

Um Novo Esquema para Controle de Admissão de Chamadas em Redes Móveis sem Fio baseadas no Protocolo IP

Kelvin Lopes Dias, Judith Kelner, Djamel F. H. Sadok

Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7851 – 50732-930 – Recife – PE – Brasil

{kld, jk, jamel}@cin.ufpe.br

Resumo. Este artigo propõe um novo esquema para controle de admissão de chamadas (CAC) em redes móveis sem fio baseadas no IP que evita a sinalização de reservas antecipadas. A proposta utiliza sistemas de posicionamento para previsão de mobilidade, informações sobre o tráfego agregado estimado para a célula de admissão e que é proveniente de sua vizinhança. Um novo conceito denominada banda de guarda virtual é empregado para evitar sinalização de reservas antecipadas para a próxima célula predita. Resultados de simulação demonstram um melhor desempenho de nossa proposta em termos de probabilidade de descarte, probabilidade de bloqueio e utilização efetiva do que outros esquemas baseados em reservas.

Abstract. This paper proposes a novel call admission control (CAC) scheme that avoids per-user reservation signalling overhead and meets the scalability requirement advocated for future all-IP wireless and mobile networks. The novel CAC uses mobility prediction based on mobile positioning systems and takes into account the expected bandwidth to be used by calls handed off from neighbouring cells within an estimation time window. A novel concept called virtual guard band is introduced which aims at preventing pre-reservation in the next predicted cell. Simulation results show that our proposal performs better than reservation-based schemes in terms of handoff dropping probability, call blocking probability and effective utilization.

1. Introdução

Um grande esforço em pesquisa e desenvolvimento vem sendo realizado para proporcionar aos usuários móveis¹ meios que viabilizem a comunicação transparente nas futuras redes móveis sem fio totalmente baseadas no protocolo IP (*All-IP wireless/mobile networks*). Tais redes visam substituir a tecnologia tradicional de comutação por circuitos utilizada nas redes de acesso sem fio e de núcleo das operadoras, por tecnologia de comutação por pacotes utilizando o protocolo IP fim a fim. Estas redes vêm sendo moldadas através das evoluções da atual rede de telefonia celular denominada de segunda geração ou 2G (*2nd Generation*) [1]. Alguns exemplos

¹ Utilizaremos os termos *usuário* ou *móvel* significando o mesmo que dispositivo móvel, estação móvel ou terminal móvel.

de tecnologias empregadas nas sucessivas gerações são: GPRS (*General Packet Radio Service*) na 2,5G, UMTS (*Universal Mobile Telecommunications System*) na 3G e as futuras redes denominadas 4G, com taxas de transmissão superiores as da 3G e que serão totalmente baseadas no protocolo IP.

Pesquisas em redes móveis sem fio baseadas no IP são especialmente desafiadoras quando o tópico envolvido é o provimento de qualidade de serviço (QoS – *Quality of Service*). As aplicações dos usuários móveis podem sofrer degradações de desempenho devido às propriedades dos enlaces sem fio e devido ao processo de *handoff*, que é o mecanismo que transfere uma chamada em andamento de uma célula para outra à medida que o usuário se movimenta através da área de cobertura do sistema. Caso a nova célula para onde o usuário acabara de migrar não possua capacidade suficiente para acomodar os requisitos de banda da aplicação, a chamada será descartada. Da perspectiva do usuário, descartes de chamadas (*dropped calls*) em andamento são menos desejáveis do que o bloqueio de novas chamadas.

Um outro aspecto importante que deve ser levado em consideração no projeto destas redes é a questão da escalabilidade no provimento de QoS. Os arcabouços atuais da Internet para provimento de QoS recaem, basicamente, em duas arquiteturas: serviços integrados (IntServ – *Integrated Services*) e serviços diferenciados (DiffServ – *Differentiated Services*). Apesar do IntServ prover garantias de QoS estritas através de sinalização explícita, falha em alcançar o objetivo de escalabilidade devido à sua abordagem baseada em reservas. Por outro lado, DiffServ objetiva prover garantias menos estritas de QoS através da classificação de pacotes nos roteadores de borda da rede e diferenciação no tratamento dos pacotes nos roteadores internos de acordo com um conjunto de classes denominadas PHB (*Per Hop Behavior*), em favor de uma melhor escalabilidade da rede [2]. O BB (*Bandwidth Broker*) [10] é a entidade responsável por implementar as políticas de gerenciamento de recursos na arquitetura DiffServ, incluindo o controle de admissão.

A reserva de recursos tem efeitos mais negativos em redes móveis sem fio do que nas redes fixas de alta velocidade devido à escassez de largura de banda dos enlaces sem fio. Em nossa opinião, uma arquitetura escalável para as futuras redes móveis sem fio baseadas no IP, deveria evitar o uso excessivo de sinalização por usuário para efetuar reservas de recursos de rádio. Promovendo, desta forma, um uso mais eficientemente dos recursos dos enlaces sem fio.

Para obter comunicação transparente, as propostas de arquiteturas para redes móveis sem fio baseadas no IP devem contemplar meios para manter níveis adequados de QoS para as aplicações durante o processo de *handoff*. Um componente importante para tal arquitetura é o mecanismo de controle de admissão de novas chamadas (CAC – *Call Admission Control*). O CAC deve ser utilizado para atacar o aspecto de mobilidade, decidindo se novo usuário deve ser admitido ou rejeitado. Para que o usuário seja admitido, tanto o novo usuário deverá receber garantias de manutenção dos requisitos de QoS para sua aplicação, bem como, sua admissão não deverá degradar o desempenho das aplicações já presentes na rede.

Neste artigo propõe-se um novo esquema para controle de admissão de chamadas em redes móveis sem fio baseadas no protocolo IP que evita a tradicional

reserva de recursos por usuário e o *overhead* de sinalização associado. As características chave do CAC proposto são:

- ✓ Utiliza predição de mobilidade baseada em GPS (*Global Positioning System*) para determinar a próxima célula a ser visitada pelo móvel;
- ✓ Estima a carga agregada que migrará para a célula de admissão², dentro de uma janela de tempo, com o intuito de evitar que a admissão deste novo usuário provoque o descarte de chamadas em andamento, caso estas efetuem *handoffs* para a célula de admissão, e;
- ✓ Introduce um novo conceito denominado *banda de guarda virtual* (VGB – *Virtual Guard Band*), através do qual evita-se que o usuário sinalize e efetue reservas de banda na próxima célula a ser visitada.

