

Gustavo de Assis Costa

**Serviços Baseados na Localização do
Usuário em um Ambiente Móvel**

Dissertação apresentada ao Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

Belo Horizonte

Julho de 2001

*Dedico este trabalho aos meus pais, Genair e Enilde,
e aos meus irmãos Álvaro, Lívio e Lúcio.*

Agradecimentos

Agradeço imensamente a Deus por ter me iluminado e guiado durante toda a minha vida e por ter me concedido tão valiosa oportunidade de poder ter conhecido e vivido tantos desafios e superações. Agradeço também ao apoio e carinho de toda a minha família, em especial aos meus pais, por terem me dado o exemplo de honestidade, respeito, humildade e pela oportunidade que me deram de ter uma vida tão digna e feliz. Agradeço muito aos meus irmãos pelo grande apoio e respeito que me deram durante todos os momentos de minha vida.

Agradeço especialmente ao meu orientador, Prof. Antônio Alfredo Ferreira Loureiro, pela paciência, dedicação, orientação e principalmente pelo exemplo de um profissional competente e muito dedicado. Sem dúvida alguma, as suas observações e os seus questionamentos me proporcionaram um grande amadurecimento em relação à pesquisa. Eu não poderia deixar de agradecer também ao meu co-orientador, Clodoveu A. Davis Júnior, pela paciência, pelo grande apoio durante a minha pesquisa e também pelos seus grandes conhecimentos que foram de extrema importância a este trabalho. Agradeço também ao professor Dorgival Olavo pelas dicas, orientações e pelo apoio.

Agradeço ao apoio e à grande amizade dos colegas do mestrado e doutorado, os quais tive a oportunidade de conhecer durante este período. Não posso deixar de citar amigos como o Cleiber Garcia, pelas dicas, pelo apoio e pelos conselhos; e os amigos da “República do Bicalho”, Autran, Paulo Sérgio, Pavel e mais recentemente o Hervaldo, pela convivência harmoniosa e sobretudo divertida que me proporcionaram a tranquilidade necessária ao longo do meu trabalho.

Não tão menos importantes foram as grandes amizades que tive a oportunidade de fazer com o pessoal da “TurmaPos99”. Infelizmente, citar o nome de todos neste momento não é tarefa muito fácil, mas eu gostaria de agradecer às pessoas que direta ou indiretamente contribuíram com meu crescimento pessoal e profissional durante o curso de mestrado e que conviveram comigo quase que diariamente. Em especial eu gostaria de agradecer aos grandes amigos Alisson, Edierley, Filipe, Hilza, Irna, João, Juliana, Júlio Conway, Karine Louly, Karine Versieux, Luciano, Nilma, Pedro Ivo, Rêmulo, Silvana, Tatiana e Vinícius. Peço desculpas àqueles que eventualmente eu tenha esquecido de citar mas eu gostaria de deixar aqui os meus agradecimentos.

Agradeço à Prodabel por ter cedido uma parte da base de dados georeferenciados da cidade de Belo Horizonte, sem a qual eu não poderia ter dado prosseguimento ao meu trabalho.

Agradeço a todos os colegas e amigos do CEFET por terem acreditado e dado apoio ao desenvolvimento deste trabalho.

Resumo

Com o atual desenvolvimento do mercado de dispositivos de comunicação sem fio e também da infra-estrutura e tecnologia relacionados a demanda pela disponibilização de serviços e aplicações úteis vem aumentando cada vez mais. A rede de comunicação celular, uma das grandes forças propulsoras deste mercado, já vem possibilitando o desenvolvimento de tecnologias que permitem o acesso a dados.

Entretanto, mesmo com as facilidades oferecidas por estas tecnologias, o meio de comunicação sem fio apresenta sérias restrições que influenciam na oferta de serviços aos usuários. Para que as aplicações específicas para dispositivos móveis sem fio possam estar disponíveis, os desenvolvedores deverão agora se preocupar com as restrições impostas pelo ambiente, que traz à tona uma nova abordagem na construção das mesmas.

Considerando a relevância dos aspectos mencionados acima, este trabalho tem como proposta principal o desenvolvimento de uma arquitetura que utiliza técnicas de adaptação como uma solução para que as aplicações possam contornar os problemas existentes no meio sem fio. Como estudo de caso para a validação da arquitetura e das técnicas propostas foi desenvolvida uma aplicação de localização de endereços mais próximos do usuário para o ambiente móvel.

Esta aplicação pertence a uma das categorias de aplicações móveis mais atraente no mercado. Ela é composta pelos chamados serviços LBS (Location Based Services) que oferecem facilidades relacionadas à própria localização física do usuário através de tecnologias específicas da infra-estrutura existente. Aplicações do tipo "Páginas amarelas", por exemplo, oferecem uma grande riqueza de informações ao usuário possibilitando que o mesmo possa localizar endereços e outras ocorrências geográficas em uma cidade

No decorrer do texto da dissertação são descritas as especificações e as soluções de implementação utilizadas no desenvolvimento da arquitetura e da aplicação, além da validação destas realizada com um simulador.

Abstract

In the recent development stage of wireless devices, infra-structure and related technologies, demand for availability of useful services and applications is growing quickly. The cellular communication network, one of the most propelling impulses of the market, is already enabling the development of data access technologies.

However, either with facilities offered by these technologies, the wireless communication environment has serious restrictions in services offer to mobile users. Specific applications for mobile wireless devices can be made available only if developers consider restrictions imposed by mobile communication environment through a new approach in applications design.

Considering the relevant aspects mentioned above, the main proposal of this dissertation is to develop an architecture that implements adaptation techniques as a solution for mobile applications. An application for finding nearest address related to a mobile user was developed as an example of service for the architecture and adaptation techniques validation.

This application is included in one of the most attracting mobile applications category, the LBS (Location Based Services) services. These applications can offer facilities related with own fisical location of mobile users through use of specific infra-strucutre technologies. For example, applications like "Mobile Yellow Pages" provide a rich variety of information that mobile users can use to locate addresses and another geographical ocurrences in a city.

In the remainder of this text there is a description of the specification and implementation solutions used in the architecture and application deployment, as well as the validation process accomplished with one simulator.

Sumário

Lista de Figuras	viii
Lista de Tabelas	ix
1 Introdução	1
1.1 Motivação	2
1.1.1 Sistemas de informação geográficos	4
1.1.2 Adaptação	4
1.2 Objetivos	5
1.3 Trabalhos relacionados	6
1.4 Organização da dissertação	10
2 Sistemas de localização	11
2.1 Regulamentação do serviço 911	12
2.2 Serviços de localização	12
2.3 Precisão de posicionamento	13
2.4 Tecnologias de posicionamento	14
2.4.1 Soluções baseadas no terminal	16
2.4.2 Soluções baseadas na rede	23
2.5 Estratégias de localização	31
2.5.1 IP móvel	31
2.5.2 IS-41 e GSM MAP	32
3 Arquitetura de serviços	33
3.1 Considerações sobre o ambiente móvel	33
3.2 Descrição da arquitetura	34
3.2.1 Localização de serviços	37
3.2.2 Adaptação	38
3.2.3 Cliente	40
3.2.4 Localização de usuários	40
4 Aspectos de implementação	41
4.1 Plataforma de desenvolvimento	41
4.2 Especificação e modelagem dos dados	42
4.2.1 Coleta e criação da base de dados	46
4.3 Estrutura de dados utilizada	47
4.4 Aplicação	49
4.4.1 O algoritmo de Dijkstra	50

4.5	Descrição dos módulos da arquitetura	52
4.6	Adaptação	53
4.6.1	Níveis de adaptação	54
4.7	Simulador	57
4.8	Parâmetros de qualidade de serviço	60
5	Testes e validação do sistema	62
5.1	Consultas ao serviço	62
5.2	Adaptação	66
5.2.1	Bateria	66
5.2.2	Largura de banda	67
5.2.3	Memória no terminal	68
5.3	Considerações sobre os testes	70
6	Conclusões e Trabalhos Futuros	71
6.1	Conclusões	71
6.2	Trabalhos futuros	72
	Bibliografia	74
	Glossário	79
A	Padronização dos serviços de localização	82
A.1	Definições de padrões	82
A.1.1	Location estimate	82
A.1.2	Mobile Assisted positioning	82
A.1.3	Mobile Based Positioning	83
A.1.4	Location request	83
A.1.5	LoCation Services (LCS)	83
A.1.6	Padronização de tecnologias de localização para GSM	83
A.1.7	Padronização da arquitetura GSM	84
A.1.8	Padronização de tecnologias de localização para UMTS	86
A.1.9	Padronização da arquitetura UMTS	86

Lista de Figuras

2.1	Método TDOA	17
2.2	Posicionamento <i>Enhanced Observed Time Difference</i> : posicionamento hiperbólico . . .	18
2.3	Posicionamento <i>Assisted GPS</i>	22
2.4	Arquitetura do sistema Cell-STK	25
2.5	Arquitetura do sistema Cell-ID IN	25
2.6	Arquitetura do sistema Cell-ID WAP	26
2.7	Método Cell-ID+TA	27
2.8	Determinação de localização pelo <i>Time Of Arrival</i>	29
2.9	Posicionamento pelo método <i>Angle Of Arrival</i>	30
3.1	Modelo cliente-proxy-servidor	35
3.2	Ambiente de comunicação para LBS	36
3.3	Arquitetura proposta	37
4.1	Representação de circulação por trechos e nós.	44
4.2	Modelo de dados usado na aplicação.	45
4.3	Distância euclidiana	47
4.4	Representação da estrutura de dados	48
4.5	Exemplo de um grafo a ser percorrido	50
4.6	Rota completa entre origem e destino	51
4.7	Gerenciador de recursos	57
4.8	Simulador da aplicação	58
4.9	Lista dos endereços mais próximos	58
4.10	Informação de rota a ser seguida nos formatos texto e figura	59
4.11	Tela para definição dos parâmetros das condições do meio e do dispositivo	60
5.1	Ocorrências de endereços mais próximos a localização do usuário	63
5.2	Rota entre a origem e a primeira ocorrência de endereços mais próximos	63
5.3	Caminho permitido entre a origem e o destino	64
5.4	Rota entre os endereços de número 2285 e 2311 na Avenida Amazonas	65
5.5	Exemplo de um logradouro fragmentado pelo recorte	65
5.6	Tela para indicação de quantidade insuficiente de carga da bateria	67
5.7	Gerenciador configurado para uma largura de banda de 300 Kbps e com tempo de resposta de 5 segundos	67
5.8	Tela de log do servidor	68
5.9	Tela de <i>log</i> do <i>proxy</i>	69
5.10	Requisição de um dispositivo com pouca memória disponível	70

Lista de Tabelas

2.1	Tabela de requisitos de precisão de serviços	14
4.1	Divisão do campo endereço	42
4.2	Tabela de apoio para representação dos logradouros	43
4.3	Forma normalizada de representação do endereço	43
4.4	Tabela de trechos	44
4.5	Tabelas de nós	44
4.6	Representação das cardinalidades entre relacionamentos.	45

Capítulo 1

Introdução

No mundo inteiro está havendo um crescimento constante do mercado de dispositivos de comunicação sem fio e também da infra-estrutura e tecnologia relacionadas. O ambiente computacional onde dispositivos portáteis (que se comunicam por uma rede sem fio) podem se comunicar com a parte fixa da rede e possivelmente com outros dispositivos é chamado de computação móvel. A computação móvel é uma área de pesquisa que despertou muito interesse na comunidade acadêmica e industrial nos últimos anos e muitos são os desafios que envolvem esta pesquisa, pois são necessárias modificações e adaptações de diversos aspectos das mais variadas áreas da computação.

Aliada a este crescimento, a possibilidade de acesso a *World Wide Web* através de dispositivos móveis vem se tornando uma realidade para milhões de usuários por todo o mundo. Com o atual desenvolvimento de tecnologias como **WAP** (Wireless Application Protocol) [53], **i-Mode** (Tecnologia proprietária de acesso a *web* para dispositivos celulares) [27] e **3G** (Sistemas de Terceira Geração) [28] várias companhias de telefonia celular têm investido na disponibilização de infra-estrutura e serviços para atender aos seus usuários. Contudo, mesmo com o atual estágio de desenvolvimento dessas tecnologias, muitos problemas ainda existem.

Esse tipo de ambiente é caracterizado pelas limitações existentes tanto no dispositivo de comunicação móvel quanto na infra-estrutura que o serve. Os dispositivos são bastante limitados quanto a recursos de processamento, memória, bateria e interface com o usuário. Já a infra-estrutura de comunicação se caracteriza pela alta variação e limitação de banda, alta latência e alta taxa de erros. Restrições como estas são muito comuns pois em sistemas de comunicação sem fio, como as redes de telefonia celular por exemplo, várias requisições

de usuários que se movem por uma área de cobertura podem ser atendidas, gerando assim perfis distintos de mobilidade.

As diferenças existentes entre as redes sem fio e as redes fixas implicam em mudanças principalmente nos projetos de aplicações. Diante disso, a capacidade de se oferece conjuntos de serviços comparáveis aos existentes em redes fixas com computadores estáticos torna-se um grande desafio. O tipo de informação necessária em um ambiente móvel pode ser sensivelmente diferente do que é usualmente oferecido em redes fixas. Tal ambiente pode dar margem ao desenvolvimento de aplicações e serviços completamente diferentes daqueles existentes na *web* atualmente [36].

Várias soluções foram propostas e uma das principais investigações tem seguido o princípio de que as aplicações devem ser capazes de se adaptarem às constantes mudanças apresentadas pelo ambiente sem fio. "A mobilidade requer adaptabilidade"[32]. Por exemplo, um usuário utilizando seu PDA para visualizar a imagem de um produto que está em lançamento faz sua requisição numa rede com baixa capacidade de transmissão de dados. O mais provável seria que o sistema pudesse oferecer uma imagem de baixa resolução e em preto e branco. Para tanto, o sistema deve prover a funcionalidade de adaptação, ou seja, o sistema deve ser capaz de reagir às mudanças do ambiente de forma dinâmica.

1.1 Motivação

As grandes metrópoles do Brasil e do mundo são compostas por complexos sistemas de circulação viária e de endereçamento que, apesar de já apresentarem uma forma automática de manipulação de dados implementada pelos sistemas de georeferenciamento, ainda assim não oferecem muitas formas de divulgação e utilização desses dados para o usuário final [48].

Várias aplicações foram desenvolvidas utilizando esse acervo de informações pela *web* e muitas delas oferecem várias facilidades como visualização de mapas, traçamento de rotas, localização de serviços de acordo com a proximidade, além de outros. Com o surgimento dos diferentes tipos de dispositivos móveis de acesso a dados a demanda por aplicações desse tipo está aumentando ainda mais pois a mobilidade dos usuários no sistema móvel revela uma demanda natural de serviços que os atenda levando-se em consideração a sua própria localização física.

A possibilidade de adição de informações georeferenciadas aos serviços que forem disponibilizados para dispositivos móveis poderá permitir buscas do tipo "Qual o cinema mais

próximo de um determinado local X , onde possivelmente X é o local corrente". Para se implementar tal tipo de serviço é necessário utilizar informação específica para que sejam obtidos os parâmetros de georeferenciamento. Isso pode incluir soluções bastante diferentes em função dos recursos de hardware disponíveis no dispositivo, como por exemplo o **GPS** (Global Positioning System), informações que podem ser obtidas da companhia operadora da rede sem fio (célula ativa, etc.) ou do próprio usuário. "Um dos pontos chave em tais serviços é a transparência de acesso. A informação solicitada passa a ter um valor semântico implícito facilitando a consulta e diminuindo o tamanho da mensagem. Do ponto de vista das aplicações os desafios envolvidos abrangem desde novos esquemas de endereçamento de serviços baseados na localização, até aspectos da infra-estrutura e tráfego da rede sem fio"[36].

Com o atual avanço e disponibilização de tecnologias de acesso à informação da *web* através de dispositivos portáteis como celulares, palms, smartphones e outros, considera-se que o desenvolvimento de aplicações desse tipo é de alta relevância no cenário mundial. Além disto, a necessidade de se desenvolver sistemas que se adaptem ao ambiente móvel já é bastante reconhecida na literatura [4, 5, 12, 16, 23, 24, 25, 32, 33, 35, 40, 42, 45, 46].

Como será visto na próxima seção, a facilidade em oferecer suporte automático para o gerenciamento de informações georeferenciadas motivou a adoção de um **SIG** (Sistemas de Informação Geográficos) para a solução de serviços. Informações relacionadas a endereçamento e orientação em rotas considerando sentidos, conversões e outras ocorrências espaciais podem ser facilmente representadas. Além disso, o SIG também suporta a associação de imagens a informações geográficas como mapas indicando a localização de endereços e vias de locomoção. Considera-se que estas funcionalidades são de grande interesse para os usuários pois estes podem explorar informações dependendo de onde estiver usando a aplicação através de uma interface amigável.

Diante das possíveis funcionalidades que podem ser oferecidas pelos serviços, deve-se inicialmente considerar os problemas relacionados ao ambiente móvel pois trata-se do principal obstáculo para a disponibilização destes. Como já visto anteriormente, existem diversos problemas ligados ao ambiente e também ao dispositivo de comunicação móvel que fazem com que sistemas e aplicações existentes não possam ser usados em aplicações móveis. Uma das soluções encontradas para tratar este tipo de problema foi a aplicação de algumas técnicas de adaptação, como descrito na seção 4.6.

1.1.1 Sistemas de informação geográficos

O acesso rápido a informações corretas e confiáveis é de extrema importância para as pessoas. Por serem geralmente encontradas em grande volume, qualquer base de informação pode ser gerenciada de forma eficiente e confiável através de sistemas de informação utilizando-se o computador. “O gerenciamento de sistemas de informação é uma das principais áreas de desenvolvimento da informática” [13].

Informatizados ou não, os sistemas de informação são concebidos de modo a representar o mundo real através de entidades e objetos. Em qualquer caso, é necessário um modelo simplificado que seja adequado ao uso por meio de computadores levando-se em consideração somente as características que sejam indispensáveis a essa representação. Além disso, essas formas de representação geralmente possuem uma forma de referência espacial pois constata-se que a maioria das entidades do mundo real tem uma localização espacial.

Apenas com o advento dos SIG tornou-se possível realizar análises e pesquisas apoiadas na distribuição espacial das informações. Para facilitar o entendimento do termo SIG faremos uma tentativa de definição do mesmo. Uma definição bem abrangente seria: “*SIG são sistemas automatizados usados para armazenar, analisar e manipular dados geográficos, ou seja, dados que representam objetos e fenômenos em que a localização geográfica é uma característica inerente à informação e indispensável para analisá-la*” [14]. Sempre que a questão da localização for importante nos problemas que precisam ser resolvidos por um sistema informatizado, existe uma possibilidade de se considerar a adoção de um SIG.

1.1.2 Adaptação

Aplicações que podem alterar seu comportamento de acordo com certas restrições impostas pelo ambiente são chamadas de aplicações adaptativas [4]. De modo geral, estas aplicações têm como objetivo oferecer ao cliente a informação da forma como a mesma se encontra no servidor e o usuário assim também o exige. A este tipo de compromisso dá-se o nome de *fidelidade* [42, 46]. Fidelidade é o grau de compromisso entre a cópia de um dado apresentada ao cliente em relação ao dado original que se encontra em um servidor. Em aplicações típicas da rede fixa estes requisitos são facilmente atendidos pois a disponibilidade de recursos é relativamente constante durante a comunicação, desde que estabelecidos os parâmetros de qualidade de serviço. Em contrapartida, as aplicações da rede móvel terão que se adequar dinamicamente às diversas restrições do ambiente de forma a oferecer um parâmetro mínimo de qualidade de serviço aos usuários.

Os níveis de fidelidade a serem oferecidos pela aplicação não podem ser determinados unicamente pelo tipo do dado pois muitos aspectos do processo de adaptação são desenvolvidos pela aplicação. Dependendo da natureza da aplicação irão existir diferentes níveis de fidelidade relacionados com a apresentação dos dados. Por exemplo, se uma aplicação de vídeo fosse específica para edição ao invés de visualização, diminuir a taxa de quadros seria mais apropriado do que simplesmente retirar alguns quadros para preservar a taxa de exibição. Desta forma a adaptação pode ocorrer tanto no dado quanto no próprio resultado de uma computação.

Segundo André [4], Friday [16] e Satyanarayanan [45] as propostas existentes podem ser divididas em duas principais categorias:

- Soluções que oferecem mecanismos básicos que permitem que as aplicações se adaptem a diferentes condições de execução. As aplicações deste tipo são responsáveis pelos mecanismos de detecção do contexto e de adaptação.
- Soluções que oferecem mecanismos mais avançados para monitorar o contexto de execução e se adaptarem às mudanças. Estas soluções são consideradas avançadas pois implementam um subsistema de detecção e notificação do contexto, políticas de acesso a dados e outras funcionalidades relacionadas ao gerenciamento dos recursos pelo sistema que oferecem um nível de transparência às aplicações.

