

Em [43] é apresentado um modelo analítico para análise do *handoff* no tráfego. O modelo assume tráfego uniformemente distribuído e, conseqüentemente, assume que o *handoff* para dentro da célula seja o mesmo que o *handoff* para fora da célula. A mesma equipe apresenta em [11] um algoritmo de controle de admissão baseado em limiar de aceitação e utiliza um modelo com programação não-linear.

Em [11] é apresentado um algoritmo que procura solução que maximiza o rendimento por programação linear e simultaneamente atende os requisitos de QoS.

Na tese de doutorado em [5], é desenvolvido um simulador (*MACSIM*) para permitir a análise de comportamento de dois componentes básicos da pilha de protocolos: o protocolo *MAC*, *Medium Access Control*, e o algoritmo *SAC*, *Service Admission Control*. O algoritmo de controle de admissão de serviços (*SAC*) é o componente do sistema de comunicação responsável por definir se uma requisição de uma unidade móvel para ter acesso ao protocolo *MAC* deve ser aceita ou não.

2.5

Conclusão Sobre os Trabalhos Relacionados

Com a tendência da elevação da demanda por redes móveis, sendo o espectro limitado, deve haver sempre preocupação em aumentar a utilização dos recursos rádio. Embora o tema Controle de Admissão de Chamadas tenha sido objeto de estudos nas duas últimas décadas, ainda há pontos a se explorar.

Os algoritmos de controle de admissão dinâmicos são capazes de rastrear as mudanças de tráfego e adaptar a aceitação de usuários de forma a atender aos requisitos de qualidade de serviço.

Este trabalho propõe vários algoritmos dinâmicos, tanto para uma classe, quanto para várias classes de usuários, sempre com o objetivo de atender aos requisitos de qualidade de serviço e simultaneamente manter a maior utilização possível dos recursos. Uma das características que diferem o trabalho realizado nesta tese dos trabalhos relacionados apresentados neste capítulo, é a preocupação simultânea com cada uma das restrições de qualidade de serviço, no que se refere as probabilidades de bloqueio de chamadas novas e de falha de chamadas em *handoff* de cada uma das classes individualmente.

O estudo detalhado e a determinação dos limites de intensidade de tráfego de entrada para os quais é possível atender aos requisitos de qualidade de serviço é parte integrante deste trabalho. A determinação dos limites foi feita tanto para uma classe, com as suas duas restrições, quanto para duas classes incluindo suas quatro restrições.

Encontra-se na literatura muitos trabalhos que consideram o tempo de permanência do usuário na célula como uma distribuição exponencial. Este trabalho considera o tempo de permanência do usuário na célula sendo modelado pela distribuição exponencial e também por outras distribuições adequadas. São apresentados o desenvolvimento analítico e o comportamento dos algoritmos propostos simulados nestes cenários.