Nossa proposta é comparada por meio de simulações com esquemas baseados em reservas. Os resultados das simulações demonstram que nossa proposta garante melhores resultados em termos de probabilidade de descarte em casos de *handoffs* mal sucedidos e probabilidade de bloqueio de novas chamadas, além de proporcionar uma melhor utilização efetiva dos recursos dos enlaces sem fio do que os esquemas baseados em reservas utilizados nas comparações.

O restante deste artigo é organizado da seguinte forma. Na seção 2, descrevemos os trabalhos relacionados. Em seguida, apresentamos a proposta do novo CAC na seção 3. A seção 4 descreve o modelo de simulação e suas suposições. Os resultados de desempenho são apresentados na seção 5. Apresentamos as conclusões e trabalhos futuros na seção 6.

2. Trabalhos relacionados

O controle de admissão de chamadas em redes celulares deve decidir sobre a aceitação ou rejeição de novos usuários no sistema garantindo que, mesmo que o usuário movimente-se para outra célula, sua chamada continue em andamento com alta probabilidade. De uma maneira geral, a idéia básica das estratégias para controle de admissão existentes na literatura é reservar recursos *a priori* nas células do sistema para reduzir a probabilidade de descarte no caso do móvel efetuar um procedimento de *handoff*[3][4][6][7][8]. As propostas existentes na literatura para CAC em redes móveis sem fio podem ser divididas, basicamente, em duas categorias: fixas e dinâmicas. Estratégias fixas, como o *canal de guarda* (GC - *Guard Channel*) [3], proporcionam tratamento diferenciado para chamadas de *handoffs*, reservando um número fixo de canais exclusivamente para tais chamadas. Este esquema reduz a probabilidade de descarte devido ao *handoff*, entretanto, aumenta a probabilidade de bloqueio de novas chamadas. A vantagem desta estratégia é sua simplicidade porque não há a necessidade de trocar de informações entre as estações rádio-base. Por outro lado, não é flexível para

² Utilizaremos o termo *célula de admissão* para indicar a célula onde o usuário está requisitando a admissão.

tratar cargas variáveis, uma vez que não utiliza informações sobre a carga da célula atual e das vizinhas.

Propostas de esquemas dinâmicos baseados em reservas estendem o esquema básico da estratégia fixa de acordo com estimativas da taxa de *handoffs* derivadas do número de chamadas na vizinhança e do padrão de mobilidade destas chamadas para reservar banda na próxima célula a ser visitada pelo usuário. De uma maneira geral, o usuário será admitido se existir banda suficiente na célula de admissão e se for possível efetuar reservas na vizinhança desta célula.

O esquema proposto em [4] utiliza o histórico do agregado de *handoffs* em cada célula para prever a probabilidade de que uma chamada efetuará o *handoff* para uma determinada célula vizinha. Baseado na predição de *handoffs*, o número de canais necessários é reservado antecipadamente. Cada estação rádio-base registra o número de falhas na predição de *handoffs* e ajusta a reserva modificando o tamanho da janela de tempo utilizada para fazer as estimativas. Esquemas baseados em históricos para prever *handoffs* são problemáticos no que se refere ao *overhead* para desenvolvimento, armazenamento e atualização do tráfego para diferentes células. Estes históricos não serão totalmente confiáveis uma vez que podem ocorrer mudanças de curto prazo (modificação do tráfego devido a acidentes), mudanças de médio prazo (reroteamento de tráfego devido a construções e reparos) e de longo prazo (construção de um novo *shopping center*).

O controle de admissão proposto em [5] leva em consideração o número de chamadas nas células adjacentes, além do número de chamadas existentes na célula onde o usuário requisita a admissão. Os autores desenvolveram um modelo teórico para computar os requisitos dos pedidos de *handoff* com o objetivo de manter uma determinada probabilidade de descarte de *handoff*. O modelo proposto assume que todas as chamadas são idênticas, o que deixa de ser uma suposição válida em ambientes com aplicações multimídia com requisitos de banda variáveis. Este esquema não utiliza qualquer informação de mobilidade para prever a próxima célula na trajetória do usuário.

Em [6] é proposto um esquema preditivo de reserva de canal (PCR – *Predictive Channel Reservation*) baseado no uso do GPS (*Global Positioning System*) para prever a próxima célula a ser visitada. Este esquema realiza predição baseada na posição atual e orientação do móvel. A reserva é disparada apenas se o móvel alcança um determinado limiar de distância com relação à próxima célula. Reservas podem ser invalidadas (reserva falsa) devido a mudanças na trajetória. Neste caso, o cancelamento de reservas deve ser enviado para desalocar o canal reservado. Este esquema exige uma quantidade excessiva de mensagens de sinalização para reserva e cancelamento de reservas falsas. Além disso, reservas podem reduzir a probabilidade de descarte de *handoffs* com o efeito colateral de aumentar a probabilidade de bloqueios, o que pode acarretar uma utilização ineficiente dos recursos da rede.

3. Descrição da proposta para controle de admissão

3.1 O esquema básico em duas fases

O novo esquema para controle de admissão proposto neste artigo se beneficia de sistemas de posicionamento (GPS), como feito em [6], para a determinação da próxima célula a ser visitada pelo móvel. Da mesma forma como em [5], utiliza-se informação proveniente das células vizinhas àquela onde deseja-se efetuar uma nova admissão. Contudo, melhoramos a idéia em [5] para levar em consideração uma estimativa de carga que migrará para a célula de admissão durante uma determinada janela de tempo para aplicações com requisitos variados de taxa de transmissão e não apenas considerando taxas idênticas como em [5].

Para realizar as estimativas, utiliza-se um mecanismo de janela de tempo de tamanho τ durante a qual cada estação rádio-base que controla uma célula atualiza a quantidade de carga (somatório das taxas de transmissão requeridas) dos móveis destinados a uma determinada célula vizinha. Por exemplo, como ilustrado na Figura 1, $\Omega_{1,2}$ e $\Omega_{1,5}$ representam as cargas dos móveis presentes na célula 1 que migrarão para a célula 2 e 5, respectivamente. Da mesma forma, $\Omega_{5,1}$ e $\Omega_{2,1}$ representam as cargas dos móveis com chamadas em andamento que migrarão para a célula 1 a partir das células 5 e 2, respectivamente. Através do envio periódico da carga estimada de cada célula para as suas vizinhas, ao final da janela de tempo, nosso esquema é robusto com relação às mudanças de carga na rede. É importante destacar que as mensagens de carga estimada são enviadas através do *backbone* da rede de acesso o que evita o desperdício dos escassos recursos do meio sem fio para efetuar tal troca.