Em alguns casos estas duas abordagens são combinadas [25] para oferecer maior flexibilidade ao sistema como um todo.

1.2 Objetivos

O objetivo deste trabalho é estudar as principais considerações na disponibilização de aplicações baseadas na localização de usuários no ambiente móvel. Para tanto, partiu-se inicialmente para a pesquisa das principais tecnologias que possibilitam a localização de usuários em uma rede de comunicação sem fio. Foi feito também um levantamento sobre as implicações de uso deste tipo de aplicação no ambiente móvel. Considerando-se o paradigma da computação móvel, o desenvolvimento de qualquer tipo de aplicação implica na consideração da influência do ambiente sobre a operação da mesma. A partir destes princípios, propôs-se uma arquitetura do tipo cliente-proxy-servidor [33] que trata dos aspectos de adaptação relacionados à capacidade das aplicações se adequarem às condições

do meio. Como será visto mais adiante no texto da dissertação, a arquitetura é formada por módulos e cada um dos módulos possui uma função específica.

O trabalho também descreve a especificação e modelagem de uma aplicação, da arquitetura proposta e as soluções de adaptação que foram empregadas. Como resultado do estudo procurou-se identificar a eficácia do uso das soluções de acordo com alguns cenários. Diferentes combinações das características do meio sem fio foram utilizadas através de simulações para caracterizar alguns dos possíveis cenários.

A arquitetura proposta é uma importante característica do trabalho. Na literatura existem várias propostas que tratam de arquiteturas de adaptação para aplicações em geral, como já visto na seção 4.6. Grande parte destes trabalhos considera os aspectos relacionados principalmente ao gerenciamento eficiente dos recursos do sistema através de um *middleware* que controla a comunicação entre o cliente e o *proxy*. O *middleware* é responsável por monitorar e controlar a alocação de recursos do sistema que juntamente com a camada superior de aplicação ficam responsáveis por executar as devidas adaptações.

Estas plataformas consideram a existência de recursos de sistema implementados em todos os clientes. Na verdade, o que existe atualmente são diferentes dispositivos com recursos heterogêneos e que praticamente não oferecem suporte a aplicações como os dispositivos que utilizam a tecnologia WAP por exemplo. Os clientes WAP mais simples implementam recursos de navegação através de *browser* específico que executa requisições a um servidor na rede fixa oferecendo páginas com texto e imagens de baixa qualidade como resposta.

O trabalho aqui descrito procura contemplar a disponibilização dos serviços para dispositivos WAP e também para dispositivos que têm maior poder de processamento como os dispositivos 3G que irão surgir no mercado nos próximos anos. Cada um destes clientes implementa diferentes tipos de recursos e a especificação de um modelo único de desenvolvimento de aplicações torna inviável a disponibilização destas para a gama heterogênea de equipamentos existentes.

1.3 Trabalhos relacionados

Por ser uma área relativamente recente, várias pesquisas em computação móvel estão ainda sendo desenvolvidas. O estudo do problema de oferta de serviços baseados na localização do usuário em ambientes móveis conta com relativamente poucas contribuições do meio acadêmico mas está sendo bastante citado atualmente devido aos constantes avanços

da tecnologia de equipamentos e da infra-estrutura de acesso a rede sem fio.

Em Acharya *et al.* [3] é definido um conjunto de primitivas para o desenvolvimento de aplicações baseadas na *web* dependentes da localização de usuários no ambiente móvel. Foi desenvolvido um serviço de navegação inteligente em um prédio onde eram disponibilizadas informações como plantas dos pavimentos, serviços de direção, instalações elétricas e hidráulicas para os usuários.

Este trabalho especifica um arcabouço composto por elementos que implementam algumas funções para disponibilizar serviços baseados na localização. Estas funções tratam desde os elementos para determinação da localização dos usuários na rede até mesmo do sistema de informação que oferece documentos baseados na informação de localização através da definição de algumas especificações.

Estas especificações envolvem alterações nos clientes e servidores HTTP, nas sintaxes das URLs e a implementação de um protocolo de *multicast* para localização de servidores mais próximos. O sistema de informação utilizado é organizado como uma coleção de páginas *web* e a interface para o usuário é baseada em *hyperlinks*.

Em resposta à mensagem de *multicast* gerada por um cliente, o servidor mais próximo envia sua informação de localização que será utilizada nas requisições URL subsequentes. Cada página devolvida como resposta está associada a um escopo. O escopo é uma lista de estações base em que o documento é válido. Desta forma a localização do cliente estará sempre associada ao último escopo pesquisado. Caso a localização do usuário seja alterada a interface cliente irá checar a validade de todos os documentos e invalidar somente aqueles que o escopo não corresponde a atual localização.

Algumas considerações sobre as restrições dos dispositivos móveis e do meio sem fio foram feitas mas as soluções para o tratamento destes problemas foram baseadas em estratégias relacionadas aos protocolos e a mecanismos de entrega de páginas pelos servidores. Além disso, a metodologia de especificação do ambiente e da aplicação diferenciam-se em vários aspectos em relação ao que se propõe neste trabalho de dissertação.

A modificação de especificações existentes ou mesmo a proposta de novas especificações para o uso do serviço são alguns dos pontos desfavoráveis encontrados nesse trabalho. As funções de adaptação propostas baseiam-se no conceito de *hyperlinks* adaptáveis a recursos. Esses *hyperlinks* são uma generalização da funcionalidade do HTML em exibir textos alternativos ao invés de imagens no caso de *browsers* que suportam somente texto. Cada página é associada a um conjunto de pré-requisitos de recursos que devem ser satisfeitos no cliente para que o mesmo possa recebê-la. Estes pré-requisitos são relacionados tanto a

restrições de recursos do cliente como do meio de comunicação. Neste trabalho não foram encontrados registros de testes relacionados a estes recursos de adaptação em específico e desta forma nenhum resultado foi explicitamente relatado.

Em Voelker and Bershad [52] é proposto um serviço de informação *World Wide Web* para usuários do ambiente móvel, chamado **Mobisaic**. O Mobisaic estende a *web* permitindo que os documentos possam referenciar e reagir a mudanças de contexto como a alteração da localização do usuário pela rede sem fio. Para suportar o uso de informação dinâmica, a infra-estrutura *web* existente foi modificada para incluir os seguintes recursos: um servidor de rede que mantém “contextos” dentro de um domínio específico do usuário, um mecanismo de *callback* assíncrono para a notificação dos clientes *web* quando houver alguma mudança nas variáveis de contexto dinâmicas e uma sintaxe para referenciar informação dinâmica em URLs e documentos.

O conceito de variáveis de ambiente, como as utilizadas em um *shell*, é utilizado para representar os contextos de usuários mas neste caso as variáveis são associadas a usuários ou lugares ao invés de processos de *login* como é feito normalmente num shell. Os contextos representam o ambiente de onde o usuário atualmente acessa o serviço e uma das variáveis de contexto, por exemplo, informa a localização física do usuário. Para representar esta informação o sistema conta com um mecanismo de publicação para os clientes da rede.

A notificação de modificações nos contextos dos usuários é feita através de documentos ativos. Os documentos ativos permitem que mudanças no ambiente sejam refletidas na informação visualizada pelo usuário. As notificações contêm os nomes das variáveis que sofreram alterações e desta forma o cliente pode utilizar os novos valores das variáveis para as requisições subsequentes, refletindo assim as mudanças. Estas novas requisições podem ser feitas através de URLs modificadas para aceitar as variáveis de ambiente como parâmetro. O sistema de informação é baseado em páginas HTML modificadas para interpretar as variáveis de ambiente que são passadas através da URL.

Ambos os trabalhos citados anteriormente se baseiam principalmente na alteração de clientes *web*, na alteração da sintaxe URL e no conceito de documentos ativos através de modificações nas páginas HTML. Os serviços disponibilizados têm como base de acesso documentos em linguagem de marcação e não existem mecanismos de adaptação eficientes para tratar dos problemas no ambiente móvel.

Existem também algumas iniciativas provenientes do mercado que utilizam de métodos próprios. Empresas como SnapTrack [37], Cellpoint [10] e Qualcomm/Lucent [43] anunciaram nos últimos dois anos suas soluções propostas e que já estão sendo testadas e utilizadas

em alguns países da Europa e nos Estados Unidos. Estas soluções são todas baseadas em tecnologias como GSM (Global System for Mobile communication) e GPS (Global Positioning System). O que estas empresas propõem são soluções de serviços baseados na localização do usuário que levam em consideração desde a aplicação do processo de obtenção da informação de localização do usuário até possíveis modificações na infra-estrutura de rede e nos próprios dispositivos.

A plataforma Ericsson Mobile Positioning System (MPS) descrita em [49], não requer modificações. Esta solução inclui servidores que permitem que serviços de posicionamento sejam utilizados em redes do padrão GSM com os mais diversos métodos de posicionamento. Além do subsistema de posicionamento, esta solução também é composta por um subsistema *gateway* (MPC - Mobile Positioning Center) que funciona como um mediador entre a rede celular e o cliente convertendo dados do subsistema de posicionamento em informações de posicionamento. As aplicações são registradas no *gateway* MPC.

A solução da SignalSoft [47] apresenta uma plataforma modular, semelhante a apresentada anteriormente, composta por produtos específicos. Os dois principais produtos são o **Location Manager** e o **local.info** que são responsáveis respectivamente pelo gerenciamento das várias técnicas de localização e pela combinação de informações de localização com conteúdo.

Em [34] é descrita uma estrutura de serviços baseados na localização proposta pela Oracle. Esta estrutura é formada por dois módulos principais: o Oracle Spatial que é um SGBD estendido para gerenciar dados espaciais e o Oracle Portal to Go, um servidor que combina serviços de transformação avançada de dados com a funcionalidade de um portal para usuários móveis permitindo que sejam oferecidos vários tipos de conteúdo, inclusive aplicações baseadas em XML (eXtensible Markup Language), como o WML (Wireless Markup Language). Esta solução da Oracle oferece suporte a aplicações e ferramentas para gerenciamento de informações de localização em uma estrutura de n camadas aberta.

A solução da Webraska [54] oferece uma API (Application Programming Interface) que agrega como principais funcionalidades: rota pelo caminho mínimo diferenciada pelo meio de locomoção; exibição de mapas pré-definidos com informações sobre tráfego em tempo real; *ranking* dos serviços mais próximos do usuário; além de outras funções relacionadas. As informações manipuladas por estas funcionalidades são todas gerenciadas em um banco de dados. Todas as informações geradas por essas funções são adaptadas para a visualização adequada aos diferentes dispositivos. A API que permite o uso destas funcionalidades é acessada via requisições HTTP ao servidor da Webraska que pode ser acessada por outros

servidores ou mesmo aparelhos celulares. Após uma requisição a este servidor, a resposta é gerada em formatos específicos como XML ou WML, dependendo do cliente.

1.4 Organização da dissertação

Este trabalho de dissertação é dividido em seis capítulos. O capítulo 1 é composto pela motivação, considerações sobre os sistemas de informação geográficos, considerações sobre adaptação em ambientes móveis e trabalhos relacionados.

O capítulo 2 apresenta as características gerais das tecnologias de posicionamento de usuários. São descritos alguns serviços sendo desenvolvidos no mercado, as diferentes formas de precisão de posicionamento e suas implicações, as principais tecnologias utilizadas atualmente pelas companhias operadoras e por fim uma discussão sobre estratégias de localização em uma rede celular.

O capítulo 3 descreve a arquitetura proposta no trabalho. Nesse capítulo, faz-se algumas considerações sobre o ambiente móvel e sobre o modelo de arquitetura sendo utilizado.

O capítulo 4 apresenta uma discussão sobre a modelagem de dados feita para o sistema de georeferenciamento utilizado e sobre as soluções de implementação empregadas no desenvolvimento da arquitetura, do protótipo da aplicação e do simulador. É feita uma explanação sobre o emprego de técnicas de compressão e de parâmetros de qualidade de serviço no desenvolvimento da arquitetura e da aplicação.

O capítulo 5 apresenta os testes feitos sobre a arquitetura e o protótipo da aplicação, onde são avaliados alguns parâmetros de variação de condições e de definição de parâmetros de qualidade de serviço. Apresentamos também uma análise dos resultados obtidos.

No capítulo 6 são apresentadas as conclusões e considerações sobre trabalhos futuros.

Capítulo 2

Sistemas de localização

Produtos relacionados a localização farão parte de uma das maiores classes de serviços de valor agregado que serão oferecidos pelas operadoras de telefonia móvel. Os serviços a serem disponibilizados não serão necessariamente novos mas também poderão ser melhorias nos atuais serviços tais como os serviços pré-pagos baseados na localização ou então serviços de informação.

É claro o fato de que a localização de um usuário em uma rede móvel celular é um processo muito mais complicado do que a localização de um usuário na rede de telefonia fixa. Enquanto os telefones fixos possuem um endereço também fixo, os telefones móveis são caracterizados pelo fato de que eles podem estar em qualquer lugar, podendo mudar de localização a qualquer momento. Além disso, a localização de terminais móveis está sujeita a problemas como as intempéries naturais de um ambiente de comunicação via rádio e também à presença de diferentes padrões de comunicação celular atualmente disponíveis.

Entretanto, mesmo na presença de tais dificuldades, várias tecnologias de posicionamento têm sido ou estão sendo desenvolvidas. Vários produtos comerciais já estão sendo oferecidos principalmente no mercado americano, especialmente para os serviços de emergência **911** (seção 2.1).

Os serviços baseados na localização representam uma das principais categorias de aplicações que serão oferecidas aos usuários do ambiente móvel devido a sua característica de oferecer facilidades relacionadas a sua própria localização. Desta forma, estes serviços deverão ser capazes de obter a informação de posicionamento dos usuários de diversas maneiras, podendo-se utilizar desde uma descrição textual até sofisticadas técnicas de posicionamento. Serão descritas a seguir algumas das principais técnicas de posicionamento

e alguns dos fatores que estão impulsionando o desenvolvimento das mesmas.

2.1 Regulamentação do serviço 911

O FCC (Federal Communications Commission) dos EUA adotou uma lei em junho de 1996 que requer que todas as operadoras de telefonia móvel ofereçam informação de localização em todas as chamadas de emergência americano [15]. O FCC exige que a partir do dia primeiro de outubro de 2001, todas as chamadas 911 feitas por aparelhos sem fio devem ser rastreadas com uma precisão de 125 metros por 67% do tempo de uso. Em 24 de dezembro de 1998, o FCC complementou a lei permitindo o uso de soluções baseadas no terminal e também das baseadas na rede, explicadas na seção 2.4.

2.2 Serviços de localização

Com o advento da tecnologia de acesso à informação da *web* por dispositivos móveis a oferta de serviços baseados na localização neste tipo de ambiente é cada vez maior devido principalmente ao potencial dos sistemas de informação. A possibilidade de acesso à grande quantidade e variedade de dados que estão disponíveis aliada ao contexto do usuário de um ambiente móvel abre um leque enorme de possíveis aplicações que podem ser desenvolvidas. Neste trabalho está sendo considerando um tipo de aplicação que se caracteriza pelo acesso aos recursos oferecidos pelos sistemas de informação geográficos. É importante salientar que os serviços baseados na localização (LBS - *Location Based Services*) não se restringem somente aos serviços que utilizam de dados georeferenciados. Alguns serviços podem ser oferecidos somente com a utilização da informação de localização do usuário [26]. Dentre estes podem ser citados:

- Segurança: serviços de emergência, assistência rodoviária, etc.
- Cobrança: cobrança baseada na localização, escritório sem fio, uso residencial, cobrança por eventos, etc.
- Informação: informação de entretenimento, navegação, páginas amarelas, eventos, tráfego, etc.
- Rastreamento: gerenciamento de frotas, rastreamento de objetos e pessoas, etc.

- Apoio à gerência de redes: aplicações para planejamento de redes, gerenciamento de QoS, etc.

Existem atualmente diversas empresas que estão desenvolvendo e oferecendo plataformas, aplicações e tecnologias para suporte aos serviços baseados na localização. Estas empresas podem ser divididas em categorias de acordo com o tipo de produto que oferecem no mercado. No levantamento feito neste trabalho foram encontradas três principais categorias de produtos: tecnologias de posicionamento, plataformas/ *middleware* e aplicações para o usuário final.

As empresas que oferecem produtos relacionados a tecnologia de posicionamento trabalham basicamente com sistemas de posicionamento baseados nas tecnologias que serão citadas na seção 2.4. A principal meta dessas empresas é desenvolver sistemas e equipamentos que sejam precisos e adequados para os diferentes tipos de infra-estrutura existentes nas redes de telefonia móvel celular. Existem também plataformas de integração de sistemas de posicionamento e de aplicações que oferecem maior suporte quanto à flexibilidade do sistema. Grande parte dessas empresas garante a compatibilidade com qualquer sistema de posicionamento empregado além de oferecer toda a infra-estrutura de sistemas para o gerenciamento dos serviços. Por fim, encontram-se no mercado empresas que desenvolvem especificamente aplicações para o usuário final dedicando-se ao desenvolvimento de serviços diferenciados, APIs, *toolkits* e outras ferramentas.

2.3 Precisão de posicionamento

Diferentes aplicações baseadas na localização exigem diferentes níveis de precisão. Por exemplo, é evidente a importância para os navios saberem exatamente a distância até a costa além da profundidade da água. No caso de rastreamento de pessoas a precisão de centenas de metros já seria aceitável. A tabela 2.1 exibe algumas das várias aplicações baseadas na localização e mostra algumas das exigências de precisão para estes serviços inicialmente e quando já estiverem sendo empregados. Da esquerda para a direita, as colunas correspondem respectivamente ao tipo da aplicação, parâmetros iniciais de precisão, requisitos de precisão quando os serviços já estiverem em uso e os objetivos de cada serviço.

Guias de ruas e estradas, assistência rodoviária e rastreamento de frotas, objetos e pessoas irão cada um tornar-se preponderantes a partir do momento em que as soluções permitirem precisões maiores. O princípio para se incorporar precisão num serviço de

localização baseado em rede é a habilidade em combinar tecnologias de localização com outros métodos menos precisos como, por exemplo, Cell-ID (*Cell Identification*) de tal forma que as operadoras de rede possam atingir escalabilidade rapidamente. Além disso, elas também podem atender amplamente aos usuários atuais ao invés de atenderem somente aos assinantes com telefones mais avançados. Serviços de atendimento ao consumidor, de cobrança e de comutação podem ser integrados com aplicações de localização sem ter de se esperar pela ampla cobertura da rede ou com a total penetração de novos aparelhos. Sendo assim, os consumidores podem utilizar imediatamente os novos serviços de localização em toda a área de cobertura, desconsiderando as inconsistências inerentes às precisões de localização, além da própria disponibilidade.

<i>Aplicação</i>	<i>Parâmetros iniciais</i>	<i>Requisitos</i>	<i>Objetivos</i>
Cobrança	Célula/Setor	250m	Preço competitivo
Assistência em auto-estradas	500m	125m	Assistência ao usuário
Páginas amarelas	Célula/Setor	250m	Serviço mais próximo
Informação de tráfego	Célula/Setor	Célula/Setor	Fluxo de tráfego
Serviço de mensagens	Célula/Setor	125m	Avisos, alertas, informes
Rastreamento de frotas	Célula/Setor	30 - 125m	Gerenciamento de recursos
Rastreamento de objetos	Célula/Setor	Célula/Setor	Localização e direção
Rotas	125m	30m	Guia de rotas

Tabela 2.1: Tabela de requisitos de precisão de serviços

2.4 Tecnologias de posicionamento

Nos Estados Unidos, os aspectos legais têm agido como uma força de impulsão para a padronização de tecnologias de posicionamento do padrão GSM. O órgão mais atuante desta batalha tem sido o FCC. Organizações que representam brigadas de fogo, hospitais e outros serviços de emergência têm também participado na formulação dos requisitos da FCC e tanto a indústria como autoridades têm concordado com o andamento das padronizações para os sistemas de posicionamento que irão servir o mercado. As exigências ainda não foram completamente finalizadas mas já foi especificada uma descrição com as principais características destas exigências.

Como resultado foram definidas algumas regulamentações do FCC que diferenciam os sistemas baseados na localização entre soluções *terminal-based* (baseadas no dispositivo) e soluções *network-based* (baseadas na infra-estrutura de rede) [49]. A combinação dos dois tipos de soluções (soluções híbridas) é também utilizada pois ela pode combinar os benefícios de ambas as técnicas enquanto limita algumas de suas desvantagens [7].