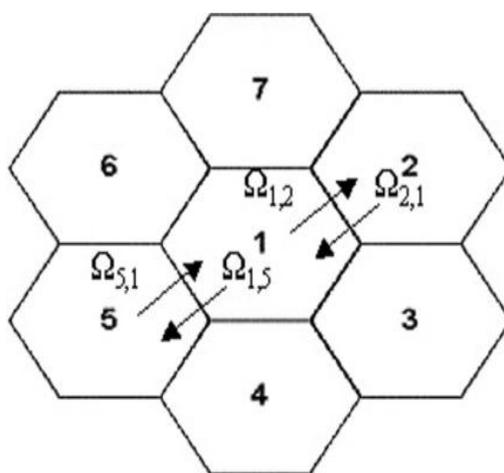


Figura 1: Troca de mensagens de carga estimadas entre células.

O CAC proposto contempla duas fases para a admissão de uma nova chamada. A primeira fase determina se a célula onde o usuário requisita admissão tem capacidade suficiente para suportá-lo e quanto da carga agregada dos usuários presentes nas células vizinhas migrará para a célula de admissão durante a janela de tempo. A fase final é baseada na disponibilidade de banda suficiente na próxima célula predita na trajetória do usuário utilizando um novo conceito que denominamos de *banda de guarda virtual* e que será descrito posteriormente.

Na primeira fase, as duas condições a seguir devem ser satisfeitas:

$$\sum_{i=1}^N Bi + Bk \leq Cj \quad (1)$$

e

$$E(\Omega_{i,j}) + E(\Omega_{j,j}) + Pr_K * B_K \leq \delta C_J \quad (2)$$

onde :

$$0 < \delta \leq 1,$$

$$E(\Omega_{i,j}) = \sum_{i=1}^M \sum_{s \in S_{i,j}} (Ph_s * B_s), \text{ e} \quad (3)$$

$$E(\Omega_{j,j}) = \sum_{i=1}^N (Pr_{i,j} * Bi_{j,j}) \quad (4)$$

Na equação (1), N representa o número de conexões existentes na célula de admissão. C_j é a capacidade do enlace sem fio da célula de admissão. Bi é a banda da i -ésima conexão presente na célula de admissão e Bk é a banda requerida pela nova conexão do usuário k .

O CAC proposto considera situações de sobrecarga devido ao tráfego excessivo proveniente de células vizinhas. O objetivo é evitar que conexões em andamento na vizinhança da célula de admissão sejam descartadas durante *handoffs* para esta célula. Com este fim, a equação (2) estima a carga esperada na célula de admissão durante uma janela de tempo, que inclui ambas, a quantidade de banda utilizada por chamadas já existentes na célula de admissão e a banda necessária para chamadas que efetuarão *handoffs* para esta mesma célula. Neste ponto, precisamos explicar as suposições adotadas para os modelos de mobilidade e duração de chamadas. O tempo gasto por um usuário na célula (tempo de residência) antes de efetuar *handoff* e a duração de chamadas são exponencialmente distribuídos com médias $1/h$ e $1/\mu$, respectivamente. Então, baseados nesta suposição, a probabilidade de que um usuário efetuará *handoff* durante uma janela de tempo de τ unidades de tempo, é dada por : $P_h = 1 - e^{-h\tau}$. A probabilidade de que um usuário permaneça na célula durante esta janela de tempo é: $P_r = e^{-(h+\mu)\tau}$. Note que $(1 - Ph - Pr)$ é a probabilidade de que uma chamada seja finalizada em τ unidades de tempo. A seguir, descrevemos os componentes da equação (2).

Cada usuário admitido, s , é associado a uma célula vizinha dependendo de sua direção atual de movimento. Cada célula tem uma base de dados que armazena a carga esperada que deve migrar para cada uma de suas vizinhas. Periodicamente, cada célula recebe a informação resumida sobre a carga estimada que receberá da vizinha que lhe enviou tal informação, bem como, envia sua própria para suas vizinhas. É importante frisar, novamente, que estas mensagens são enviadas através do *backbone* da rede de acesso constituído de enlaces com fio (*wired*), o que evita desperdício de utilização dos

enlaces de rádio sem fio (*wireless*) para este propósito. O usuário pode mudar a direção do movimento e, assim, a célula que está servindo tal usuário deve refletir esta mudança, atualizando a base de dados através da indicação da nova célula para a qual o usuário, provavelmente, efetuará o *handoff*, para que na fase de troca de mensagens de cargas, a nova célula seja avisada. Um grupo de usuários registrados com a célula i que têm j como a próxima célula predita em suas trajetórias, é representado por $S_{i,j}$.

A carga esperada devido a *handoffs* provenientes das células vizinhas à célula de admissão é dada em (3), onde $E(\Omega_{i,j})$ representa a carga esperada que migrará da célula vizinha i para a célula de admissão j . Ph_s e B_s são a probabilidade que um usuário na célula vizinha i efetuará *handoff* para a célula j durante uma janela de tempo e sua largura de banda requerida, respectivamente. M é o número de vizinhos da célula de admissão j .

Além da carga proveniente de *handoffs* para a célula de admissão, a primeira fase do CAC deve prever a carga que permanecerá nesta célula decorrente de usuários que, provavelmente, não efetivarão *handoffs* durante a janela de tempo. Então, a expectativa de carga de usuários já presentes na célula de admissão e que permanecerão nesta célula é dada por (4), onde $E(\Omega_{j,j})$ representa a carga prevista para permanecer na célula de admissão, $Pri_{j,j}$ a probabilidade que os usuários permanecerão na célula j e, $Bi_{j,j}$ é a largura de banda requisitada por cada usuário i .

O teste de admissão na primeira fase do CAC deve levar em consideração a probabilidade de *handoff* da nova requisição de conexão para o usuário k denotada como Pr_k , e sua requisição de banda B_k , como descrito na equação (2). Esta informação específica do usuário deve ser adicionada à carga de *handoff* esperada das células vizinhas e à carga remanescente na célula de admissão. O fator δ é um parâmetro de projeto utilizado para reforçar o grau de predição de congestionamento e sua sintonia deve ser realizada dependendo dos objetivos da operadora da rede. Um valor alto para δ beneficiará a utilização global da rede a um custo de aumentar a probabilidade de descarte de *handoff*. Pequenos valores para δ devem gerar uma baixa utilização da rede, contudo, diminuindo a probabilidade de descarte.

A segunda fase de nossa proposta utiliza informação das células vizinhas para implementar o que denominamos de banda de guarda virtual (*virtual guard band*). Este é um esquema virtual porque não há reservas de canal. A estação rádio-base da célula onde o usuário requisita admissão realiza uma estimativa da banda que estará disponível na célula predita para um futuro *handoff* do usuário. Esta estimativa é feita com base em informações recebidas periodicamente das vizinhas, as mesmas trocadas para viabilizar a primeira fase de nossa proposta. Estas informações correspondem ao somatório das expectativas de carga que migrará para a célula predita no próximo *handoff*, obtidas a partir das células que formam a vizinhança da célula predita em questão. O objetivo é evitar que a nova chamada seja admitida quando estimasse que a próxima célula no caminho do usuário estará sem banda suficiente para acomodar sua conexão em caso de *handoff*. Desta forma, minimizasse o descarte de chamadas devido à ausência de banda na célula para onde o usuário efetuará o *handoff*. Para que a segunda fase do mecanismo de controle de admissão tenha sucesso, a equação abaixo deve ser satisfeita:

$$E(\Omega_{jh}) + E(\Omega_{jr}) \leq \beta C_j \quad (5)$$

onde $E(\Omega_{jh})$ representa a carga esperada de futuros *handoffs* para a célula j , que é a célula predita para o *handoff* do usuário que está requisitando admissão. $E(\Omega_{jr})$ representa a carga que deverá permanecer na célula j predita para o *handoff*. C_j representa a capacidade da célula predita para o *handoff*. β é o fator utilizado para reforçar quanto da capacidade da célula predita para o novo *handoff* deveria ser virtualmente reservada para priorizar *handoffs*. Por exemplo, se β fosse igual a 0,8, isto poderia ser interpretado como uma forma de aproximar o comportamento da estratégia de canal de guarda (GC) onde 20% da capacidade da célula é reservada para *handoffs*, sem, de fato, efetuar qualquer reserva.

Diferente da proposta apresentada em [5], onde cada estação rádio-base é diretamente responsável pelo envio e recepção de mensagens de suas células vizinhas, em nossa proposta utilizamos uma arquitetura hierárquica de servidores, numa idéia similar a de arquiteturas de gerenciamento de recursos escaláveis usando *bandwidth brokers* [13][14], que é responsável por receber as informações enviadas pelas estações rádio-base e enviar de volta o somatório resumido para viabilizar a segunda fase do CAC, juntamente com as informações necessárias para a primeira fase[11]. Dessa forma, a informação sobre carga predita para a provável célula de *handoff* não é trocada entre as estações rádio-base a cada nova requisição de admissão ou *handoff*. Esta informação é trocada periodicamente através da rede de *backbone* entre as estações rádio-base sem comprometer o objetivo de escalabilidade de nossa proposta.

3.2 Otimização do esquema básico

Além do uso do GPS para prever a próxima célula na direção do movimento do usuário, nossa proposta utiliza-o para otimizar o cálculo das estimativas de carga que devem migrar dentro da janela de tempo. Ao invés de usar o GPS para verificar o *limite de distância* a partir do qual a reserva deve ser feita na próxima célula a ser visitada, como em [6], nossa abordagem utiliza um *limite de tempo* para calcular a estimativa de carga que migrará para a célula predita dentro da janela de tempo. Se o tempo de residência restante do móvel na célula atual for menor que um *limite de tempo* pré-determinado (o que indica a iminência da execução de um *handoff*), a carga que migrará deve ser considerada como total. Dessa forma incrementa-se a carga estimada para a próxima célula predita no caminho do usuário com o valor total do seu requisito de banda. Isto difere do esquema básico da seção 3.1, que segue a idéia descrita a seguir, no próximo parágrafo.

Caso contrário, utiliza-se o esquema básico. Se o tempo restando para o *handoff* for superior ao *limite de tempo*, determina-se uma carga probabilística dada pelo produto entre a probabilidade de *handoff* da classe de mobilidade do usuário durante uma janela de tempo e a banda requerida pelo usuário k ($Ph_k \times B_k$ para atualizar carga de *handoff* ou $Pr_k \times B_k$ para atualizar carga de permanência). Esta otimização está relacionada à forma como as cargas direcionais (previstas de uma célula para todas suas vizinhas) são atualizadas para, posteriormente, ao final da janela de tempo, serem trocadas entre as células. O valor armazenado nas bases de dados de carga direcionais é atualizado

localmente nas estações rádio-base durante as admissões de novos usuários, *handoffs* e finalizações das chamadas.

É importante enfatizar que além das filosofias de uso do GPS serem para objetivos diferentes (disparar reservas, na proposta em [6] vs. estimar carga que migrará dentro do intervalo de tempo, no caso da nossa proposta), os limiares adotados para otimizar tais filosofias são também diferentes em termos de unidades de medida (distância, em [6] vs. tempo, em nossa proposta). A escolha de um limiar de tempo ao invés de um limiar baseado na distância da vizinha, permite que nossa proposta contemple classes de mobilidade de usuários, o que não é contemplado na proposta em [6], uma vez que mesmo que o limiar de distância seja alcançado, indicando a proximidade da vizinha, o usuário pode permanecer estacionário próximo à fronteira da vizinha sem realmente efetuar o *handoff*, o que pode gerar desperdício de banda com reservas e sinalizações desnecessárias na proposta [6].

4. Modelo de simulação e suposições

Para a avaliação de nossa proposta, foi desenvolvido um simulador escrito em SMPL (*Simulation Program Language*)[12], uma extensão da linguagem C para desenvolvimento de simulações baseadas em eventos discretos. O modelo de simulação inclui um modelo de topologia de células do sistema, modelo de tráfego e modelo de mobilidade. Além do simulador desenvolvido para nossa proposta, desenvolvemos também quatro simuladores para esquemas de controle de admissão baseados em reservas, descritos posteriormente.

4.1 Topologia

A Figura 2 apresenta a topologia celular simulada que é composta de 19 células, utilizando hexágonos para representar as relações de adjacência entre as células. Cada célula tem a capacidade de 20 BUs (*bandwidth units*)[9]. Para evitar o efeito de borda utilizamos a mesma abordagem adotada na maioria dos trabalhos relacionados da literatura: quando um usuário mover-se para fora da topologia deverá entrar novamente na rede pelo outro lado. Por exemplo, usuários nas células de borda 8 e 12 que estejam deslocando-se para fora da topologia deverão entrar novamente no sistema através das células 14 e 18, respectivamente.