As soluções de posicionamento baseadas no terminal estão relacionadas à participação ativa do terminal nos cálculos de posicionamento. Estas técnicas normalmente exigem que sejam instalados *softwares* e/ou *hardwares* específicos tanto no terminal quanto na infra-estrutura da rede. Na prática, isso significa que uma vez instalados os sistemas, os assinantes terão que adquirir novos aparelhos para se beneficiarem da tecnologia. A penetração no mercado terá um crescimento de forma gradual assim que os equipamentos forem substituídos em um período que pode durar de quatro a cinco anos. Como exemplo de soluções baseadas no terminal podemos citar o o GPS, o A-GPS (*Assisted GPS*) [19], o DGPS (*Differential GPS*), o STK ou SIM Toolkit (*Subscriber Identity Module Application Toolkit*) [21] e E-OTD (*Enhanced Observed Time Difference*) [19].

Ao contrário das soluções baseadas no terminal, as soluções de posicionamento baseadas na rede exigem uma inteligência intrínseca dos equipamentos da infra-estrutura da rede, significando que a penetração de mercado poderá ser maior desde o primeiro dia de lançamento do sistema. Como exemplos de soluções baseadas na rede podemos incluir os métodos de CGI (*Cell Global Identity*), também conhecido como Cell-ID [19], e TOA (*UpLink Time Of Arrival*) [19]. Devido às diferentes características entre os sistemas de posicionamento baseados em rede e baseados no terminal o FCC estipulou requisitos separados para cada um deles. Atualmente os únicos métodos que satisfazem estes requisitos são a solução TOA, que é representante dos sistemas baseados em rede, e a solução A-GPS, que é representante dos sistemas baseados no terminal.

O comitê T1P1, um sub-comitê do ANSI (*American National Standards Institute*) [6], e o ETSI (*European Telecommunications Standards Institute*) [22] se reuniram para a padronização de sistemas de localização usando E-OTD, TOA e A-GPS, além do já existente CGI.

Para se referenciar de forma homogênea ao método de identificação de células em rede celular será utilizado o termo Cell-ID.

2.4.1 Soluções baseadas no terminal

Existem vários sistemas disponíveis para o propósito de localização de terminais. Entre eles estão o TDOA (*Time Difference Of Arrival*), o E-OTD, Bluetooth, tecnologias baseadas em satélites, além de outros. Usuários de um sistema global de posicionamento como o GPS podem calcular sua localização em qualquer lugar da superfície terrestre.

Existem atualmente dois sistemas públicos de GPS. O sistema NAVSTAR é de propriedade do governo americano e é administrado pelo departamento de defesa. O sistema GLONASS é de propriedade da federação russa. Embora ambos os sistemas são para posicionamento global, o sistema NAVSTAR é geralmente referido (principalmente nos Estados Unidos) como GPS porque ele foi disponibilizado primeiro. Os dois sistemas oferecem dois conjuntos de sinais de posicionamento.

TDOA

O UE (*User Equipment*) faz o rastreamento das BTSs (*Base Transceiver Station*) que o cercam para definir os TDOAs dessas diferentes BTSs que são acessadas por ele. Estas medidas, juntas com outras informações relativas a localização geográfica dos transmissores e o RTD (*Relative Time Difference*) da atual transmissão dos sinais de *downlink* podem ser usadas para calcular uma estimativa da localização do UE. Cada medida TDOA para um par de transmissões de *downlink* descreve uma linha de constante diferença na qual o UE pode ser localizado. A localização do UE é determinada pela interseção destas linhas para pelo menos dois pares de estações base.

As medidas TDOA primárias (feitas pelo UE) são enviadas ao PCF (*Position Calculation Function*) no SRNC (*Serving Radio Network Controller*). O PCF precisa saber das medidas TDOA, das localizações geográficas das estações base que tiveram seus sinais medidos (esta informação é obtida do LSIF - *Location System Information Function*) e o RTD (*Real Time Difference*) atual entre as transmissões das estações base no momento que as medidas TDOA foram feitas. Há diferentes formas de se determinar o RTD. Uma primeira se dá pela sincronização das transmissões das estações base. Os tempos de transmissão podem ser todos alinhados a uma referência comum (como o UTC) no caso em que todos os RTD têm um valor comum. Entretanto, as transmissões podem geralmente ter um *offset* fixo de referência ao UTC e desta forma os valores RTD são diferentes de zero e podem ser armazenados em um banco de dados para uso das funções de cálculo.

A localização do UE pode ser calculada tanto na rede (UE *assisted*) como no próprio UE (UE *based*). No modo UE *assisted*, o UE mede a diferença nos TDOA das células e

envia as medidas obtidas para a rede onde um elemento se encarrega de fazer o cálculo da localização (PCF). No modo UE *based*, o UE faz as medidas e também se encarrega dos cálculos de localização. Para isso ele precisa de informação adicional (como a localização das estações base) para o cálculo.

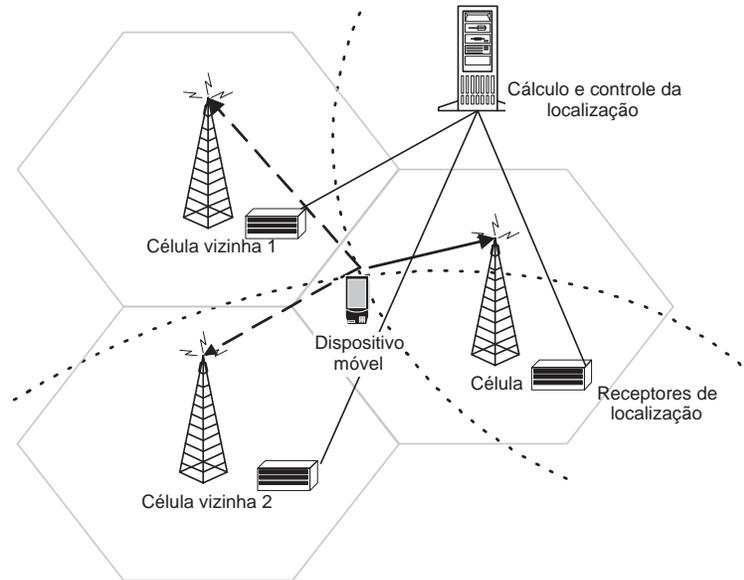


Figura 2.1: Método TDOA

E-OTD

Sistemas E-OTD operam através do uso de receptores de localização ou sinalizadores de referência, distribuídos numa rede celular como uma LMU (*Location Measurement Unit*) em múltiplos locais dispersos geograficamente em uma ampla área.

Cada um destes sinalizadores tem uma fonte precisa de temporização e quando um sinal de no mínimo três estações base (triangulação) é recebido por um terminal móvel com o software E-OTD e pela unidade de medida de localização, as diferenças de OTD (*Observed Time Difference*) provindos de cada BTS no aparelho e na unidade de medida de localização são calculados. Desde que os quadros de transmissão das BTS não são sincronizados, a rede deve medir a RTD. As diferenças em marcas de tempo são então combinadas para se produzir a interseção de linhas hiperbólicas de onde a localização é estimada. Para obter uma triangulação precisa, as medidas de OTD e RTD são necessárias para pelo menos três BTS distintas separadas geograficamente.

Sistemas E-OTD podem entretanto estar expostos a vários dos mesmos problemas de ruído no sinal (multicaminhamento) como sistemas de AOA (*Angle of Arrival*). Neste caso, o multicaminhamento distorce os sinais e o grupo de retardos, causando algumas dificuldades pelo sistema E-OTD em determinar precisamente o sinal a ser medido por todos os receptores.

Estas soluções oferecem maior precisão no posicionamento do que a tecnologia Cell-ID, entre 50 e 200 metros, mas têm um tempo de resposta menor, aproximadamente 5 segundos, e exigem equipamentos com software modificado, o que significa que eles não podem ser usados para oferecer serviços de localização específicos à base de clientes.

O método E-OTD pode ser:

- assistido pela rede (*network assisted*), caso em que o dispositivo móvel mede o sinal OTD e computa sua própria localização (para tal a rede deve dar ao dispositivo informações adicionais como as coordenadas das BTS e valores RTD); ou
- assistido pelo dispositivo (*handset assisted*), caso em que o dispositivo móvel mede o sinal OTD e envia suas medidas para a rede que então computa a localização do terminal.

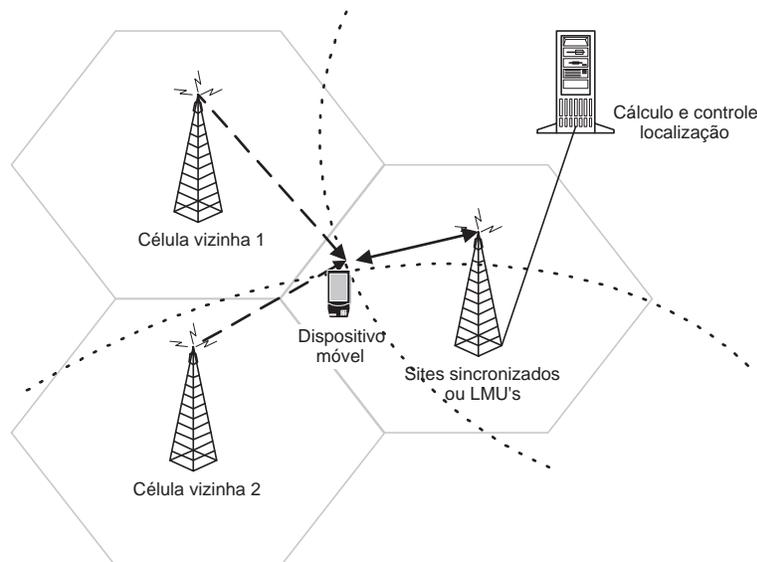


Figura 2.2: Posicionamento *Enhanced Observed Time Difference*: posicionamento hiperbólico

Tecnologia Bluetooth

O Bluetooth [7, 9] é uma interface de rádio universal que opera na faixa de frequência de 2.45 GHz que permite que dispositivos eletrônicos portáteis se conectem e se comuniquem em pequenas distâncias. Cada unidade pode se comunicar simultaneamente com sete ou mais outras unidades em uma *piconet*.

O usuário com um dispositivo Bluetooth pode obter sua informação de localização graças a pontos de acesso fixo ou através de um segundo dispositivo móvel que tenha estado bem próximo de outro dispositivo recentemente (até a faixa de 30 metros). Desta forma, é necessário que se tenha ponto de acesso fixo para esse tipo de acesso. A técnica de localização Bluetooth é baseada na proximidade (as medidas de posicionamento serão as coordenadas do ponto fixo mais próximo ao dispositivo). Para serviços de localização, espera-se que se tenha uma infra-estrutura fixa nos centros urbanos, *shopping centers*, estações e aeroportos, locais com atrações turísticas, etc. Estas áreas são as mais problemáticas quanto a questão de precisão de localização. Por isso, uma solução Bluetooth nestas áreas poderia oferecer um posicionamento de alta precisão se a infra-estrutura fixa fosse implementada de forma correta.

O SIG (Bluetooth Special Interest Group) identificou 5 diferentes cenários de uso da técnica de localização que seguem abaixo:

- Um dispositivo Bluetooth determina sua localização através de um ou mais dispositivos. Por exemplo, uma loja que possui uma rede de dispositivos Bluetooth para oferecer informações para funcionários e clientes provendo assim um serviço de localização *indoor*.
- O dispositivo aprende e lembra de sua posição. Quando dois ou mais dispositivos se cruzam eles "aprendem" a sua localização devido ao fato de que alguns destes dispositivos podem ter mecanismos de auto-localização como GPS embutido ou então posicionamento pela rede celular e o dispositivo então calibra a infra-estrutura. As localizações são "lembradas" assim que os dispositivos permanecem em contato e se tornam disponíveis a qualquer outro novo dispositivo.
- Interpolação adaptativa entre dispositivos. Em um prédio que contém vários dispositivos Bluetooth, a informação se propaga entre os dispositivos que incrementalmente interpolam suas posições individuais.

- Economia de energia. Um GPS é levado a um ambiente Bluetooth e troca a informação de localização com outros dispositivos. O processo de posicionamento é executado e depois o GPS pode ser desligado ou atualizado sem muita frequência para economizar a energia dos dispositivos.
- Telefones móveis com Bluetooth usados em um carro (com GPS já embutido). Quando os sinais GPS não estão disponíveis, o aparelho celular pode usar o sistema de posicionamento celular e passar a informação de posição para o display do GPS utilizando Bluetooth. Para aplicações que usam informação de tempo real e aplicações de rastreamento, a informação de localização do carro é passada do GPS para o telefone através do Bluetooth e depois para a rede celular.

NAVSTAR/GPS

O NAVSTAR/GPS (GPS) é um sistema de posicionamento espacial baseado em rádio que oferece informação de localização em três dimensões, velocidade e tempo 24 horas por dia em qualquer lugar. O sistema consiste de 26 satélites operados pela força aérea americana sob a supervisão do departamento de defesa, de uma central de comando no Colorado. A qualquer momento pelo menos quatro satélites são visíveis de qualquer posição na superfície da Terra.

Um sistema de posicionamento global usa as características de transmissão de rádio para determinação de localização. Ao contrário de sistemas de navegação anteriores que usavam transmissores terrestres, os transmissores baseados em satélites são usados para cobrir a superfície terrestre com maior precisão.

É necessário um receptor de rádio especial, um receptor GPS, para receber as transmissões dos satélites. Os receptores GPS contêm um computador especializado que calcula a localização baseado nos sinais do satélite. O usuário não tem que transmitir nada para o satélite e o satélite não sabe aonde está o usuário. Não há limite de número de usuários que podem usar o sistema ao mesmo tempo.

Existem vários fatores que afetam a eficiência dos serviços de localização, especialmente quando se trata de serviços de emergência. Dentre estes fatores estão a precisão, tempo de resposta, tempo de correção, cobertura do serviço e integridade. As maiores vantagens dos sistemas baseados em GPS são a cobertura global e a grande precisão, além do impacto mínimo nas redes de comunicação existentes. Por outro lado, existem problemas relativos a potência fraca do sinal de GPS dentro de edificações, o consumo de potência e o tempo

de correção na inicialização do receptor, embora os dois últimos não afetam aplicações em veículos.

Differential GPS

A proposta do *Differential GPS* é aumentar a precisão de posicionamento do GPS padrão corrigindo ou reduzindo bastante alguns tipos de erro que afetam a solução GPS em particular: *selective availability* (processo que degradava a precisão do sinal retirado em maio de 2000), erros ionosféricos, erros troposféricos e o *clock* do satélite.

Isto pode ser conseguido levando-se em conta que sobre *Baselines* limitadas entre uma estação de referência e um receptor, os erros serão os mesmos em qualquer localização. Este método faz uso de dois receptores GPS: um estacionário e localizado em um ponto conhecido (receptor de referência), o outro é operado como um receptor móvel. Desde que as coordenadas dos receptores de referência sejam conhecidas, pode-se corrigir a pseudo-faixa do satélite para a faixa real. Faixa real menos a pseudo-faixa é igual a correção diferencial. Através da comunicação entre os receptores de referência e fixos, a pseudo-faixa dos receptores móveis são corrigidas para as faixas reais. Existem basicamente dois métodos que aplicam correções diferenciais e que conseguem uma melhor precisão de posicionamento dos receptores móveis: no modo direto a correção diferencial é difundida para os receptores móveis (por exemplo, através de uma rede GSM), e o segundo método chamado de DGPS (*Differential GPS*) inverso consiste da transmissão de medidas do receptor móvel para as estações de referência, aonde a correção de posicionamento é feita.

A maior vantagem do posicionamento DGPS comparada com o GPS padrão é a alta precisão que pode ser alcançada (de 1 a 7 metros). Entretanto, o DGPS tem a desvantagem natural dos sistemas de posicionamento por satélite (fraca potência do sinal dentro de prédios, etc.).

Assisted GPS

O GPS é usado normalmente para propósitos de navegação. Unidades *stand-alone* de GPS são frequentemente encontradas em automóveis, barcos e aviões. Uma unidade GPS recebe sinais de quatro ou mais satélites. Cada sinal contém uma marca de tempo (*time stamp*) e uma descrição da posição do satélite. Comparando-se estas informações, a unidade GPS pode calcular sua própria posição. O principal obstáculo no funcionamento do GPS é que os sinais dos satélites são relativamente fracos e podem talvez nem oferecer a cobertura adequada de todos os ambientes. Entretanto, a rede GSM pode oferecer informações de

assistência que dá aos receptores GPS integrados melhor cobertura do que receptores GPS *stand-alone*.

O método *Assisted* GPS é basicamente uma variação do GPS diferencial com o uso da rede celular (GSM/PCS) para uso de sinais adicionais. Isto implica numa arquitetura distribuída: um receptor GPS de referência, um servidor de localização e um terminal móvel com GPS embutido (um telefone celular por exemplo).

Qualquer sistema de posicionamento GPS tem quatro funções principais: determinar as fases de codificação (pseudo-faixas) para os vários satélites; determinar o tempo de aplicação das pseudo-faixas; demodulação da mensagem de navegação do satélite; cálculo da posição da antena receptora que utiliza estas pseudo-faixas, e dados das mensagens de navegação e temporização.

A maioria dos receptores GPS comerciais executam todas estas operações sem assistência externa. Nestes receptores convencionais, a mensagem de navegação dos satélites e sua sincronização são extraídos de sinais GPS depois que foram reastreados e obtidos. A coleta desta informação normalmente leva de trinta segundos a vários minutos, e espera-se um alto nível do sinal recebido.

Em um sistema *Assisted* GPS estas funções são distribuídas entre um receptor de referência GPS, um servidor de localização e um dispositivo sem fio com GPS embutido.

A figura 2.3 ilustra o método de posicionamento A-GPS.

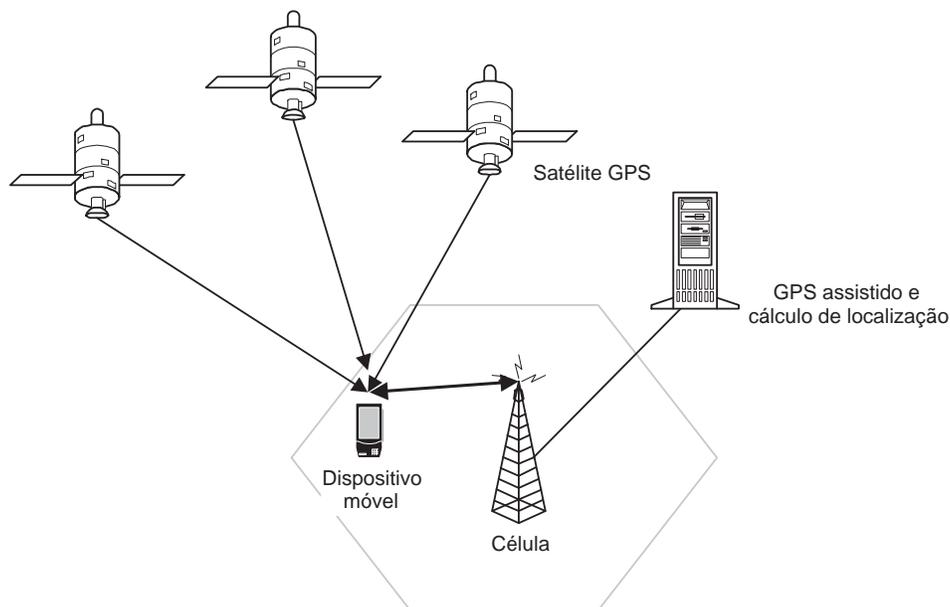


Figura 2.3: Posicionamento *Assisted* GPS

2.4.2 Soluções baseadas na rede

Para se utilizar os serviços baseados na localização, as soluções baseadas na rede não exigem alteração dos terminais. A seguir serão descritos os métodos Cell-ID, TA (*Timing Advance*) [19], TOA e AOA.

O padrão Cell-ID e três métodos Cell-ID proprietários (não padronizados) serão descritos. Os métodos proprietários recuperam a informação de identificação da célula de diferentes maneiras. A primeira utiliza o SIM Toolkit no terminal para enviar as informações via SMS (*Short Message System*), a segunda obtém a informação através do sistema IN (*Intelligent Network*) e a terceira utiliza uma conexão WAP para a transmissão da informação. Em todos os métodos citados, os resultados NMR (*Network Measurement Results*) podem ser adquiridos para aumentar a precisão de posicionamento.