4.2 Modelo de tráfego

Em nosso modelo de simulação utilizamos chamadas heterogêneas com requisitos iguais a 1, 2, 3 e 4 BUs. A probabilidade de início de uma chamada de usuário é uniformemente distribuída entre as quatro classes, isto é, 25% de cada tipo de chamada, em média, estará presente no sistema a qualquer momento. A taxa de chegadas de chamadas segue um processo de Poisson com taxa média de chegadas λ , que será a carga da simulação expressa em número médio de chamadas em cada célula por segundo. A duração média das chamadas é exponencialmente distribuída com média $1/\mu$ igual a 180s[6].

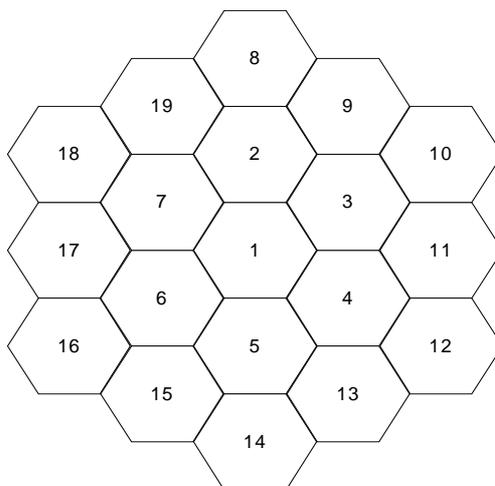


Figura 2: Topologia celular adotada nas simulações.

4.3 Modelo de mobilidade

Neste estudo, dividimos os usuários em duas classes: usuários com alta mobilidade e baixa mobilidade. Estas classes servem para caracterizar o tempo de residência na célula corrente. Estes tempos de residência são distribuídos exponencialmente com valores médios de 60s e 420s para usuários com alta e baixa mobilidade, respectivamente.

Estudamos dois cenários de mobilidade: No cenário I, a célula alvo do *handoff* é escolhida no momento da admissão e após os sucessivos *handoffs*, dessa forma, o móvel não muda sua direção enquanto cruza uma célula. No cenário II, utiliza-se uma probabilidade de 50% para que o móvel mude sua trajetória enquanto cruza uma célula, isto é, o móvel pode mudar sua trajetória a qualquer momento enquanto cruza uma célula com probabilidade igual a 50%. Para ambos cenários de mobilidade, se a direção do móvel mudar, seja no cenário I (após um *handoff*) ou no cenário II (enquanto o usuário viaja na célula), uma nova célula dentre suas seis vizinhas é aleatoriamente (uniformemente) escolhida, bem como, um novo tempo de residência é escolhido usando a distribuição exponencial com a média igual a da classe do móvel em questão.

4.4 Configuração dos parâmetros dos esquemas de CAC

Para avaliar nossa proposta (VGB) foram implementados quatro esquemas de CAC baseados em reservas, sendo um fixo, dois dinâmicos e um híbrido: canal de guarda fixo (GC)[3], reserva antecipada (RA), reserva preditiva de canal (PCR)[6] e uma variação do PCR denominado de HPCR (Hybrid PCR)[6] que mescla PCR com GC para obter melhores resultados em termos de probabilidade de descarte (P_d). Estes esquemas serão rotulados nos gráficos dos resultados das simulações através de suas siglas.

Os parâmetros do VGB foram configurados em função do esquema básico de canal de guarda (GC) que reserva 10% da capacidade das células para *handoffs*. Com este objetivo, nossa proposta utiliza os fatores δ e β configurados com o valor 0,9. O tamanho da janela de tempo τ , o limiar de tempo para o cálculo da carga estimada do

usuário e a periodicidade entre trocas de mensagens de carga agregada, foram todos configurados com o valor de 30s. Adotamos um valor para τ e a periodicidade entre trocas de mensagens superior ao adotado em [5] que foi de 20s, o que beneficia nossa proposta com menos *overhead* de sinalização para a disseminação de tais mensagens.

O esquema de reservas antecipadas (RA) representa um modelo geral e sem otimizações das propostas de reservas dinâmicas antecipadas presentes na literatura. O móvel só é admitido no sistema se for possível reservar banda na próxima célula da sua trajetória. Em caso de *handoff*, o móvel tentará novamente reservar banda na próxima célula, vizinha àquela onde está tentando efetivar o *handoff*. A chamada do móvel não será descartada se não for possível realizar a reserva antecipada nos casos dos *handoffs*. Assim, as reservas são compartilhadas pelos móveis e usuários podem utilizar reservas feitas por outros usuários caso não haja recursos livres no momento do *handoff*. A chamada do móvel só será descartada, no *handoff*, caso não exista nenhuma reserva suficiente para seus requisitos de banda e nem banda livre (não reservada) na célula. Não impusemos um limite na quantidade máxima de banda que pode ser reservada antecipadamente em cada célula com o intuito de garantir uma baixa probabilidade de descarte decorrente de *handoffs* mal sucedidos para o esquema RA e, assim, verificar se nossa proposta consegue garantir níveis similares.

Para uma comparação justa com nosso esquema, adaptamos os esquemas PCR e HPCR para utilizar limiares de tempo, ao invés de limiares de distância para controlar o disparo de requisições de reservas nestes esquemas. Os esquemas diferem apenas na inclusão de uma reserva fixa de 10% da capacidade da célula para o HPCR. Ambos utilizam um limiar de tempo de 30s, indicando que o móvel disparará um pedido de reserva apenas quando o seu tempo restante de permanência na célula for inferior ou igual a 30s.

5. Avaliação de desempenho

5.1 Métricas

As métricas avaliadas neste artigo são:

- ✓ Probabilidade de bloqueio de novas chamadas (P_b): Razão do número de novas chamadas admitidas sobre o número total de requisições de novas chamadas ao sistema.
- ✓ Probabilidade de descarte de *handoff* (P_d): Razão do número de *handoffs* descartados sobre o número total de tentativas de *handoff*.
- ✓ Utilização ($Util$): razão da quantidade de recursos de banda em uso no sistema sobre a capacidade total do sistema.
- ✓ Utilização efetiva ($Util\ Efet$). Razão da banda utilizada por usuários que finalizaram com sucesso suas chamadas sobre a capacidade total do sistema, isto é, não são consideradas chamadas admitidas que foram descartadas durante um *handoff*.

Diferente das análises comumente presentes na literatura, que avaliam apenas a utilização ($Util$), nós acreditamos que uma melhor avaliação da qualidade de esquemas de controle de admissão possa ser traduzida em termos dos recursos que foram

consumidos por usuários que conseguiram finalizar suas chamadas com sucesso (*Util Efet*).