O método TA é um método utilizado em todos os métodos de posicionamento sendo utilizado principalmente como um procedimento de segurança. Ele é padronizado pela ETSI e é baseado no parâmetro TA do GSM. O método TOA também é padronizado pela ETSI e lida com medidas do tempo de propagação de sinais em pelo menos três BTS. O método AOA não é padronizado ainda e depende fortemente do desenvolvimento de “antenas inteligentes”.

Cell-ID padrão

No caso de redes GSM, o TA com Cell-ID é padronizado como um método de segurança. Por outro lado, no caso do UMTS (*Universal Mobile Telecommunications System*), o método Cell-ID é padronizado. O SRNC determina a identificação da célula que provê a cobertura de acesso ao UE. Dependendo do status operacional do UE, operações adicionais são necessárias:

- A informação da célula ativa (cell-ID) não é conhecida no momento em que a requisição do LCS (*LoCation Services*) é recebida no SRNC. O UE deve receber um sinal de *paging* para localizar sua Cell-ID atual. Uma outra alternativa seria o Cell-ID poder ser determinado através da última conexão ativa do UE. Esta determinação é acompanhada pelo *time-of-day* da última conexão com a célula.
- A informação da célula ativa (cell-ID) é conhecida. Se o UE não está em processo de *handover* então o cell-ID é determinado. De outra forma, o SRNC combina a informação sobre todas as células associadas com o UE. Em *handover*, o UE pode ter

várias derivações de sinais conectados a diferentes cell-ID. Um cell-ID é determinado pelo SRNC baseado na área de cobertura de cada célula.

Após este processo de identificação, o cell-ID encontrado é mapeado em coordenadas geográficas ou então em uma SAI (*Service Area Identifier*), que pode incluir uma ou mais células. Este mapeamento pode ser feito tanto em uma SRNC, como em um sistema de gerenciamento de rede ou por cooperação de vários elementos de acesso a rede.

Cell-ID STK (sem padronização)

O STK, ou SIM ToolKit, é um toolkit que permite a interface com cartões SIM de usuários de dispositivos móveis. A solução de Cell-ID STK usa um script STK armazenado no SIM card da MS (*Mobile Station*) para extrair a informação de posição dos resultados NMR (*Network Measurement Results*). A célula de origem é usada para determinar o Cell-ID. A informação de posição, podendo ser o Cell-ID ou o LAC, é enviada para o gerenciador de localização através do SMS-C (*Short Message Service Center*).

O gerenciador de localização traduz o Cell-ID ou o LAC para um sistema de coordenadas geográficas (WGS 84 ou Lambert) através de um banco de dados. Além disso, o gerenciador de localização é a interface para as aplicações e serviços providos pela operadora ou um terceiro. O banco de dados do gerenciador de localização contém o Cell-ID, LAC, as coordenadas geográficas e as informações dos tamanhos das células. Este banco de dados tem que ser atualizado com frequência devido às mudanças da rede. As coordenadas geográficas podem corresponder à posição da BTS, a um ponto de interesse ou ao centro da célula.

A principal desvantagem da solução Cell-ID STK é a variação da precisão de posicionamento que pode variar numa faixa entre 100 metros em micro-células até 35 quilômetros em áreas rurais. A análise dos resultados NMR (por exemplo, a atenuação do sinal) e TA podem aumentar a precisão se algoritmos apropriados forem desenvolvidos e disponibilizados no futuro.

O método de localização Cell-ID STK é relativamente fácil de se implementar em uma rede GSM já existente sem a necessidade de gastos e atualizações no sistema.

Cell-ID IN (sem padronização)

Em redes GSM, a informação de localização de um terminal móvel pode ser gerada através de uma atualização no VLR (*Visitor Location Register*). Este processo garante que

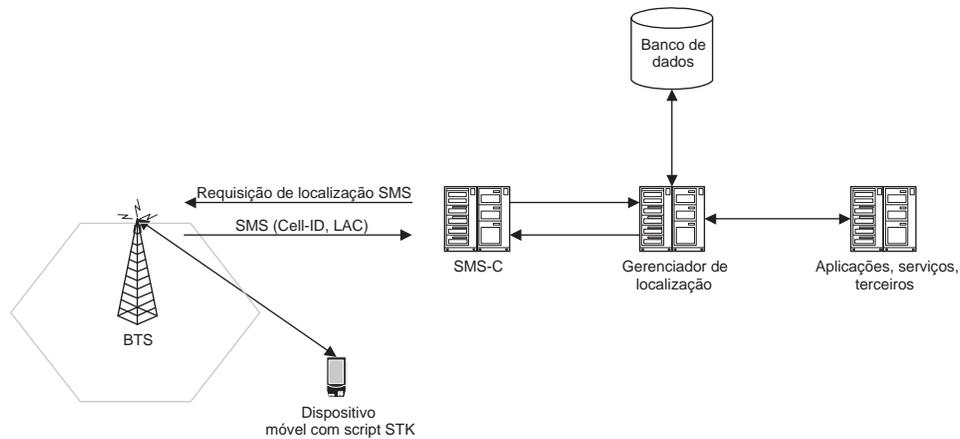


Figura 2.4: Arquitetura do sistema Cell-STK

o usuário possa ser localizado tão logo uma chamada seja feita a ele.

A informação de localização, armazenada no VLR, contém as informações de MCC, MNC, LAC e CI (*Mobile Country Code, Mobile Network Code, Location Area Code, Cell-ID*). Para se recuperar a informação de localização do VLR, o VLR pode ser acessado pelo comando MAP (*Mobile Application Part*) *Any_Time_Interrogation*, através de um serviço IN. MAP é um protocolo padrão da rede GSM. O comando MAP requisita a informação de localização do HLR (*Home Location Register*) e o HLR recupera os dados do VLR relacionado e retorna a informação de localização ao sistema IN.

A determinação de localização através do IN é uma solução que pode ser aplicada a qualquer usuário independente do aparelho celular usado ou qualquer contrato de uso do serviço. A precisão de posicionamento é limitada ao Cell-ID, satisfazendo grande parte das aplicações e serviços. Uma das grandes desvantagens é a precisão de tempo que limita esta tecnologia somente a alguns serviços.

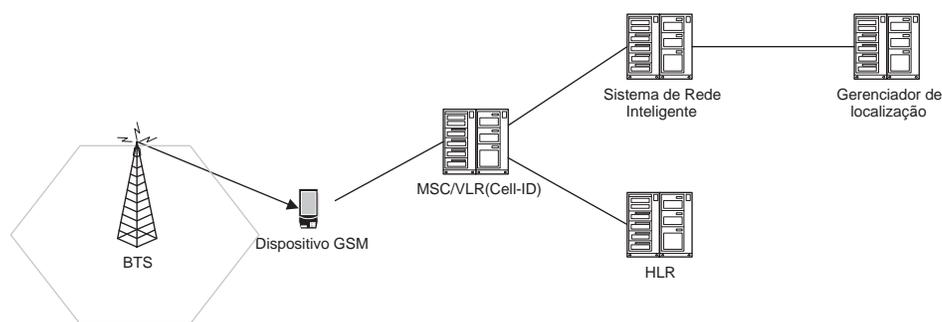


Figura 2.5: Arquitetura do sistema Cell-ID IN

Cell-ID WAP (sem padronização)

Quase todos os operadores de redes GSM na Europa oferecem acesso à tecnologia WAP. O sistema WAP é diretamente conectado ao subsistema de comutação da rede GSM e consiste principalmente de um *gateway* WAP e alguns mecanismos de autenticação e segurança incluindo um banco de dados de autenticação. O banco de dados de autenticação contém dados do usuário como MSISDN (*Mobile Station ISDN Number*), endereço IP e o *time stamp*. Adicionando-se um campo de dados para o Cell-ID, a informação de localização pode ser armazenada no banco de dados WAP. Durante uma chamada WAP o Cell-ID poderá ser recuperado do comutador da rede, repassado ao sistema WAP e finalmente armazenado no banco de dados WAP. O banco de dados WAP pode ser acessado pelo gerenciador de localização.

O acesso a informação de localização através do WAP é uma solução que pode ser desenvolvida rapidamente devido às poucas modificações que são necessárias na rede GSM. Para serviços como serviços de informação de localização onde o acesso é feito através do WAP, o método de posicionamento é ótimo. Para outros tipos de serviço onde a localização deve ser determinada fora de uma chamada WAP, o posicionamento através do WAP não é utilizado.

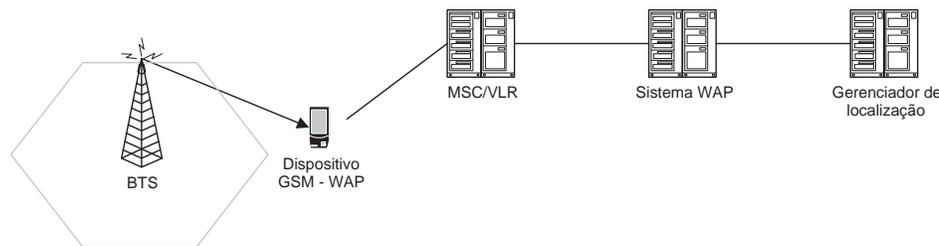


Figura 2.6: Arquitetura do sistema Cell-ID WAP

Cell-ID+TA

O método Cell-ID+TA é baseado no parâmetro GSM existente já conhecido pela BTS. O TA só é possível em redes GSM e é normalmente utilizado como um procedimento de segurança para outros métodos.

Ele utiliza os parâmetros de CGI, ou Cell-ID, e de TA para determinar a localização de dispositivos móveis. O Cell-ID identifica a célula onde o dispositivo móvel está localizado. Uma célula pode ser um setor circular ou triangular. O parâmetro TA é uma estimativa

da distância (em incrementos de 550 m) do dispositivo móvel até a estação base. A medida é baseada no atraso de acesso entre o início de um *slot* de tempo e a chegada de rajadas vindas do dispositivo. O atraso de acesso é proporcional à distância entre a BTS e o dispositivo. A precisão deste método varia de acordo com o tamanho da célula. O raio de uma célula pode variar de 100 m a 35 km.

O método de posicionamento Cell-ID+TA é um simples avanço do Cell-ID. A localização estimada é informada em termos de longitude, latitude, e uma faixa aonde o dispositivo móvel é localizado. Mesmo os outros esquemas apresentando um maior grau de precisão de posicionamento, suas principais vantagens são a velocidade de resposta rápida para se conseguir a localização (em torno de 3 segundos) e o fato de não ser necessário modificações no aparelho ou na rede.

Neste tipo de sistema, a área de cobertura da célula de uma estação base é usada como localização do usuário. A precisão da localização depende geralmente do tamanho da célula mas posições abaixo de 150 metros são possíveis em áreas urbanas com o uso de pico-células.

Atualmente está sendo investigado formas de aumentar a precisão do método Cell-ID+TA. Esta técnica avançada usa o método como base de determinação de posicionamento. Resultados de testes indicam uma granularidade entre 100 e 200 metros.

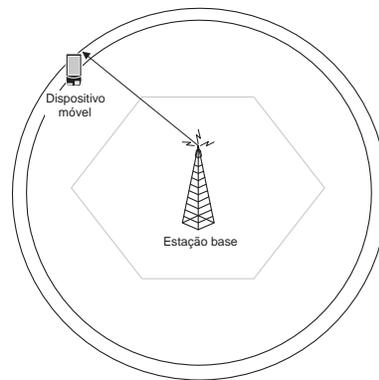


Figura 2.7: Método Cell-ID+TA

TOA

O método de posicionamento TOA é baseado na medida do tempo de chegada de um sinal de um dispositivo móvel para quatro ou mais unidades de medida. O sinal é gerado em rajadas através de uma tentativa de *handover* assíncrono feita pelo dispositivo. O método

de posicionamento TOA pode ser utilizado com todos os dispositivos móveis existentes, ou seja, nenhuma modificação é necessária nos aparelhos. As unidades de medida de localização (LMU) encontradas nas BTSs recebem as rajadas e medem o valor do tempo de chegada do enlace de subida (UL- TOA). O SMLC (*Servicing Mobile Location Centre*) calcula o TOA subtraindo os valores UL-TOA e a posição é obtida através da triangulação hiperbólica utilizando-se o PCF. Neste esquema, entretanto, podem não haver redes cobertas usadas como a unidade de medida de localização. Ao invés disso, esta funcionalidade é oferecida pela sincronização da rede celular, usando um GPS ou relógio atômico em cada BTS.

Os pré-requisitos para o cálculo de posicionamento através da utilização do método de posicionamento TOA são:

- As coordenadas geográficas das unidades de medida devem ser conhecidas;
- O *offset* de tempo entre as unidades deve ser conhecido pelo uso do tempo absoluto de GPS nas unidades de medida, ou pelo uso de unidades de medida de referência (também conhecidas como terminais de referência) para determinar o RTD.

O posicionamento através do TOA é muito caro de ser implementado desde que para todas as BTSs é necessária uma LMU, e para um servidor SMLC um PCF deve ser instalado. O método não é muito útil em áreas rurais com grandes células porque não pode se garantir que a MS esteja em contato com três BTSs. Em áreas urbanas e áreas com células de pequeno diâmetro a precisão da medida é degenerada devido aos ruídos na propagação do sinal. A precisão deste método varia de acordo com o ambiente e o número de unidades de medida de localização utilizados, e pode variar normalmente entre 50 metros (área rural) e 150 metros (área urbana densa). Uma outra desvantagem deste método é o impacto significativo na capacidade do sinal de rádio. Devido ao impacto na rede e o alto custo da implementação e manutenção, esta solução não é muito recomendada.

AOA

O AOA foi inicialmente desenvolvido para organizações governamentais e militares e pode ser operado sem modificações nos dispositivos móveis, e de início foi aplicado em sistemas celulares analógicos. Para usá-lo em sistemas digitais a brevidade dos sinais e o compartilhamento do canal o tornam um esquema complicado.

A versão mais comum desta técnica é conhecida como localização de direção de pequena abertura, a qual exige um complexo arranjo de 4 a 12 antenas em cada uma das várias

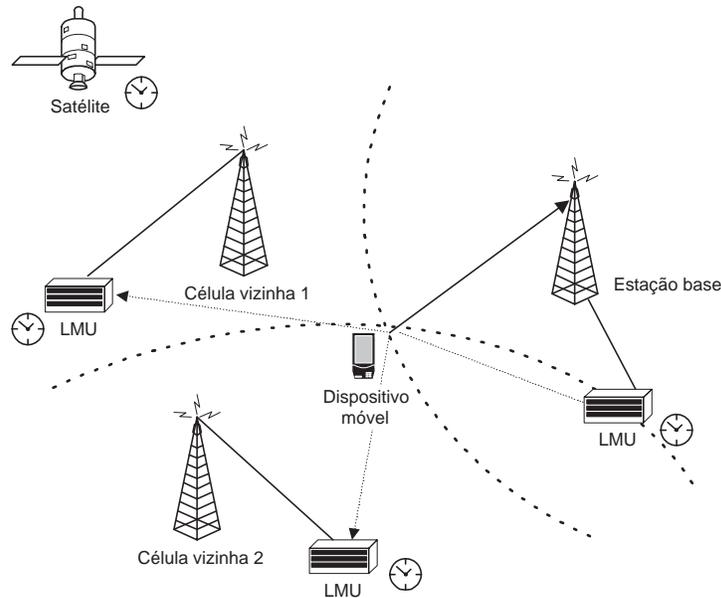


Figura 2.8: Determinação de localização pelo *Time Of Arrival*

localidades. Estes arranjos de antenas podem em princípio funcionar juntos para determinar o ângulo (relativo ao local da célula) do qual um sinal celular originou-se. Quando as várias células podem determinar seus respectivos ângulos de chegada, a localização de um telefone celular pode ser estimada do ponto de interseção das linhas projetadas de células no ângulo do qual o sinal foi originado.

Sistemas de ângulo de chegada de pequena abertura sofrem de distorções nas ondas dos sinais celulares, causadas por multicaminhamento e outros fatores ambientais. A tecnologia pode ter desempenho aceitável quando se rastreia uma transmissão contínua, como uma transmissão de voz, se tempos de integração mais longos forem usados. Como o usuário se move de célula para célula, o sistema deve seguir cada canal de voz assim que a chamada passar de canal para canal (*handoff*). Isto pode se tornar difícil se os ângulos de chegada das antenas não estiverem posicionados para interpretar a sinalização do canal de voz.

Uma desvantagem significativa dos sistemas de ângulo de chegada é o dilema logístico e estético de se colocarem arranjos de antenas nas BTS devido a problemas de regulamentação. As antenas celulares existentes não são apropriadas para o sistema de ângulo de chegada. Outra desvantagem dos sistemas de ângulo de chegada é que o erro angular do arranjo de antenas pode ser traduzido em um erro significativa na distância lateral se o telefone celular estiver longe das células.

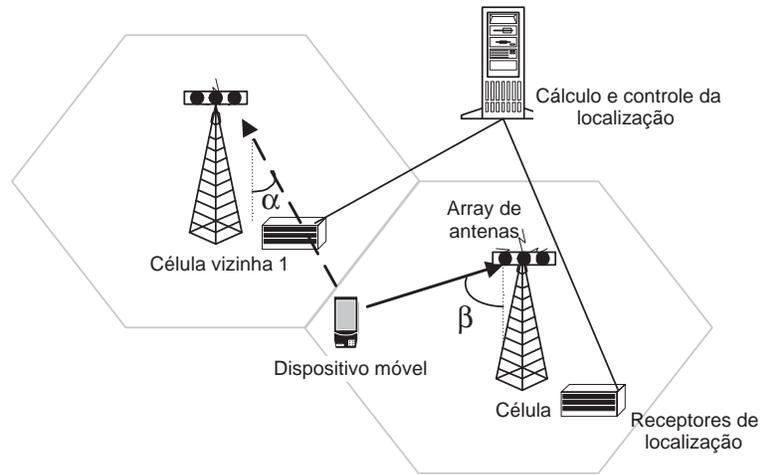


Figura 2.9: Posicionamento pelo método *Angle Of Arrival*

Atenuação do sinal

Esta técnica usa os sinais existentes para estimar a localização de telefones celulares e opera no início da atenuação do sinal a partir do momento em que o telefone se move de ou para uma estação base. A maioria das antenas de telefones celulares são omni-direcionais e a potência é dissipada rapidamente em todas as direções. Se a potência transmitida pelo telefone é conhecida, e a potência for medida em outro ponto, a distância pode ser estimada utilizando um dos vários modelos de propagação.

Entretanto, esta técnica é geralmente considerada o método menos confiável para se estimar a localização por diversas razões. Descobrir a potência transmitida acarreta uma sobrecarga significativa no sistema e é complicada de se fazer por setorização de células, inclinação de antenas ou sintonização contínua de sistemas sem fio. O sinal sofre atenuação por razões diversas além da própria distância, tais como passar por trás de paredes, folhagem, ou vidros e metais. A atenuação do sinal também sofre pelas variações sazonais devido ao clima, mudanças na folhagem e outros fatores ambientais. Circuitos de medida de potência geralmente não podem discriminar entre potências recebidas de múltiplas direções, tal como caminhos direcionados, reflexões em construções ou até mesmo em caminhões em movimento.

2.5 Estratégias de localização

A rápida evolução da comunicação celular levou ao desenvolvimento de uma nova geração de rede de comunicação móvel chamada PCS (*Personal Communications Service*). Este tipo de sistema permite que um maior número de assinantes móveis possam utilizar a rede. Outra característica importante dos sistemas PCS é que diferentes tipos de serviços podem ser oferecidos aos usuários além dos que já são usados nos sistemas celulares tradicionais.

Assim como na rede de telefonia fixa, as redes de telefonia celular necessitam de informações sobre os terminais para que haja a comunicação. Na rede fixa existe um relacionamento entre um terminal e sua localização. As chamadas para um determinado terminal são sempre roteadas para a localização associada a ele, pois não há distinção entre o terminal e sua localização.

Em redes do tipo PCS a mobilidade dos terminais é uma constante e o ponto de acesso de um terminal à rede muda a medida que ele se desloca sobre a área de cobertura. Dessa forma, o número do assinante PCS não fornece informação sobre sua localização. Conseqüentemente, torna-se necessário um mecanismo de gerenciamento de localização para a entrega de chamadas ao usuário móvel.

Um mecanismo de gerenciamento de localização é composto por três componentes: um sistema de banco de dados, que mapeia o número do assinante PCS em sua localização atual; uma operação de pesquisa para a localização do usuário móvel com base na informação armazenada no sistema de banco de dados; e uma operação de atualização que mantém o sistema de banco de dados atualizado sobre a localização do usuário móvel e faz a autenticação do usuário quando uma informação de localização está disponível.