5.2 Resultados das simulações

Para a simulação de todas as propostas, utilizamos a mesma topologia, modelo de tráfego e mobilidade, descritos anteriormente na seção 4. Os gráficos apresentam as métricas de interesse em função da carga da rede descrita em termos da taxa média de chamadas por célula por segundo (chamadas/cel/s) que varia desde uma carga baixa até situações de congestionamento. Os resultados das simulações foram obtidos através da técnica de médias em lotes (*batch means*) removendo-se estados transientes do início da simulação, computando a média dos lotes e calculando a média final a partir das médias obtidas destes lotes. O tempo de simulação foi de 10^5 s. As barras verticais para os valores médios apresentados nos gráficos foram calculadas para representar um intervalo de confiança de 95%.

A Figura 3 apresenta os resultados obtidos para o cenário I, isto é, o usuário sempre efetuará *handoff* para a célula que foi predita no momento que entrou na célula corrente. O pior resultado para a probabilidade de descarte (Pd) foi o do esquema básico de canal de guarda fixo (GC) como mostrado na Figura 3(a). Apesar da melhor utilização do GC (Figura 3(c)), o ganho é, na verdade, observado quando verificamos a utilização efetiva na Figura 3(d), indicando quanto dos recursos foram utilizados por usuários que finalizaram com sucesso suas chamadas, excluindo assim, a banda desperdiçada por usuários que foram descartados durante *handoffs*. A utilização efetiva do GC (Figura 3 (d)) só foi superior que a do RA em situação de carga baixa.

Os melhores resultados de Pd foram obtidos por nossa proposta (VGB) e pelo esquema de reservas antecipadas simples (RA), que foram aproximadamente os mesmos ao longo de todo o espectro de cargas, exceto para cargas mais baixas (inferiores a 0,2 chamada/cel/s), onde o esquema RA obteve um Pd ligeiramente melhor. Entretanto, o efeito colateral de um menor Pd , em situações de cargas muito baixas obtido pelo esquema RA, é o uso ineficiente dos recursos da rede, comprovado por uma maior probabilidade de bloqueio de novas chamadas Pb (Figura 3 (b)). Este efeito colateral fica ainda mais visível através dos menores valores obtidos para a utilização (Figura 3 (c)) e utilização efetiva (Figura 3 (d)) devido às reservas feitas sob demanda para o usuários admitidos na rede e que podem gerar bloqueios desnecessários de novas chamadas que desejem admissão.

É importante ressaltar que nossa proposta alcança níveis similares de Pd aos do esquema RA, sem que para isso seja necessário sinalizar e efetuar reservas antecipadas. Como descrito na seção 3, em nossa proposta há uma troca periódica (a cada 30s em nossas simulações) de informações sobre estimativas de migração de carga entre as vizinhas. Estas informações proporcionam um ajuste fino para evitar estados de congestionamento (1ª fase), bem como, evitam descarte na célula predita para *handoff* (2ª fase) no momento da admissão.

Os esquemas baseados em reservas dinâmicas, PCR e HPCR, obtiveram resultados de desempenho inferiores aos obtidos por nossa proposta com relação a todas as métricas analisadas (exceto a utilização - Figura 3(c)), mas obtiveram melhor utilização e utilização efetiva do que o RA. Com relação à probabilidade de descarte,

nosso esquema é superior porque baseia-se em informação de carga das células vizinhas, bem como, não admite excessivamente novos usuários que poderiam ser descartados na célula vizinha predita para o *handoff* na fase de admissão, uma vez que estima se haverá capacidade suficiente para o usuário em caso de *handoff*. Já no esquema PCR, usuários podem ser admitidos, mas não, necessariamente, sinalizar reserva para alocar recursos na célula predita na sua trajetória. Esta postergação na reserva melhora a utilização da rede, mas pode degradar a *Pd*: requisitar a sinalização para reservar banda na próxima célula, apenas no momento em que o limiar de tempo for alcançado, pode ser inútil devido a situações de congestionamento com provável falta de recursos na célula predita para o *handoff* deste móvel.