Existem vários métodos de gerenciamento de localização que utilizam diferentes estratégias para efetuar estas operações básicas. Nas seções seguintes será feita uma breve discussão sobre alguns dos principais padrões de gerenciamento de localização.

2.5.1 IP móvel

O grupo de trabalho IP móvel da IETF (*Internet Engineering Task Force*) partiu do princípio de manter a compatibilidade com o protocolo de rede já existente no modelo de referência TCP/IP. Vários itens foram incluídos neste protocolo para dar suporte à mobilidade. A proposta era que nenhuma mudança nos *hosts* deveria ser necessária, sistemas que não possuíssem extensões de mobilidade deveriam ser tolerados (interoperabilidade com a

infra-estrutura existente) e o suporte a mobilidade deveria ser completamente transparente para os protocolos das camadas de transporte e aplicação.

O IP móvel implementa um esquema onde a mobilidade é gerenciada com a ajuda de duas estações de gerenciamento, uma na rede de origem e a outra na rede estrangeira, chamadas de agente de Origem (*Home Agent*) e agente estrangeiro (*Foreign Agent*). O *host* móvel, após registrar-se na rede estrangeira, informa ao agente de origem sua nova localização. O agente de origem encapsula todos os pacotes recebidos endereçados ao *host* móvel e os encaminha ao agente estrangeiro, que se encarrega de entregar os pacotes ao *host*.

2.5.2 IS-41 e GSM MAP

As redes que seguem os padrões IS-41 (*Interim Standard 41*) [17] e GSM MAP [38] são tipicamente redes que adotam uma arquitetura celular. As redes de comunicação celular possuem a característica de reuso utilizando um número relativamente pequeno de canais sem fio e possibilitando o suporte a um grande número de assinantes PCS. Nesse esquema, a área de cobertura da rede é dividida em células e um número de canais sem fio é atribuído a cada célula. O reuso de canais é a possibilidade de uso dos mesmos canais de uma célula em outra célula desde que esta última esteja a uma distância suficientemente grande da primeira para evitar a interferência. Para cada célula existe uma estação base, também chamada ERB (Estação Rádio Base) que tem a função de atender aos usuários do sistema através de enlaces sem fio. A área de cobertura de uma rede celular é dividida em áreas de registro, ou áreas de localização. Cada área de registro consiste em um número de células.

Os padrões IS-41 e GSM MAP são baseados em uma hierarquia de dados de dois níveis. O HLR é uma base de dados que contém registradas informações dos usuários incluindo a informação de localização e geralmente existe um HLR para cada rede PCS. Cada usuário está permanentemente associado a um HLR na rede. O VLR por exemplo, é a segunda base de dados que armazena informações (obtidas do HLR) dos terminais móveis visitando sua área. A quantidade de VLRs em uma rede pode variar. As VLRs são gerenciadas pelas CCC (Centrais de Comutação e Controle). A CCC é o elemento de uma rede que tem a função de coordenar operações de comutação e as operações de pesquisa e atualização.

Capítulo 3

Arquitetura de serviços

Neste capítulo será feita uma descrição das principais características da arquitetura de serviços proposta. Serão discutidos o modelo tradicional de arquitetura cliente-servidor utilizado e a modelagem que foi proposta. Definiu-se a arquitetura como um conjunto de módulos onde cada um dos módulos tem uma função específica como será visto a seguir.

Contudo, antes de iniciar a discussão, algumas considerações sobre o ambiente móvel serão feitas para que fique clara a compreensão sobre algumas características da arquitetura.

3.1 Considerações sobre o ambiente móvel

Como já mencionado no início do capítulo 1, o ambiente móvel apresenta várias restrições que comprometem a qualidade dos serviços oferecidos aos usuários, dentre as quais podem ser citadas as seguintes:

- taxa de bits errados
- largura de banda
- conectividade
- energia no dispositivo

Os principais problemas relacionados à computação móvel são decorrentes principalmente de questões relacionadas a mobilidade do usuário, variação e instabilidade nas condições de comunicação e na limitação de energia dos dispositivos.

Um usuário pode estar acessando o sistema de diversas localidades cobertas pela rede e a cada mudança de ponto de acesso podem haver desconexões. Os cenários mais comuns são geralmente o próprio usuário evitar intencionalmente o acesso à rede, ocorrer uma variação de ruído no canal ou então a falta de energia disponível no dispositivo. Caso haja uma desconexão o usuário fica sem acesso momentaneamente aos serviços como normalmente acontece nas redes celulares de voz atuais.

Além do problema de conectividade, o meio sem fio apresenta restrições quanto à taxa de transmissão de bits decorrentes de ruídos e atenuação nos sinais de transmissão. Esta taxa é de cinco a dez ordens de grandeza maior que a taxa de transmissão numa rede fixa, tipicamente 1 bit errado para cada 10^5 a 10^6 bits transmitidos na rede móvel comparados com 1 bit errado a cada 10^{12} a 10^{15} bits transmitidos na rede fixa. A alta taxa de erros na comunicação sem fio faz com que a eficiência do canal seja menor principalmente porque as retransmissões de pacotes serão mais frequentes. O ambiente móvel ainda apresenta problemas de largura de banda efetiva para os usuários pois, mesmo em sistemas como as redes de terceira ou quarta geração, as taxas nominais de transmissão são divididas entre os usuários de uma mesma célula.

Por último, o próprio dispositivo móvel apresenta restrições quanto à energia disponível para operações. Computadores móveis dependem de baterias para poderem funcionar. A tecnologia atual de produção de baterias não consegue acompanhar o crescimento de outros segmentos da informática além de estar muito limitada devido a fatores de escala. As baterias disponíveis atualmente são pesadas e só conseguem armazenar energia para algumas horas de uso. Muitas vezes esse problema é visto como um dos maiores obstáculos no uso de dispositivos móveis. A queda de energia em dispositivos deste tipo pode ser considerada como de comportamento exponencial no decorrer do tempo de uso dos serviços.

3.2 Descrição da arquitetura

Um dos principais desafios na disponibilização de aplicações para o ambiente móvel é a capacidade de se oferecer condições para que estas aplicações possam se adequar ao ambiente de acordo com as variações dos parâmetros citados na seção 3.1. Partindo do princípio de que os serviços a serem oferecidos devem ser adaptativos, a arquitetura deve então prover condições para essa finalidade. Dessa forma, foram investigadas algumas das possíveis soluções para este problema e observou-se que grande parte destas soluções se encaixam no modelo de arquitetura cliente-proxy-servidor [33], uma extensão do tradicional

modelo cliente-servidor. A figura 3.1 destaca os principais componentes desta arquitetura.

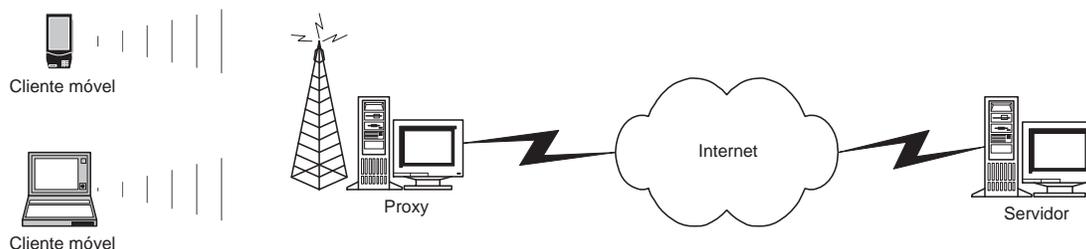


Figura 3.1: Modelo cliente-proxy-servidor

Os dispositivos móveis executam o cliente, entidade responsável pela interface e muitas vezes por uma parte da lógica da aplicação. Em uma rede de comunicação celular típica os dispositivos se comunicam com a parte fixa da rede através de uma estação base que é o ponto de acesso de todos os dispositivos que são atendidos pelas células. Para os sistemas de acesso a dados a estação base funciona como um *gateway* entre as redes móvel e fixa podendo prover funcionalidades de “conversão” entre os respectivos protocolos utilizados por estas redes. O *proxy* é o elemento central deste modelo pois ele pode ser capaz de atender às requisições do cliente e também de executar uma parte da aplicação aproveitando os recursos da rede fixa. Logicamente ele gerencia de forma transparente a complexidade do ambiente móvel permitindo que os servidores já existentes na rede fixa sejam utilizados também para o atendimento a este tipo de cliente. Geralmente o *proxy* é uma entidade que se liga diretamente às estações base.

A parte lógica dos clientes é geralmente dividida entre o dispositivo móvel e o *proxy*, ou mesmo toda atribuída a este último para que as limitações apresentadas pelos dispositivos sejam tratadas de forma mais eficiente na parte fixa da rede através de adaptações. Possivelmente o *proxy* será utilizado em máquinas robustas com maior poder de processamento, maior quantidade de memória, maior espaço em disco, além de vários outros recursos. Se características como largura de banda disponível são problemas críticos para uma rede sem fio, o *proxy* pode, por exemplo, “filtrar” ou mesmo comprimir as informações antes de enviá-las para o cliente móvel.

A figura 3.2 ilustra o ambiente no qual está sendo proposta a arquitetura. Como já foi visto no capítulo 2, o mecanismo de posicionamento de dispositivos em uma rede móvel pode se dar através do uso de subsistemas de localização da própria infra-estrutura da rede, através de GPS embutido no terminal ou mesmo da combinação dos dois anteriores. Seja qual for a tecnologia utilizada para a localização, a informação de posicionamento será

enviada para o *proxy*, que por sua vez pode enviar esta informação para um servidor de aplicações. O servidor de aplicações será responsável pela lógica da aplicação envolvendo a informação de localização e os sistemas de informação.

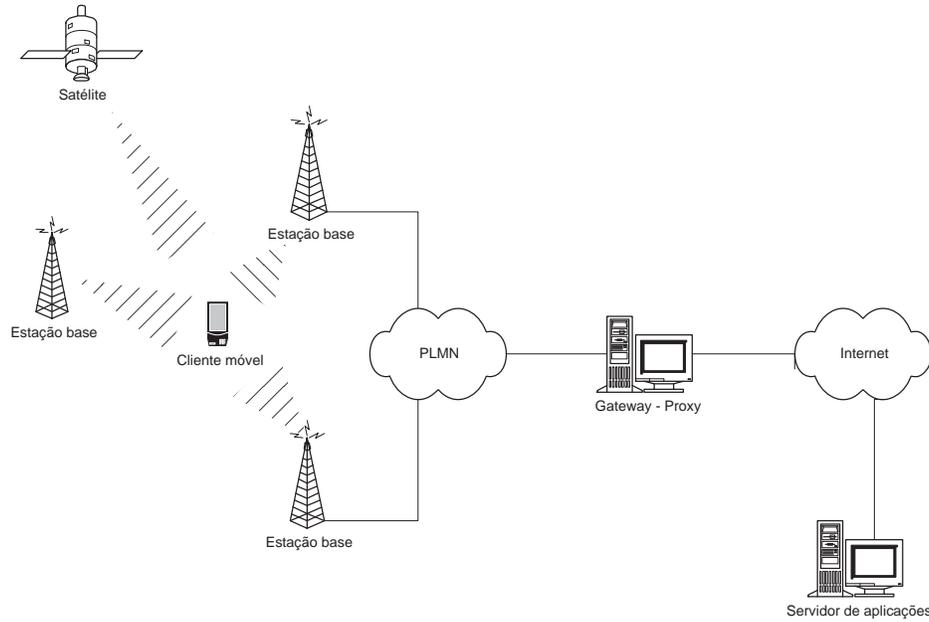


Figura 3.2: Ambiente de comunicação para LBS

Existem três grandes variáveis envolvidas no processo de utilização de um LBS. A primeira é a tecnologia de posicionamento de terminais móveis que não será tratada neste trabalho mas que será considerada como disponível. Desta forma, partindo-se do pressuposto de que a informação de localização é conhecida, o segundo ponto a ser tratado é a definição dos mecanismos de adaptação que poderão ser empregados no *proxy*. Deve-se ressaltar que a adaptação não é restrita somente ao *proxy* pois a aplicação também pode utilizar alguns mecanismos de adaptação. A terceira e última variável se resume à aplicação propriamente dita.

Para se descrever a arquitetura convencionou-se por uma divisão lógica em três módulos nas quais se encaixam os elementos citados anteriormente: localização de serviços, localização de usuários e adaptação. A figura 3.3 mostra os principais componentes da arquitetura mas antes de partir para a discussão sobre cada um dos módulos deve-se ressaltar que, como pode ser visto pela mesma, o cliente, o *proxy* e o servidor são baseados na plataforma Java. Um dos pontos importantes do trabalho é que apesar de ter sido baseada nesta plataforma, a arquitetura é aberta e pode se utilizar de outras linguagens

de programação no desenvolvimento dos serviços. Pode-se implementar um componente no *proxy* que faz este tipo de verificação.

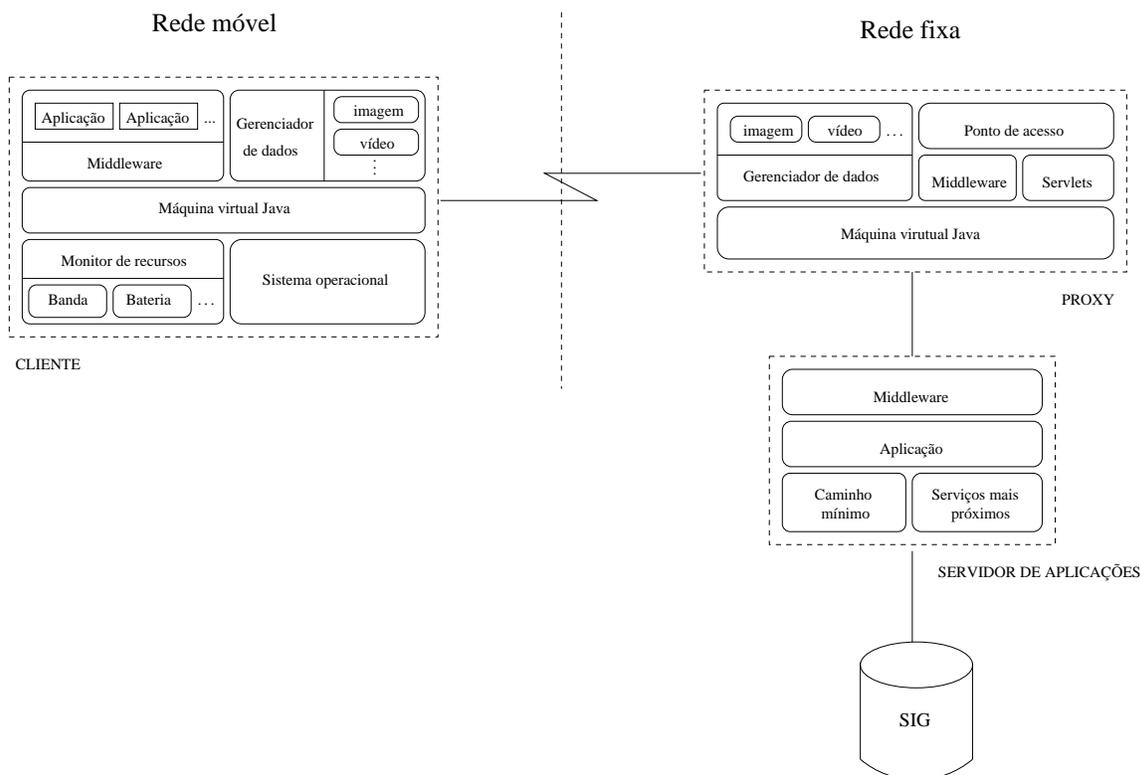


Figura 3.3: Arquitetura proposta

3.2.1 Localização de serviços

O elemento ao qual se dá a denominação de localização de serviços é formado por um servidor de aplicações, e também por uma base de dados georeferenciados que gerencia estes objetos. A seguir são descritas as propostas para estes elementos.

Servidor de aplicações LBS

O servidor de aplicações LBS gerencia as aplicações (serviços) oferecidas aos clientes. Como parte da arquitetura, é definida uma aplicação composta por dois componentes básicos de cálculo de caminho mínimo e de localização de serviços mais próximos. Qualquer eventual serviço que venha a surgir poderá utilizar estes componentes assim como outros

que venham a ser necessários. Cada aplicação implementa um *middleware* específico para que determinadas operações tenham a flexibilidade necessária aos processos de adaptação.

O *middleware* é uma camada do servidor que implementa a adaptação de recursos de processamento semanticamente inerentes às características de cada aplicação. Partindo-se do pressuposto de que os clientes na sua maioria não possuem muitos recursos de processamento, a lógica da aplicação estará em grande parte concentrada no servidor e este por sua vez terá condições de se adaptar através do *middleware*.

Por decisão de projeto decidiu-se separar a lógica da aplicação entre o cliente, o *proxy* e o servidor, diferenciando este trabalho de grande maioria das propostas encontradas na literatura. A adaptação feita pela aplicação oferece um dos mais altos níveis de adaptabilidade e também de ganho em desempenho mas, em contrapartida, torna mais difícil o processo de desenvolvimento das aplicações por exigir o conhecimento do sistema de adaptação pelo programador. Contudo, o motivo de se utilizar esta estrutura está no fato de que grande parte das aplicações será executada principalmente na rede fixa devido às restrições dos terminais sem fio. As aplicações serão capazes de executar a “adaptação de computação”.

Base de dados georeferenciada

A base de dados utilizada é composta por dados georeferenciados que serão consultados através de um gerenciador de banco de dados relacional. Essa base representa todas as informações relacionadas a endereços e a malha viária de uma cidade. Pode-se referenciar geograficamente estes endereços além de se oferecer orientação através das vias de circulação (composta por trechos de logradouros) entre quaisquer endereços.

Para realizar suas pesquisas, as aplicações devem lidar com um grande volume de dados que dizem respeito aos endereços, serviços, trechos de logradouros, objetos vetoriais, além de outros.

3.2.2 Adaptação

O módulo de adaptação foi assim chamado porque é umas das partes do sistema composta pelos principais elementos responsáveis pela adaptação. Esse módulo é composto por uma entidade responsável pela “adaptação dos dados” das aplicações utilizadas no sistema. Como será visto mais adiante, este tipo de adaptação é muito importante pois pode ser aplicada a qualquer tipo de aplicação.

Proxy

O servidor *proxy* geralmente representa o *gateway* entre a rede fixa e a rede móvel. As consultas geradas pelo cliente serão atendidas pelo *proxy* que por sua vez envia as requisições ao servidor de aplicações LBS. O resultado das requisições serão encaminhados pelo servidor ao *proxy*, que então fará a análise e aplicação da adaptação.

Como pode ser visto na figura 3.3, o *proxy* é composto por uma camada *middleware* reponsável pela negociação de fidelidade dos dados com o cliente. Além do *middleware*, existe um componente *servlet* que atende a requisições de clientes *web*. Este componente é utilizado por clientes que possivelmente não possuem capacidade de processamento necessária para negociar recursos através do *middleware* como clientes de dispositivos WAP. Para cada um destes componentes citados anteriormente existe um porto definido e que pode ser acessado pelo ponto de acesso.

Através do componente gerenciador de dados, o *middleware* é capaz de processar diferentes tipos de dados. A implementação feita está considerando atualmente dados do tipo imagem e texto mas o suporte a outros tipos de dados é possível através da adição de componentes específicos no gerenciador de dados.

A arquitetura considera processos de adaptação genéricos para a informação semântica de cada tipo de dado. Como discutido em [24], a compressão com perda é um mecanismo de adaptação eficiente que pode ser atingido através de operações específicas sobre os diferentes tipos de dados. A compressão de uma imagem, por exemplo, considera o descarte de informações de cor ou de diminuição da resolução. Já a compressão de um vídeo pode, além disso, incluir uma redução na taxa de *frames*.

A distinção entre tipos de dados também permite a priorização entre estes. Por exemplo, se uma pessoa que queira receber um artigo composto por informações no formato texto e imagem estiver sob condições "precárias" do meio, ela poderá receber este com os componentes de maior valor semântico, no caso as informações no formato texto.

Para informações do tipo imagem ou vídeo, por exemplo, os processos de compressão são mais facilmente perceptíveis devido às características inerentes aos próprios tipos de dado. Características como cor e resolução podem ser trabalhadas por diferentes técnicas de compressão. Para informações do tipo texto estes processos não são tão claros pois grande parte do *overhead* está ligado a formatação utilizada. Por exemplo, no caso de texto sem formatação os processos de adaptação podem ser aplicados de outras formas para que a quantidade de informação seja reduzida.

3.2.3 Cliente

Este é o módulo que representa o cliente de um terminal móvel. Este módulo contém a entidade cliente que é responsável pela execução das requisições ao servidor.

Cliente móvel

Como visto na seção anterior, estão sendo considerados alguns tipos de cliente devido aos diferentes recursos utilizados nos mesmos. Para cada um desses dispositivos são especificadas diferentes formas de acesso. Atualmente, os dispositivos WAP são os principais clientes existentes. Estes terminais executam aplicações através de *browsers* de acesso a *web*. Para clientes deste tipo o acesso à aplicação é possível através do componente *servlet*. A aplicação não executa nenhum processo de adaptação pois não é possível a execução de aplicações externas que possibilitem a monitoração de recursos do terminal.

Além dos atuais celulares WAP a indústria tem investido em equipamentos para as tecnologias de terceira e quarta gerações. Os equipamentos de terceira geração são caracterizados principalmente pela sua capacidade de processamento local, capacidade de armazenamento, recursos multimídia, além de outros. Considerando-se a capacidade destes equipamentos foi definido um modelo de cliente para os serviços.

A camada de *middleware* gerencia o processo de adaptação através do componente gerenciador de dados. O componente gerenciador de dados trata das peculiaridades de cada tipo de dado suportado. Ao executar uma requisição do cliente a aplicação acessa o *middleware*. O *middleware* então faz uma chamada ao gerenciador de dados para buscar o dado, mas antes que a chamada seja executada o *middleware* consulta o monitor de recursos. O monitor de recursos é o componente responsável por monitorar a disponibilidade de recursos do dispositivo através de chamadas específicas ao sistema operacional do mesmo.

3.2.4 Localização de usuários

Como já foi detalhado no capítulo 2, diversas técnicas de posicionamento podem ser utilizadas para se executar o posicionamento de um usuário em um sistema de comunicação móvel. Em nossa especificação estamos considerando que o sistema terá disponível qualquer tipo de tecnologia para a localização de usuários. Vale ressaltar que existem várias restrições quanto ao uso dessas tecnologias. Características como tempo de resposta, precisão e padrões de representação da informação de posicionamento podem ser algumas das principais limitações encontradas em sistemas desse tipo.

Capítulo 4

Aspectos de implementação

Este capítulo trata das soluções de implementação adotadas no desenvolvimento da arquitetura e do protótipo da aplicação. Será feita inicialmente uma consideração sobre a especificação e modelagem dos dados e também sobre a estrutura de dados utilizada. Logo depois serão descritas as decisões de implementação da arquitetura e das técnicas de adaptação utilizadas.

4.1 Plataforma de desenvolvimento

Este trabalho foi desenvolvido utilizando classes escritas na linguagem Java [30, 39]. O paradigma de orientação a objetos oferecido pela linguagem teve fundamental importância no desenvolvimento da arquitetura e aplicação pois códigos encapsulados em objetos oferecem uma grande flexibilidade principalmente devido ao reaproveitamento de código [4]. De acordo com a proposta deste trabalho, a modularização na forma de classes é também necessária pois os componentes dos elementos cliente, *proxy* e servidor podem sofrer eventuais modificações de acordo com a necessidade e neste caso as alterações não terão impacto no sistema como um todo. Basta para isso adicionar novas classes ou mesmo alterar algumas das classes existentes.

Como outra motivação, tem-se que os principais fabricantes de equipamentos e sistemas para dispositivos sem fio vêm anunciando o suporte de seus produtos à tecnologia Java [1, 50]. Como uma tendência atual, estes equipamentos estão inicialmente oferecendo suporte a uma categoria de classes chamada *applets*. *Applets* são pequenas aplicações que são acessadas de um servidor *web* e que podem executar algumas funções limitadas aos

recursos de uma máquina virtual.

Além disso, a linguagem Java ainda oferece classes específicas para o suporte a páginas em linguagem de marcação geradas dinamicamente. Estas classes são denominadas *servlets*. Como já mencionado no capítulo anterior, a arquitetura oferece suporte a alguns terminais com diferentes recursos. Dispositivos como celulares WAP poderão ter acesso à aplicação através de páginas geradas por *servlets* no formato da linguagem WML.

4.2 Especificação e modelagem dos dados

Em sistemas convencionais, que geralmente utilizam dados de um endereço para impressão, é comum que o endereço seja tratado como um atributo único e armazenado sem sofrer nenhum tipo de tratamento de consistência. Todos os dados do endereço, incluindo o nome e o tipo do logradouro, o número do imóvel, o bairro e complementos são considerados como um campo único em uma representação que muitas vezes pode ser armazenada de forma redundante ou até mesmo conflitante. Do ponto de vista das aplicações geográficas, o endereço deve ser tratado como uma entidade e não como um atributo. Quando se georeferencia os endereços considerando-os como entidades, todas as aplicações que se utilizam da entidade endereço passam a ser georeferenciáveis também.

De acordo com a consideração anterior, a informação que poderia estar contida em um único campo de endereço será dividida em diversos outros campos. Dependendo do tipo de aplicação que utiliza a base de dados, existem várias formas de divisão que podem ser executadas. Para a modelagem feita neste trabalho está sendo considerada a estrutura apresentada na tabela 4.1:

<i>Tipo do logradouro</i>	<i>Nome do logradouro</i>	<i>Número</i>
Rua	dos Tamoios	576

Tabela 4.1: Divisão do campo endereço

A forma normalizada é constituída pela representação do logradouro por meio de um código que será encontrado em uma tabela de apoio, onde o código será traduzido nos diversos atributos que identificam um logradouro. Os campos *Posição X* e *Posição Y* da tabela 4.2 correspondem às coordenadas geográficas dos endereços. A representação normalizada do endereço que será usada neste trabalho é mostrada na tabela 4.3.

Além dos dados de endereços e logradouros correspondentes, o sistema deve também armazenar os dados de arcos ou trechos de vias. Somente através dos dados referentes aos

<i>Código do logradouro</i>	<i>Número</i>	<i>Posição X</i>	<i>Posição Y</i>
01137	576	611827	792564

Tabela 4.2: Tabela de apoio para representação dos logradouros

<i>Código do logradouro</i>	<i>Tipo</i>	<i>Nome</i>
01137	Rua	dos Tamoios

Tabela 4.3: Forma normalizada de representação do endereço

arcos que o sistema poderá calcular e informar ao usuário sobre o caminho mínimo entre dois pontos.

Os logradouros são representados espacialmente através de um conjunto de arcos, denominados trechos, e cada trecho corresponde a uma parcela do logradouro contida entre dois cruzamentos. Os trechos possuem as mesmas orientações permitidas pela malha viária de uma cidade. Essas representações delimitam o trânsito permitido para o tráfego de veículos automotores. Na existência de circulação de sentido duplo em um logradouro, haverá um trecho para cada sentido do tráfego. As conversões permitidas entre logradouros são representadas por trechos particulares com a função de representar a conexão entre logradouros em cruzamentos. Os trechos de logradouro e de conversão são interligados através de uma rede por meio dos nós de cruzamento, de modo que cada trecho está associado sempre a dois nós, e em cada nó podem convergir diversos trechos. Podemos visualizar este cenário através da figura 4.1.

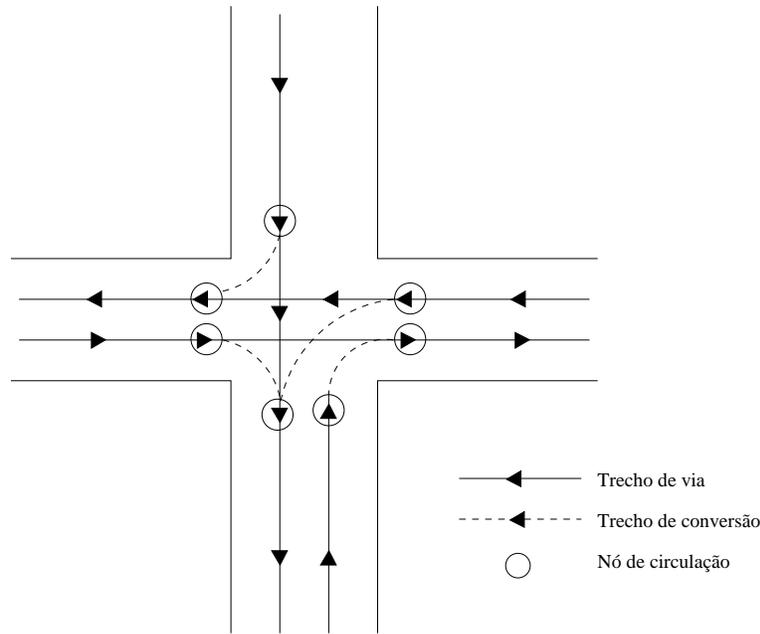


Figura 4.1: Representação de circulação por trechos e nós.

A situação acima é representada por duas entidades que representam os trechos e os nós. O relacionamento entre estas duas entidades é feito através de seus códigos.

Como pode ser visto nas tabelas 4.4 e 4.5, cada trecho terá um código de logradouro correspondente indicando a qual logradouro aquele trecho pertence. Somente no caso de trechos de conversão não haverá logradouro correspondente. O campo de comprimento informado indica a distância em **metros** do trecho. Os campos *Posição X* e *Posição Y* indicam a localização geográfica de cada nó.

<i>Código do trecho</i>	<i>Comprimento</i>	<i>Nó de origem</i>	<i>Nó destino</i>	<i>Código do logradouro</i>
68754	25	9862	10396	026718

Tabela 4.4: Tabela de trechos

<i>Código do nó</i>	<i>Posição X</i>	<i>Posição Y</i>
10871	612267	794911

Tabela 4.5: Tabelas de nós

Na figura 4.2 é apresentado o modelo de dados formado por quatro entidades com seus respectivos atributos e três relacionamentos. Essa modelagem foi baseada na especificação apresentada anteriormente nesta mesma seção.

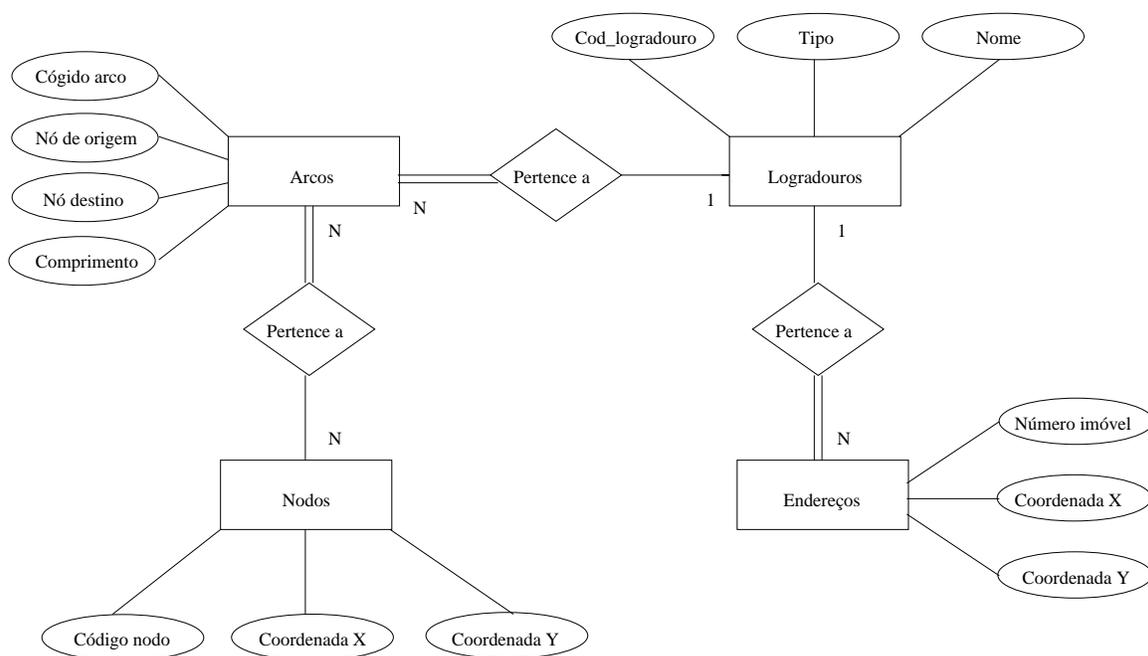


Figura 4.2: Modelo de dados usado na aplicação.

Para uma melhor compreensão do modelo será feita uma descrição das convenções utilizadas e seus respectivos significados.

Os retângulos representam as entidades e os balões ligados a esses retângulos são seus respectivos atributos. Para se expressar o relacionamento entre as entidades estão sendo utilizados losangos ligados por segmentos de retas. Entre as entidades podem haver relacionamentos que utilizam segmentos de reta simples e duplo para representar a totalidade.

As cardinalidades dos relacionamentos estão sendo representadas de acordo com a seguinte descrição apresentada na tabela 4.6:

<i>Relacionamento</i>	<i>Descrição</i>
1 ————— N	Um para muitos. A entidade que está mais a esquerda está associada a zero ou mais entidades que estão mais a direita.
N ————— N	Muitos para muitos. A entidade que está mais a esquerda está associada a zero ou mais entidades que estão mais a direita e vice-versa.
1 ————— 1	Um para um. A entidade que está mais a esquerda está associada a zero ou uma entidade que está mais a direita ou vice-versa.

Tabela 4.6: Representação das cardinalidades entre relacionamentos.

Pode-se observar que em alguns relacionamentos existe uma linha dupla que representa

a totalidade entre as associações. A totalidade indica se uma entidade do tipo de entidade relacionada participa ou não de um relacionamento. Tomemos como exemplo a entidade logradouros. A entidade logradouros, que é definida pelo seu código, pode ter zero ou mais endereços associados mas a entidade endereço tem que obrigatoriamente estar associada a algum logradouro. Podemos dizer que um logradouro pode existir independentemente se há algum endereço associado mas todo endereço deve estar relacionado a pelo menos um logradouro.

Deve-se ressaltar que cada logradouro pode conter um ou mais trechos associados mas um trecho pode estar associado a somente um logradouro. Existem casos onde um trecho pode não estar associado a nenhum logradouro, como é o caso dos trechos de conversão. A associação entre trechos e logradouros é necessária pois para os cálculos de caminho mínimo entre endereços é necessário exibir para o usuário uma identificação de fácil reconhecimento. No caso de endereços o mais comum é a divulgação de nomes de logradouros.

Para se calcular o caminho mínimo entre dois pontos é necessário que haja uma associação entre as coordenadas do usuário e também do endereço destino desejado com os nodos do grafo. Esta associação é feita através do cálculo da distância euclidiana entre os valores das coordenadas do usuário e do endereço destino e as coordenadas dos nodos. Dentre o conjunto de nodos associados aos trechos de logradouros em questão, o nodo que possuir menor distância euclidiana será o escolhido. O cálculo do caminho mínimo também pode levar em consideração restrições como tráfego, rota não disponível, etc. Neste trabalho está sendo considerado somente o atributo de distância para considerar o caminho mínimo. A figura 4.3 ilustra o processo de escolha dos nodos que serão utilizados como origem e destino no cálculo do caminho mínimo. Pode-se observar ainda que o arco em destaque define, para o exemplo, o caminho entre os dois pontos.

4.2.1 Coleta e criação da base de dados

O desenvolvimento da aplicação proposta neste trabalho teve como principal incentivo a base de informações georeferenciadas cedida pela Prodabel (Empresa de Processamento de Dados de Belo Horizonte). Sem este conjunto de informações não haveriam condições viáveis de se estabelecer de forma realística a organização viária da cidade de Belo Horizonte, que foi a cidade escolhida como exemplo.

Para se criar a base de dados foi necessário inicialmente que se definisse o SGBD a ser utilizado e optou-se pelo SGBD Oracle devido à robustez, segurança e também às ferramentas disponibilizadas pelo mesmo.

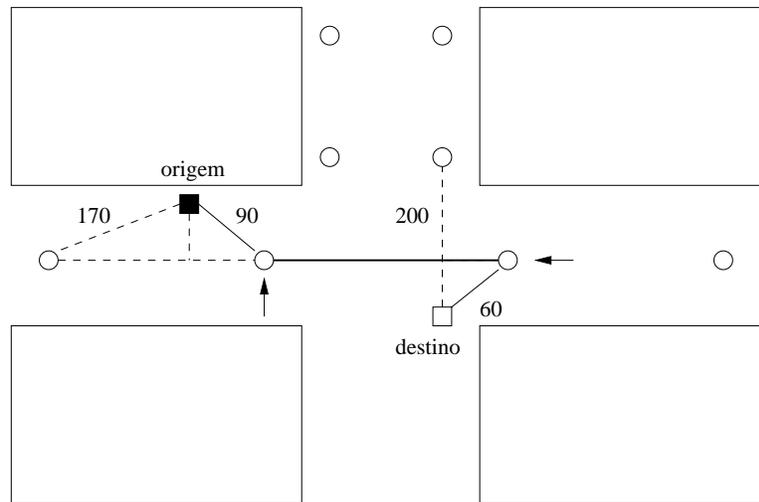


Figura 4.3: Distância euclidiana

Para o lançamento dos dados no SGBD foram desenvolvidos alguns scripts que pudessem ler a partir de arquivos no formato texto pois os dados estavam em sua grande maioria no formato texto. Estes scripts foram desenvolvidos na linguagem Perl devido ao poder e facilidade desta linguagem em relação à manipulação de dados deste tipo. Os scripts geram as entradas para criação das tabelas a serem tratadas pela ferramenta de manipulação de dados do Oracle - o Sqlplus. O Sqlplus é a ferramenta da Oracle que implementa a linguagem SQL (Structured Query Language) permitindo que o usuário possa interagir com o SGBD através de comandos específicos.

4.3 Estrutura de dados utilizada

Para representar o grafo que compõe a malha viária da cidade de Belo Horizonte propõem-se uma estrutura de dados que represente os nodos juntamente com os trechos associados a cada um desses nodos. A figura 4.4 ilustra um exemplo da estrutura de dados utilizada pelos módulos de cálculo de endereços mais próximos e de caminho mínimo entre dois pontos.

A estrutura é formada por um *hash* composto pelos nodos do grafo e cada elemento do hash é identificado pelo código do nodo, dois campos com as coordenadas geográficas, respectivamente os campos *coord_x* e *coord_y*, um campo de status para indicar a presença ou não deste nodo numa determinada rota quando se estiver usando o algoritmo de caminho mínimo (P para permanente e T para tentativo) e um quinto campo representado por um

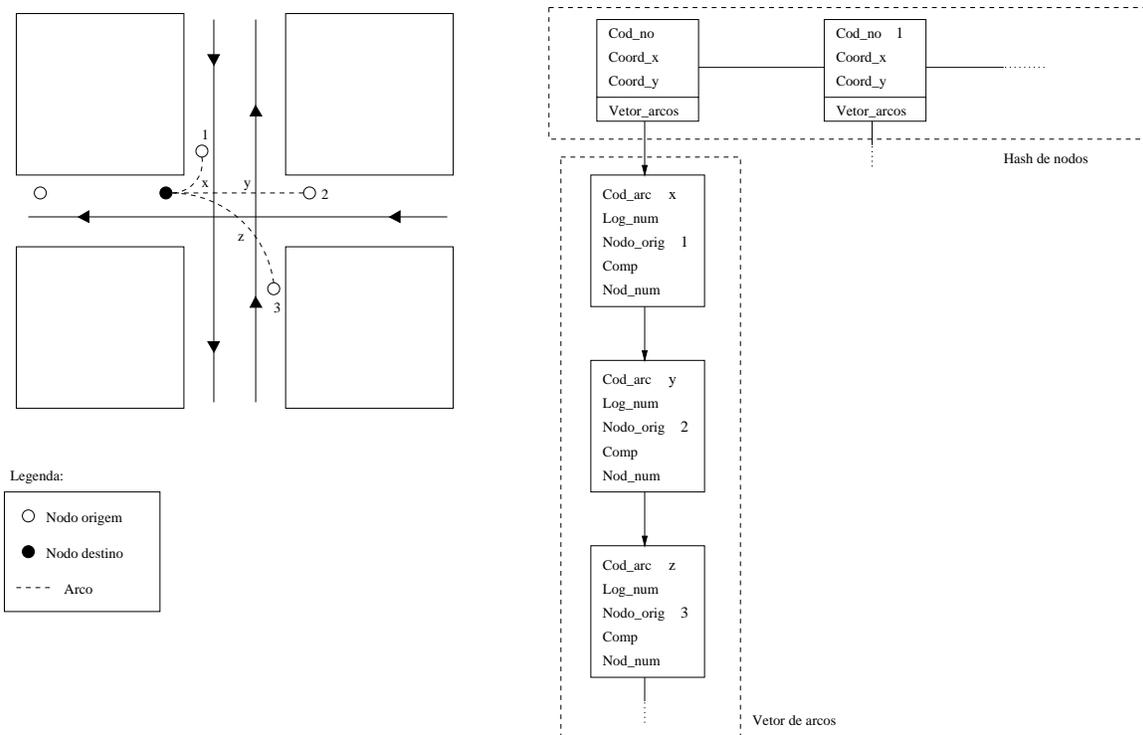


Figura 4.4: Representação da estrutura de dados

vetor de elementos. Este vetor representa os trechos que ligam estes nodos a outros nodos do grafo e cada elemento do vetor contém o identificador do arco, o identificador do nodo de origem do arco, o comprimento em metros do arco e o identificador do logradouro associado.