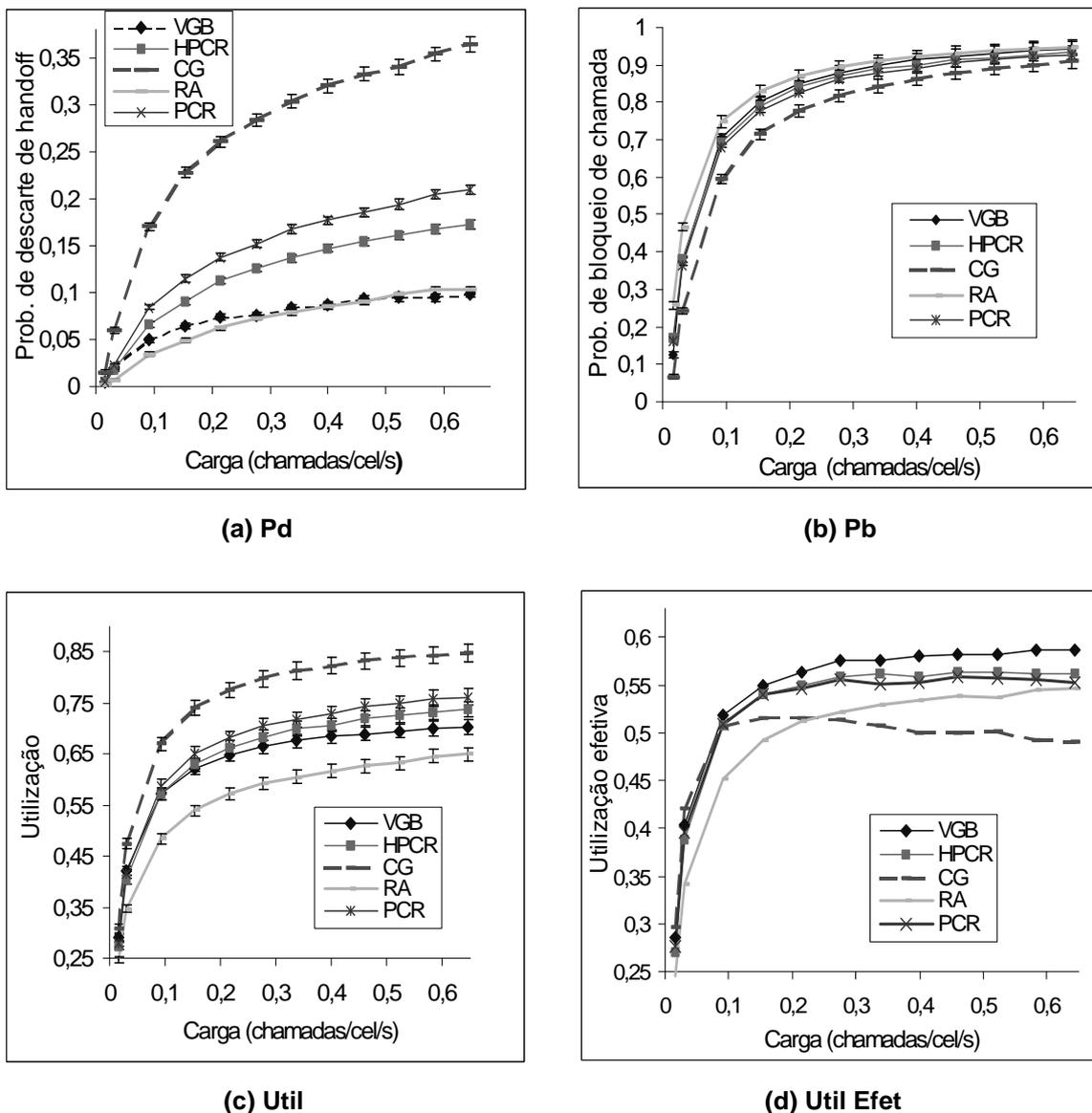


Figura 3: Cenário I.

O HPCR tem um melhor *Pd* do que o PCR, pois utiliza uma parte da capacidade de suas células para priorizar *handoffs*. Confirmamos, assim, a superioridade do

esquema PCR híbrido sobre o esquema PCR puro com relação à garantia de um menor Pd como demonstrado em [6].

É importante salientar que os esquemas PCR e HPCR obtiveram Pd maiores que o de RA devido à postergação para efetivar as reservas na célula vizinha: a reserva pode não ser disparada no momento em que o usuário é admitido ou efetua um *handoff* e, sim, quando o usuário alcança um limiar de tempo indicando a iminência de um *handoff*. Esta postergação na reserva beneficia a utilização e a utilização efetiva dos recursos da rede para os esquemas PCR e HPCR. Em [6] a avaliação da proposta é realizada através de comparação com o esquema fixo de canal de guarda (GC), não havendo comparação com outros esquemas dinâmicos baseados em reservas. Resultados de simulação não apresentados neste artigo, mostraram que um valor mais alto para o limiar de tempo de disparo de reservas no PCR e HPCR aproximará a Pd alcançada pelo esquema RA, entretanto, isto também reduz a utilização da rede tal como acontece com o esquema RA.

Os próximos resultados da avaliação do VGB são apresentados na Figura 4 e referem-se ao cenário II. Mantivemos os mesmos parâmetros discutidos anteriormente na configuração dos esquemas. Contudo, o móvel poderá mudar a nova célula escolhida para o *handoff* enquanto cruza a célula corrente com probabilidade igual a 50%. O objetivo é verificar o comportamento de nossa proposta em um cenário de mobilidade menos previsível.

Nossa proposta obteve resultados ainda melhores com relação aos resultados das demais propostas da literatura. Na Figura 4(a), a menor probabilidade de descarte (Pd) foi obtida por nossa proposta (VGB) tanto para carga baixa quanto alta e a pior pelo esquema de canal de guarda fixo (GC). Os esquemas dinâmicos baseados em reservas (RA, PCR, HPCR) obtiveram uma deterioração da probabilidade de descarte neste novo cenário de mobilidade devido a um aumento nos pedidos falsos de reserva no momento da admissão ou *handoff*. No caso do esquema RA, reservas falsas são desfeitas a partir do momento em que o usuário muda sua trajetória, no entanto, pode não haver possibilidade de reservar na nova célula para onde o usuário está apontando após a mudança de trajetória.

Para os esquemas PCR e HPCR ocorre o mesmo que no RA, mas com o agravante da reserva ser postergada até que o tempo estimado para o *handoff* seja menor ou igual ao limiar de tempo de 30s utilizado para disparar os pedidos de reservas. Dependendo do valor escolhido para este limiar, mais usuários podem ser admitidos, prejudicando assim o Pd .

Como última análise dos resultados, a utilização obtida nos dois cenários de mobilidade Figura 3(c) e Figura 4(c) demonstraram que nossa proposta só foi superior ao esquema de reservas antecipadas (RA). Contudo, é importante verificar que esta métrica pode não ser um bom indicativo da qualidade do algoritmo, traduzindo apenas a ocupação média da rede tanto por usuários que finalizarão com sucesso suas chamadas quanto por usuários que terão suas chamadas descartadas. Nossa proposta embute uma certa “*inteligência*” no sentido de admitir apenas o tráfego que provocará um menor impacto na probabilidade de descarte de chamadas, bem como, uma melhor utilização efetiva dos recursos da rede.