A estrutura de dados é criada apenas uma vez na inicialização do servidor de aplicação. Para tanto, foi implementada uma classe para realizar as consultas ao banco de dados através do JDBC (Java DataBase Connectivity). Além das consultas, a classe também se encarrega de carregar a estrutura de dados com os dados provenientes das consultas. Uma das vantagens de se realizar este processo de forma centralizada é que todas as consultas são feitas através de uma única conexão e a partir desta conexão também é carregada toda a estrutura de dados. Depois deste processo de inicialização a estrutura fica disponível em memória para as sucessivas consultas. Entretanto, apesar de ser vantajosa em algumas circunstâncias, esta estratégia impõe um custo em relação a quantidade de memória disponível no servidor. Para que ela seja eficientemente aproveitada é necessário que o servidor possua espaço suficiente em memória para armazenar e manipular a estrutura.

4.4 Aplicação

O protótipo da aplicação desenvolvido neste trabalho pode ser considerado como uma aplicação de “*Páginas amarelas*” (*Mobile Yellow Pages*). Aplicações deste tipo, apesar da simplicidade, são de grande utilidade para a população principalmente em situações cotidianas como naquelas em que a procura por provedores de serviços ou mesmo endereços se tornam cada vez mais difíceis devido ao crescimento constante das grandes cidades. O resultado da consulta forma uma lista dos endereços que sejam mais próximos a localização do usuário.

Além de indicar os endereços mais próximos, a aplicação ainda oferece a rota entre os pontos de origem e destino como um serviço de navegação para orientação. Ao escolher o endereço desejado, o sistema informa ao usuário o caminho a ser seguido entre o ponto onde ele se encontra até o endereço informado. Em sistemas mais sofisticados o usuário poderá optar por fazer o percurso a pé ou usando um automóvel, e até mesmo se deseja utilizar o percurso que possui menor número de vias de alto fluxo, etc.

Para este protótipo foram considerados dois formatos de visualização da rota: texto e figura. A descrição da rota no formato texto é formada por uma lista de logradouros obtida principalmente através de duas rotinas, como descrito abaixo. As figuras utilizadas são estáticas e foram utilizadas neste trabalho para a validação das estratégias de adaptação. Como será visto na seção 4.6, as estratégias propostas consideram as figuras como parâmetro para a aplicação da adaptação pois elas consomem a grande maioria dos recursos envolvidos.

Existem duas classes implementadas que executam as principais operações da aplicação: serviços mais próximos do usuário e cálculo da rota de caminho mínimo entre um endereço de origem e um endereço de destino. O algoritmo de localização de endereços mais próximos executa os cálculos considerando a distância euclidiana entre dois pontos. Para se calcular a rota de caminho mínimo utilizou-se o algoritmo de Dijkstra [51].

Como será visto mais adiante, foi desenvolvido um simulador do módulo cliente que considera a localização do usuário através da descrição textual do endereço de origem para a execução das consultas. Apesar do sistema considerar diretamente as coordenadas geográficas dos usuários obtidas através dos métodos já vistos anteriormente, o simulador utiliza o endereço como campo de entrada para informação de localização pois este é uma das formas mais comuns da população se referenciar a ocorrências de localização espacial.

Como descrito na seção 4.2, para que as ocorrências mais próximas do usuário possam ser calculadas é feita uma associação entre as coordenadas do usuário com os nodos do

grafo. Como todo endereço da base também é associado a coordenadas, pode-se mapear facilmente um endereço aos nodos do grafo.

4.4.1 O algoritmo de Dijkstra

O algoritmo de Dijkstra é um dos algoritmos que calcula o caminho de custo mínimo entre vértices de um grafo. Escolhido um vértice como raiz da busca, este algoritmo calcula o custo mínimo deste vértice para todos os demais vértices do grafo. Ele é bastante simples e com um bom nível de desempenho. Ele não garante, contudo, a exatidão da solução caso haja a presença de arcos com valores negativos, o que não acontece neste caso pois todos os arcos que ligam os nodos do grafo possuem pesos positivos.

A estrutura de dados descrita anteriormente foi planejada levando-se em conta principalmente a utilização do algoritmo de Dijkstra e para se compreender o seu funcionamento pode-se observar inicialmente a figura 4.5. Supondo que se queira traçar o caminho mínimo entre um nodo A e um nodo X por exemplo, tem-se que o nodo destino é o nodo X e então a busca começa por este nodo.

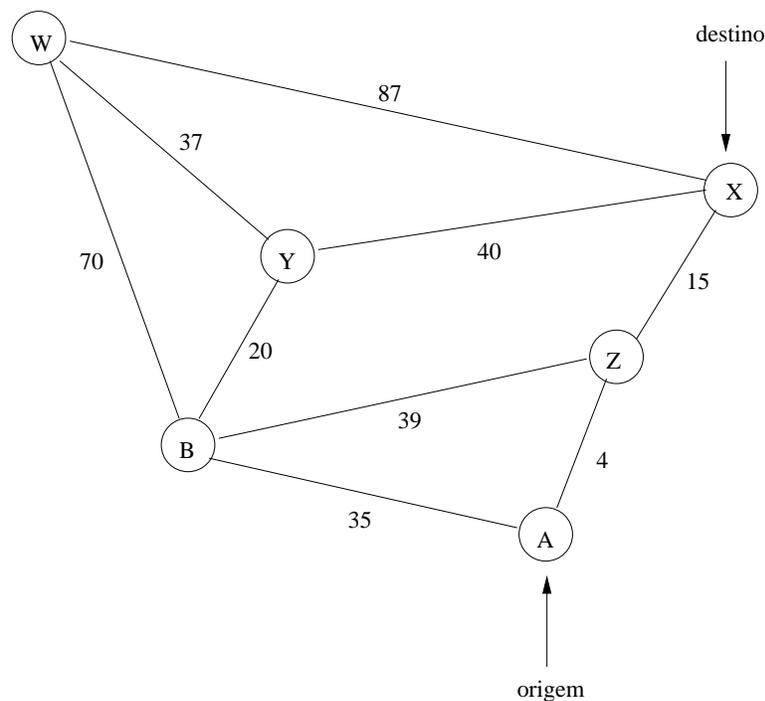


Figura 4.5: Exemplo de um grafo a ser percorrido

O primeiro passo é alterar o status do nodo corrente X de T (tentativo) para P (per-

manente) para incluir este nodo na rota final. Depois disso passa-se a verificar qual dos elementos do vetor de arcos apresenta o menor comprimento. Este passo procura garantir o arco de menor caminho entre o próximo nodo. Como pode ser verificado pela figura 4.6 o arco de menor caminho é o que liga o nodo X ao nodo Z. A partir deste ponto o nodo corrente passa a ser o nodo Z.

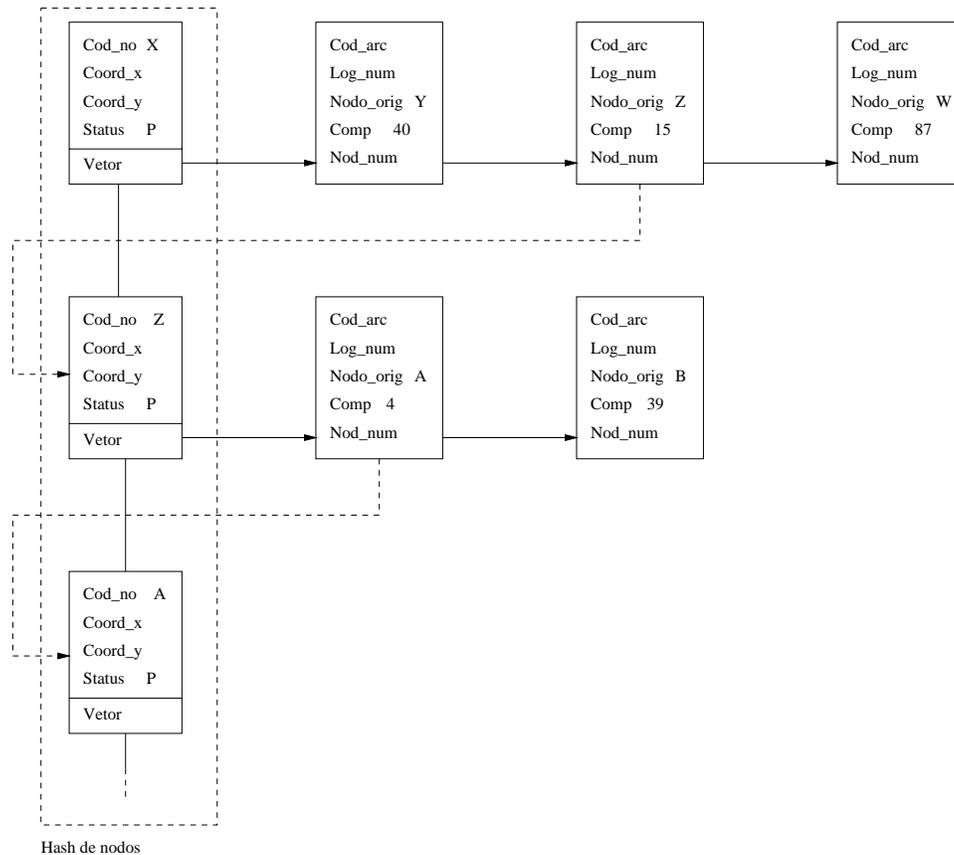


Figura 4.6: Rota completa entre origem e destino

Iniciando a pesquisa agora a partir do nodo Z verifica-se novamente qual elemento do vetor de arcos apresenta o menor comprimento. Neste caso o arco escolhido é o arco que liga o nodo Z ao nodo A. Como o nodo A é o nodo de origem então o algoritmo verifica esta condição e, antes de parar a sua execução, é modificado o status do nodo origem para garantir que este último também seja adicionado a rota.

Com os nodos inclusos na rota, agora basta percorrer a lista formada pelos elementos com o campo status = P e verificar quais os logradouros que correspondem aos arcos seleccionados. Para o usuário o conjunto dos logradouros é a informação desejada.

4.5 Descrição dos módulos da arquitetura

Cada componente das entidades cliente, *proxy* e servidor foi implementado como uma classe na linguagem Java. Uma das principais restrições existentes é que os objetos instanciados só podem acessar funções permitidas pela máquina virtual Java. Desta forma as aplicações ficam limitadas a certos recursos do sistema operacional. Uma das vantagens desta abordagem é que os aspectos relacionados com a mobilidade, por exemplo, são gerenciados por um *middleware* implementado em classes que confinam o impacto destes aspectos em códigos específicos. Isto permite que a complexidade relacionada aos detalhes de implementação sejam abstraídos dos outros módulos da arquitetura.

As entidades *proxy* e servidor são classes *daemon* que ao serem instanciadas passam a executar um *listen* em portos distintos para aceitarem conexões dos clientes. O cliente é uma classe *applet* que ao ser executada pela JVM do dispositivo automaticamente inicia o *deck* com a aplicação. Para requisitar as informações a aplicação se conecta através de um *socket* com o *proxy*. O *proxy* por sua vez também se comunica através de um *socket* com o servidor. Para o protótipo desenvolvido neste trabalho estão sendo utilizadas conexões TCP entre as entidades mas para uma simulação mais realística o uso de datagramas seria mais viável. O uso de protocolos de transmissão por datagramas é o cenário padrão nas redes de comutação por pacotes.

A seguir será feita uma descrição passo-a-passo do funcionamento da arquitetura. Esta descrição é feita a partir de uma requisição do cliente.

- Para acessar a aplicação o usuário faz uma requisição ao documento *web* que contém o *applet* da aplicação. O servidor *proxy* se encarrega de processar a requisição e devolver o documento contendo a classe *applet* da aplicação. Está sendo considerado que o servidor *proxy* implementa um servidor *web* para atender a requisições de *applets* e também de páginas geradas dinamicamente pelo componente *servlet*.
- Ao ser carregado, o *applet* instancia o componente *middleware* cliente. Neste momento o *middleware* cliente instancia um objeto do tipo monitor de recursos que se encarrega de monitorar os recursos disponíveis e passar esta informação de volta para o *middleware* cliente. O *middleware* cliente faz então a transmissão dos dados dos recursos disponíveis ao *middleware* proxy. Esta transmissão é crucial pois a adaptação depende destes dados.
- Após a notificação, o *middleware* cliente faz requisições ao *middleware* proxy através do componente gerenciador de dados. O objeto gerenciador de dados oferece métodos

para requisição por tipo de dados e estes métodos abrem conexões com o *middleware* proxy através de interface *socket*. A conexão por tipo de dados é feita pois cada um destes tipos terá um processo de adaptação específico.

- O servidor proxy faz a requisição das informações para o servidor de aplicações. O *middleware* servidor se encarrega de tratar estas requisições geradas pelo *proxy* e define para a aplicação os dados a serem devolvidos. O *middleware* servidor também se encarrega da adaptação de computação das aplicações. Cada aplicação implementa o seu *middleware* de acordo com a sua lógica de funcionamento.
- Ao receber os dados, o *middleware* proxy instancia o gerenciador de dados. Os métodos do objeto gerenciador de dados instanciam classes específicas para o tratamento dos dados e estes componentes tratam de fazer a adaptação de acordo com a informação de recursos enviado anteriormente pelo cliente.

4.6 Adaptação

Em Fox et al.[24] são identificados três princípios de projeto para tratar da variação do cliente de forma mais eficiente:

- Mecanismos de compressão com perda (*lossy compression*) específicos por tipo de dado que podem atingir melhores níveis de compressão do que mecanismos genéricos de compressão. O mecanismo de compressão com perda reduz um arquivo eliminando permanentemente algumas informações, especialmente informações que sejam redundantes. Quando o arquivo é descomprimido, somente uma parte da informação original é mantida. A utilização deste tipo de compressão permite que se defina o quanto se quer perder de informação, criando-se assim um compromisso entre o tamanho do arquivo e a qualidade da imagem.
- Adaptação “*on the fly*”¹ processando uma dada representação de um objeto de um tipo qualquer ao invés de se basear num conjunto de representações pré-computadas. Embora haja uma latência maior associada com a geração de representações do dado, a latência de comunicação entre cliente e servidor pode ser reduzida devido ao menor tamanho da representação processada.

¹Expressão utilizada para representar uma atividade realizada em tempo de execução.

- Diminuir a complexidade nos clientes e servidores pois existem várias razões técnicas e econômicas que não permitem que todos os clientes e servidores tenham muito poder de processamento, capacidade de memória, etc. A adaptação *on the fly* deve ser feita num proxy intermediário que tem acesso a recursos mais robustos e conexões mais confiáveis a Internet.

Como parte do subsistema de adaptação foi modelada uma estrutura composta por componentes que processam diferentes tipos de dados. Cada um destes componentes é responsável pela aplicação de técnicas de adaptação específicas para diferentes tipos de dados. Para o protótipo desenvolvido neste trabalho implementou-se uma classe componente para tratamento de dados do tipo imagem que utiliza uma biblioteca de classes para compressão de imagens com perda, como descrito na seção 4.6.1.

4.6.1 Níveis de adaptação

A partir da especificação da aplicação foi feita uma análise dos níveis de adaptação que poderiam ser aplicados aos diferentes tipos de dados e o que poderia ser feito pela aplicação no tratamento da adaptação. Nesta situação é necessário que se revejam os princípios e conceitos dos algoritmos tradicionais pois as aplicações desenvolvidas considerando este paradigma dificilmente seriam suportadas no ambiente móvel devido ao alto grau de restrições apresentadas tanto pelo hardware do dispositivo quanto pelo meio de comunicação sem fio. Como consequência desta necessidade introduziu-se o conceito de algoritmos multi-fidelidade [41, 46]. Algoritmos multi-fidelidade são uma sequência de passos computacionais que ao final de sua execução definem como resultado um valor que esteja dentro de uma faixa de aceitação composta por valores chamados de fidelidades.

Diferentemente da proposta inicial de se ter uma especificação única para a geração de uma saída a partir de uma sequência de passos, os algoritmos multi-fidelidade tratam os problemas de uma maneira mais flexível. Tem-se como resultado uma saída que se adequa a uma faixa de valores aceitáveis, os quais são chamados de fidelidades. Dependendo do estado de parâmetros externos ao ambiente de execução, os algoritmos multi-fidelidade podem gerar qualquer resultado que esteja dentro do nível de aceitação. Pode-se pensar que as fidelidades são níveis de degradação aplicados a uma mesma informação, ou seja, para um mesmo resultado tem-se diferentes níveis de qualidade de representação do mesmo.

Em contrapartida à fidelidade de dados, como poderia ser chamado o processo anterior, a fidelidade computacional é o processo de adaptação que se apóia na alteração de um

processo de execução, baseado em parâmetros externos, para geração de uma saída. Neste caso, têm-se como saída dados também alterados que podem representar a informação original de forma resumida.

Padrão JPEG2000

Para o tratamento de imagens está sendo utilizado um conjunto de classes do pacote JJ2000 que implementa o padrão JPEG2000 parte 1 [29]. Este pacote é resultado do projeto de implementação do padrão na linguagem Java por um consórcio formado pelo CRF (*Canon Research Center France*), pela Ericsson e pelo EPFL (*École Polytechnique Fédérale de Lausanne*) [31]. A integração deste pacote à arquitetura foi feita através da modificação das classes de codificação e decodificação de *bitstream* e da classe de entrada/saída. A partir destas modificações a classe de codificação pode alterar os atributos da imagem formada de acordo com os dados de recursos disponíveis. Vários atributos podem ser modificados de acordo com os recursos do algoritmo de codificação mas neste trabalho está sendo considerado o parâmetro de resolução em **bits por pixel**. Alterando esta taxa pode-se ter uma representação com menor ou igual resolução do que a imagem original.

O projeto do padrão JPEG2000 teve como principal motivação a possibilidade de aumentar os recursos e o desempenho na compressão de imagens principalmente para aplicações *web* e aplicações para o ambiente móvel.

Vários recursos de tratamento de imagens estão disponíveis e grande parte deles pode contribuir para oferecer maior flexibilidade ao processo de adaptação. Um estudo comparativo entre o padrão JPEG2000 e o padrão JPEG (*Joint Photographic Experts Group*) [11] mostra que para aplicações que utilizam imagens de alta qualidade (0.5-1.0 bpp) o JPEG2000 apresenta taxas de compressão 10% a 20% melhores do que o JPEG.

Fidelidade

Como já citado no capítulo 1, fidelidade é o grau de compromisso entre a cópia de um dado apresentada ao cliente em relação ao dado original que se encontra em um servidor. A classe *Encoder* do pacote JJ2000 foi modificada para receber os parâmetros de codificação da imagem para executar a compressão. Dentre estes, o parâmetro *rate* está sendo utilizado para alteração da taxa de bits por pixel (bpp) na aplicação da adaptação.

Para se alterar a fidelidade de uma imagem foi implementada uma classe que trata do compromisso entre a informação do meio de comunicação e a taxa de compressão a ser utilizada. Como já foi dito anteriormente, o protótipo desenvolvido lida com imagens

estáticas e para tanto utilizou-se uma imagem no formato PPM (*Portable Pix Map format*), um dos formatos que podem ser interpretados pela biblioteca JJ2000.

A classe citada anteriormente faz predições a respeito do tamanho da imagem que pode ser obtida com a compressão a uma determinada taxa. A predição é feita através da análise dos tamanhos obtidos em outras execuções do algoritmo de compressão utilizando uma certa taxa aplicada à imagem e aplicando-se esta mesma taxa sobre a imagem original para produzir o tamanho desejado.

Como o protótipo é baseado na mesma imagem, a variação da relação entre a taxa de compressão aplicada e o tamanho da imagem é bastante homogênea. Em média, a porcentagem de redução da taxa de compressão correspondeu diretamente proporcional à porcentagem de redução do tamanho da imagem, ou seja, reduzindo-se a taxa de compressão em 30 valor inicial. A partir da observação feita definiram-se também alguns valores que correspondem aos diferentes níveis de qualidade da informação. Como será visto mais adiante, o sistema considera alguns parâmetros de qualidade de serviço, e o parâmetro de fidelidade pode ser definido entre os valores normal e baixa, que correspondem respectivamente aos valores de taxa de compressão mínimos de 0,7 bpp e 0,2 bpp.