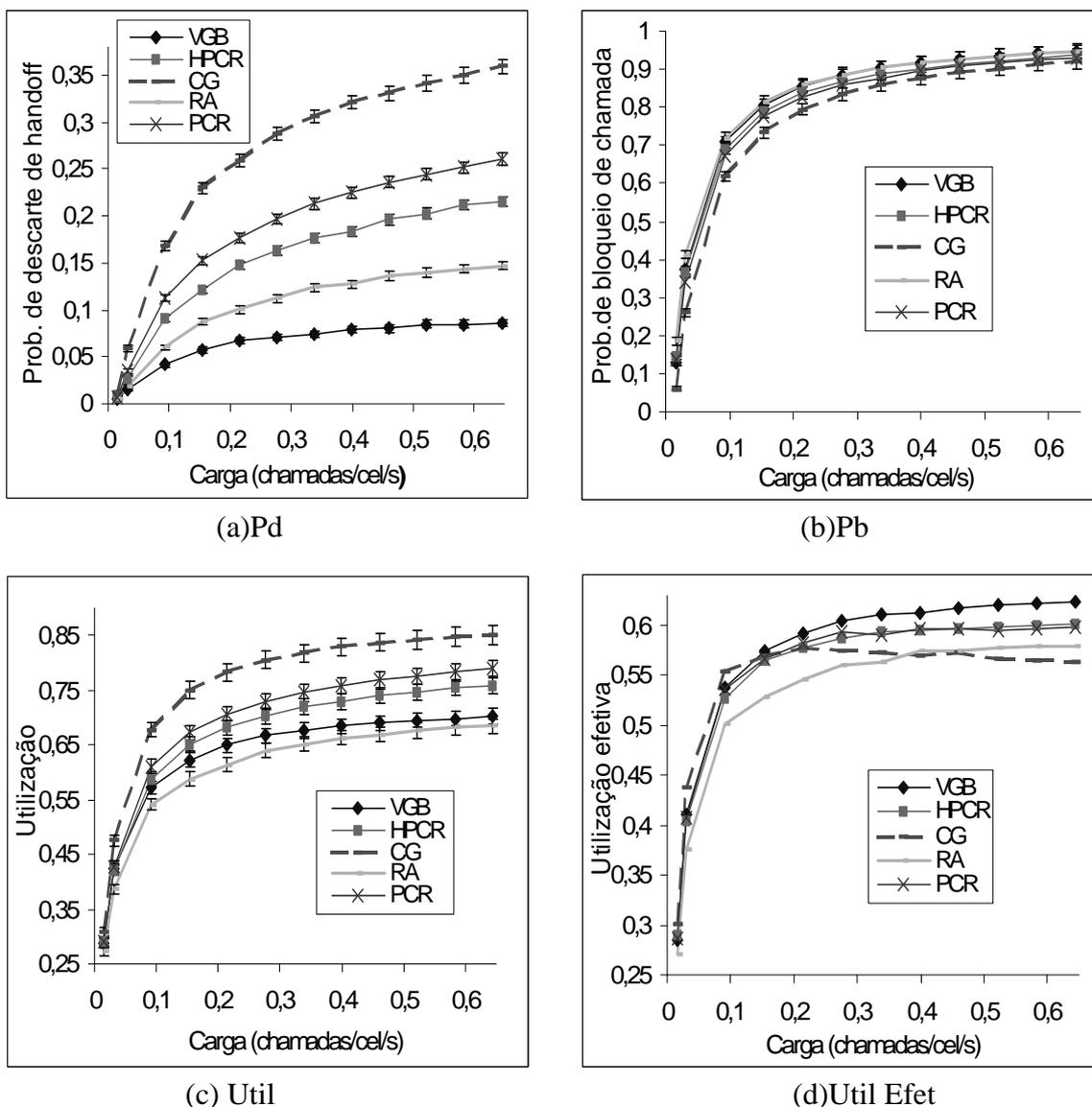


Figura 4: Cenário II.

6. Conclusões e trabalhos futuros

Neste artigo, propusemos um novo esquema de controle de admissão de chamadas (CAC) para oferecer garantias de QoS nas futuras redes móveis sem fio baseadas no protocolo IP. Nossa proposta é compatível com os atuais arcabouços escaláveis de provimento de QoS da Internet, pois evita *overhead* de sinalização de reservas antecipadas por usuário e os bloqueios desnecessários devido a reservas não utilizadas.

Nossa proposta integra predição de mobilidade utilizando GPS, uso e disseminação de mensagens sobre as expectativas futuras de cargas que migrarão entre as células da rede de maneira distribuída, utilizando para isso o *backbone* da rede de acesso. Dessa forma, evita-se o desperdício de recursos de rádio. Um novo conceito,

denominado banda de guarda virtual, foi introduzido com o intuito de evitar reservas antecipadas na próxima célula para *handoff*.

Os resultados de simulações preliminares mostraram que nossa proposta obteve o melhor compromisso em termos das métricas analisadas do que os esquemas baseados em reservas, utilizados no estudo. Dentre as métricas analisadas, é importante enfatizar uma característica interessante de nossa proposta que é a de manter os melhores níveis de probabilidade de descarte e, ao mesmo tempo, a melhor utilização efetiva dentre todas as propostas analisadas neste artigo. Como trabalhos futuros, devem ser analisados mais cenários para validar o mecanismo proposto, tais como, utilização de outros valores para o tamanho da janela de tempo e do limiar de tempo escolhido para o cálculo das expectativas de cargas entre as células.

Referências

- [1] R. Berezdivin, R. Breinig, R. Topp. Next-Generation wireless communications concepts and technologies. IEEE Comm. Magazine, vol. 40, no. 3, Mar 2002, pp. 108 – 116.
- [2] S. Blake et all. An architecture for differentiated services. IETF RFC 2475, Dec. 1998.
- [3] D. Hong and S. Rappaport. Traffic model and performance analysis for cellular mobile radio telephone systems with prioritised and nonprioritised handoff procedures. IEEE Trans. On Veh Teh., 35(3):77-92 August 1986.
- [4] S. Choi and K. G. Shin. Predictive and adaptive bandwidth reservation for handoffs in QoS sensitive cellular networks. In Proceedings of the ACM SIGCOMM'98. September 1998.
- [5] M. Naghshineh and M. Schwartz. Distributed call admission control in mobile/wireless networks. IEEE J. Selected Areas in Commun., 14(4):711-717, May 1996.
- [6] M.H. Chiu and M. A. Bassiouni. Predictive schemes for handoff prioritization in cellular networks based on mobile positioning. IEEE J. Select. Areas Commun. 18(3), March 2000.
- [7] D. A. Levine, I. F. Akyldiz and M. Naghshineh. A resource estimation and call admission control algorithm for wireless multimedia networks using the shadow cluster concept. IEEE/ACM Trans. Networking, vol. 5, pp. 1-12, Feb. 1997.
- [8] C. Oliveira, J. Kim, T. Suda. An adaptive bandwidth reservation scheme for high speed multimedia wireless networks. IEEE J. Select. Areas in Commun., vol. 16, pp 858-874, 1998.
- [9] M. Naghshineh, A.S. Acampora, "Design and Control of Microcellular Networks with QoS Provisioning for Data Traffic", Wireless Networks, vol. 3, pp 249-256, Sept. 1997
- [10] K. Nichols, V. Jacobson and L. Zhang. A two-bit differentiated services architecture for Internet. IETF RFC 2638. July 1999.
- [11] K. L. Dias, Djamel Sadok. A scalable call admission control for QoS provisioning in all-IP networks. International Conference on Third Generation Wireless and Beyond (3Gwireless'2003). San Francisco, USA, May 27-30, 2003.
- [12] M. McDougall. Simulating computer systems: techniques and tools. MIT Press. 1987.
- [13] Z-L. Zhang et al., "On scalable design of bandwidth brokers". IEICE Trans. Comm, E84-B(8), Aug. 2001
- [14] Thomas Engel *et.al.*, AQUILA: adaptive resource control for QoS using a IP-based layered architecture. IEEE Comm. Mag., Jan. 2003, pp 46-53.