Adaptação do processamento

Além da fidelidade do dado, a arquitetura contempla a capacidade das aplicações também poderem adaptar seu processamento. Como será visto na próxima seção, o protótipo permite que o usuário possa definir alguns parâmetros de qualidade de serviço. Esses parâmetros, juntamente com as informações do meio e do dispositivo, definem o cenário que impõe as restrições para a aplicação da adaptação. Dentro deste contexto, a aplicação foi modelada para que pudesse se adaptar às condições citadas através da redução de passos do processamento para a obtenção do resultado.

O processo de adaptação de dados foi especificado para ser executado no *proxy* pois desta forma qualquer aplicação pode delegar a esta entidade a função de adaptação que seja adequada ao tipo da informação. Para o projeto específico do protótipo proposto no trabalho está sendo considerado um cenário de cooperação entre a aplicação e o sistema baseado no *proxy*. Os diferentes tipos de informação são processados para cada requisição do usuário e os dados são tratados de forma “atômica”, ou seja, sob os aspectos da adaptação, a aplicação trata todos os dados de forma única.

Como consequência a aplicação também passa a ter acesso aos dados do meio e do dispositivo e a partir destas informações ela pode definir a taxa de compressão que deverá

ser aplicada sobre as imagens geradas utilizando o método de predição. Depois de definida a taxa mais adequada, a aplicação envia esta informação para o *proxy* para que este possa executar a compressão sobre as imagens. O mesmo processo poderia ser aplicado a outros tipos de dado.

Como descrito na seção 4.7, a partir de uma requisição a aplicação calcula os endereços mais próximos da localização do usuário e cria uma lista. Para cada ocorrência desta lista são criadas listas de logradouros a serem percorridos e também é associada a imagem representando a rota. No caso específico deste protótipo está sendo considerado somente o *overhead* gerado pela transmissão das imagens associadas às ocorrências da lista de endereços. Desta forma a aplicação está considerando o dado do tipo imagem para aplicar as técnicas de adaptação necessárias.

4.7 Simulador

Para possibilitar a visualização dos resultados foi desenvolvido um simulador do módulo cliente. A interface deste simulador é composta por uma janela principal com o gerenciador de recursos do ambiente. Como mostra a figura 4.7, o gerenciador de recursos é composto por três conjuntos de campos de informação e por três botões.

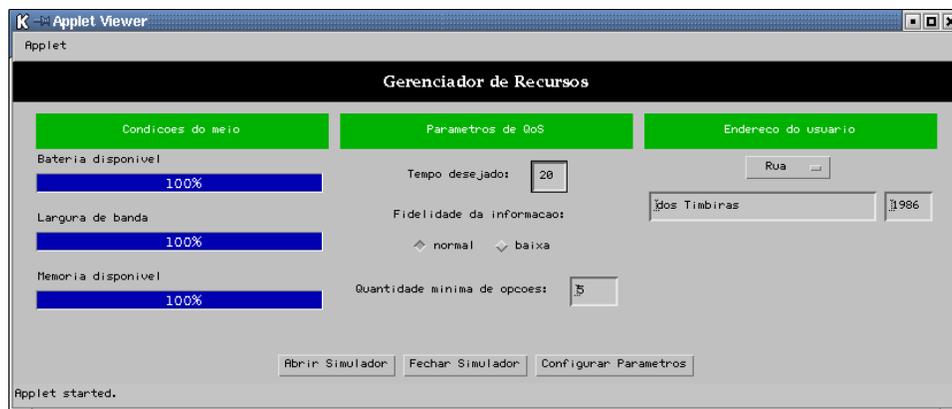


Figura 4.7: Gerenciador de recursos

O primeiro conjunto corresponde às informações de recursos do dispositivo e do meio de comunicação: quantidade de memória e bateria disponíveis e largura de banda. O segundo conjunto possui campos de definição dos parâmetros de QoS (Qualidade de Serviço). Foram definidos três parâmetros de qualidade de serviço: tempo em segundos desejado

para a execução da consulta, nível de fidelidade da informação e a quantidade mínima de ocorrências desejada como resposta. O terceiro e último conjunto de informação é formado pelos campos de definição do endereço correspondente à localização do usuário. Como não é o objetivo deste trabalho, não foi especificado e implementado nenhum tratamento em relação a consistência do endereço digitado pelo usuário do simulador.

O botão “Abrir simulador” instancia uma nova janela com a interface da aplicação no dispositivo móvel, como mostra a figura 4.8. A partir desta janela o usuário do simulador poderá iniciar as requisições de serviços ou endereços.

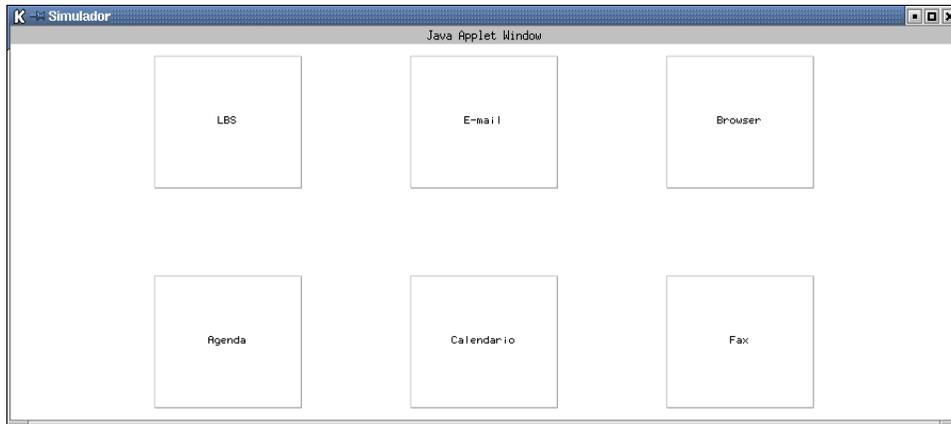


Figura 4.8: Simulador da aplicação

Ao acessar a opção de serviços **LBS** o simulador oferece a lista das ocorrências (endereços) mais próximas da localização do usuário, como visto na figura 4.9.

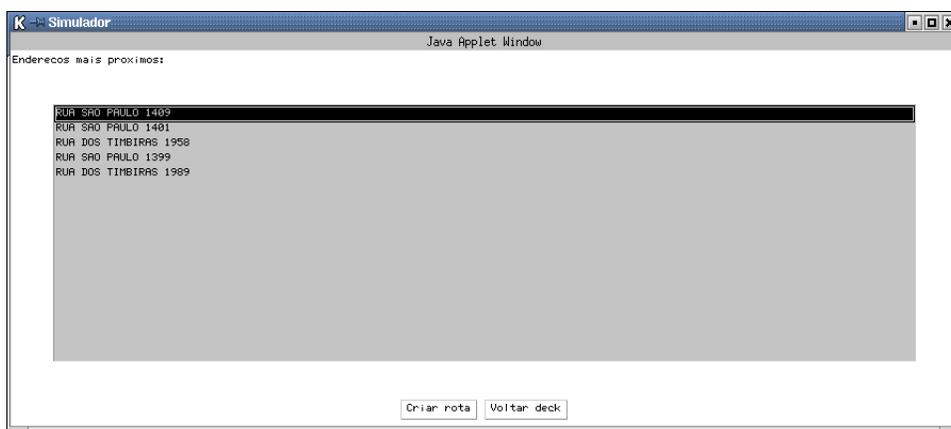


Figura 4.9: Lista dos endereços mais próximos

A partir da lista gerada o usuário poderá escolher entre qualquer um dos endereços para

obter a informação da rota a ser seguida. A figura 4.10 mostra a janela com as informações do endereço selecionado. A aplicação gera uma descrição textual e uma figura com a rota entre o endereço de origem e o endereço de destino.

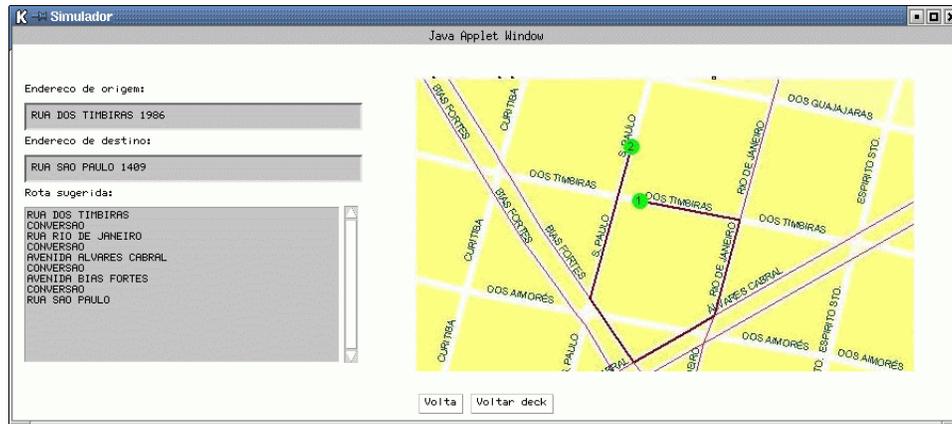


Figura 4.10: Informação de rota a ser seguida nos formatos texto e figura

A utilização do simulador é bastante simples. A partir de um estado inicial das condições do meio e dos parâmetros de qualidade de serviço o usuário poderá fazer as requisições utilizando o botão “Abrir simulador”. Para que as requisições sejam processadas de forma adequada é necessário que o cliente informe à aplicação sobre as condições do meio e de recursos. Estes dados podem ser configurados a partir da janela de configuração de parâmetros que pode ser acessada a partir do botão “Configurar parâmetros”.

Como pode ser visto na figura 4.11 existem três campos de definição de valores que foram propostos como parâmetros. Para cada parâmetro podem ser definidos o total de recursos disponível e o valor inicial que será considerado nas requisições. Cada um dos parâmetros é quantificado em unidades de medida padrão e os valores iniciais definidos foram de 500 Kb (*Kilobytes*) de memória do dispositivo, 100% de carga de bateria e uma largura de banda disponível de 14.400 Kbps (*Kilobits* por segundo).

A atualização dos dados de recursos no simulador é feita através do monitoramento dos parâmetros a cada requisição solicitada. A cada conexão o cliente acumula a quantidade de dados transmitidos em *bytes* e decrementa este valor do valor de memória disponível. O decremento de carga de bateria disponível segue um comportamento linear decrescente de acordo com o tempo e a largura de banda pode ser definida a cada requisição.

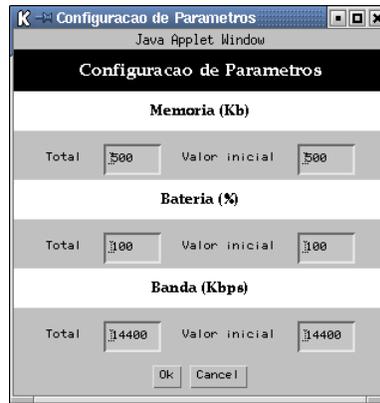


Figura 4.11: Tela para definição dos parâmetros das condições do meio e do dispositivo

4.8 Parâmetros de qualidade de serviço

Um conceito muito importante que pode influenciar bastante no processo de adaptação no ambiente móvel é o conceito de QoS (Quality of Service). Não obstante seja a importância das condições do meio, os parâmetros de qualidade de serviço definem a forma como será feita a adaptação, seja ela nos dados ou na própria computação. A partir do momento em que são definidos estes parâmetros a aplicação deverá oferecer o serviço de acordo com as preferências do usuário.

De acordo com os valores de recursos do ambiente definidos, a aplicação deverá processar as informações desde que não ultrapasse os limites mínimos impostos pelos parâmetros de QoS. Seguindo estes princípios, foram especificados três parâmetros de qualidade de serviço para o protótipo da aplicação:

- tempo mínimo de resposta
- nível de fidelidade da informação
- quantidade mínima de opções de endereços

O tempo definido pelo usuário será considerado juntamente com a largura de banda disponível para que a aplicação produza uma quantidade de informação inferior ou igual ao cálculo de multiplicação entre o tempo em segundos pela largura de banda em Kilobits por segundo.

O uso de diferentes técnicas de compressão são estudadas para a identificação de cenários que ofereçam uma melhor eficiência do canal de comunicação em um ambiente sem fio

[44]. Partindo-se do pressuposto de que a compressão aumenta a eficiência da transmissão no meio sem fio, definiu-se como um dos parâmetros de QoS o nível de fidelidade da informação, pois a fidelidade define diferentes patamares de qualidade da informação que são proporcionais à taxa de compressão aplicada.

No caso da compressão de imagens, existe um *trade-off* entre a taxa de compressão e a qualidade da informação pois quanto mais comprimida a imagem menor será a sua qualidade e desta forma é importante que o usuário possa definir um limite inferior para as suas necessidades. Os níveis de fidelidade estabelecidos foram baseados em critérios de análise observada e para as opções disponíveis foram associadas faixas de valores que correspondem a conjuntos distintos de taxas de compressão que podem ser aplicadas. Deve-se observar que as faixas de valores definidas correspondem a representações que oferecem uma qualidade mínima de visualização.

O parâmetro de quantidade mínima de opções de endereços é um dos principais parâmetros que contribuem para a definição da forma de execução da adaptação do processamento. Dependendo da escolha do usuário, a aplicação poderá oferecer pelo menos a quantidade de endereços que for definida. Com esta informação a aplicação poderá reduzir a quantidade de informação que será processada e armazenada no cliente.

Capítulo 5

Testes e validação do sistema

Neste capítulo é feita uma descrição dos testes que foram realizados com a aplicação e com o simulador para avaliar a eficácia do sistema. O objetivo dos testes inicialmente foi verificar o funcionamento da aplicação através dos resultados obtidos com a utilização de alguns parâmetros de consulta. Verificou-se também o funcionamento do sistema de adaptação através da sua reação em relação a alguns parâmetros de ambiente e de qualidade de serviço.

5.1 Consultas ao serviço

Para a verificação do funcionamento da aplicação foram utilizados alguns exemplos de consulta como entrada para o simulador. Como já foi dito anteriormente, o simulador da entidade cliente utiliza endereços no formato texto como entrada para a aplicação. Os endereços utilizados nestas consultas são dados reais que fazem parte da base de dados fornecida pela Prodabel.

A primeira consulta foi feita usando o seguinte endereço como exemplo: Rua dos Timbiras, nº 1986. A figura 5.1 mostra a lista de ocorrências de endereços mais próximos listados em ordem de proximidade com a localização do usuário.

O resultado da busca por endereços mais próximos reflete a proximidade geográfica entre os endereços através do cálculo da distância euclidiana entre as coordenadas correspondentes. Entretanto, como pode ser visto pelo exemplo da figura 5.3, para se chegar do ponto de origem ao destino, mesmo pela proximidade apresentada, o usuário deverá percorrer um caminho mais longo do que a distância real existente entre os dois pontos. Um

Rua São Paulo 1409
 Rua São Paulo 1401
 Rua dos Timbiras 1958
 Rua São Paulo 1399
 Rua dos Timbiras 1989

Figura 5.1: Ocorrências de endereços mais próximos a localização do usuário

dos fatores que ocasionam este comportamento se deve ao fato de que para o algoritmo criar a rota entre os endereços de origem e destino, inicialmente ele deve associar estes endereços a nodos distintos que façam parte do grafo. Após feita a associação a rota pode ser traçada entre os respectivos nodos.

Como a base de dados está configurada para rotas feitas por veículos automotores, considerando sentidos e conversões permitidas, e se o ponto de origem se localiza num logradouro de sentido único, então para chegar ao destino o usuário deverá seguir a rota pelos sentidos permitidos.



Figura 5.2: Rota entre a origem e a primeira ocorrência de endereços mais próximos

No resultado da consulta feita à aplicação utilizando o endereço de exemplo, como mostra a figura 5.2, a rota de menor caminho entre o endereço de origem até a primeira ocorrência de endereço mais próximo segue o caminho ilustrado na figura 5.3. Como neste trecho da Rua Timbiras o sentido é único, o usuário é obrigado a virar na Rua Rio de Janeiro, depois virar na Avenida Álvares Cabral, virar na Avenida Bias Fortes e por fim

virar na Rua São Paulo.

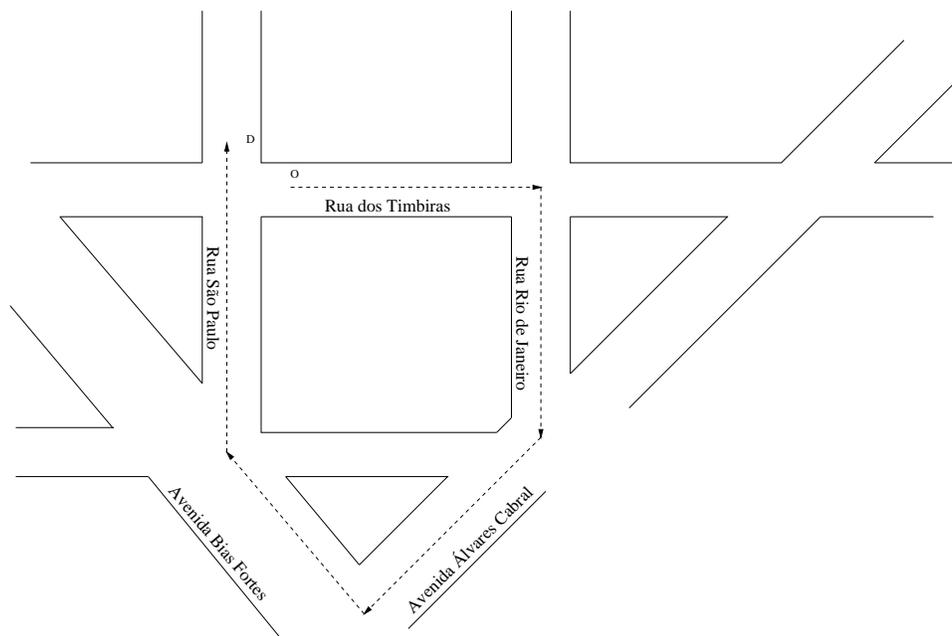


Figura 5.3: Caminho permitido entre a origem e o destino

Para um usuário que queira se locomover utilizando um automóvel e que não conhece a região ou mesmo a cidade esta informação é muito importante. Um outro exemplo que ilustra esta situação é o caso de um usuário que queira se locomover do endereço *Avenida Amazonas, nº 2285* até o endereço *Avenida Amazonas, nº 2311*. Mesmo estando situados bem próximos um do outro, estes endereços se encontram em vias de diferentes sentidos do mesmo logradouro. A rota de menor caminho entre estes endereços pode ser vista na figura 5.4.

Avenida Amazonas
 Conversão
 Rua Tompson Flores
 Conversão
 Rua Sertões
 Avenida do Contorno
 Conversão
 Avenida do Contorno
 Conversão
 Rua Bernardino de Lima
 Conversão
 Avenida Francisco Sá
 Conversão
 Avenida Amazonas

Figura 5.4: Rota entre os endereços de número 2285 e 2311 na Avenida Amazonas



Figura 5.5: Exemplo de um logradouro fragmentado pelo recorte

Como a base de dados utilizada no protótipo da aplicação foi retirada de uma região da área central da cidade de Belo Horizonte, algumas observações foram feitas em relação a integridade das respostas geradas. No caso de consultas feitas a endereços que fazem parte de logradouros compostos por somente um nodo, ou seja, segmentos de logradouros que têm apenas parte de sua extensão por se encontrarem na região periférica do recorte, o algoritmo não consegue encontrar o caminho mínimo e gera uma resposta negativa de “Rota não disponível” como mostra o exemplo na figura 5.5.

5.2 Adaptação

Nesta seção são descritos os testes que foram feitos com o simulador para se analisar os resultados da adaptação. As consultas foram feitas utilizando algumas variações dos valores dos parâmetros de QoS e das condições do meio. Entretanto, antes de iniciar a descrição dos experimentos, serão feitas algumas observações em relação aos valores escolhidos para os testes. Cada um dos parâmetros pode possuir valores dentro da faixa definida pelo próprio usuário e esses valores devem ser definidos de acordo com as unidades de medida previamente estabelecidas.

Para analisar o protótipo desenvolvido alguns valores são mais significativos para o estabelecimento de limiares no processo de adaptação. Desta forma, foram utilizados valores que exprimissem condições precárias do ambiente de acordo com o total de informação gerado pelas consultas. Como já visto no capítulo 4, o total de informação gerado corresponde à soma do tamanho das imagens de cada opção da resposta. A partir deste total o sistema pode então calcular a disponibilidade do meio e do dispositivo em conter a informação gerada. Dependendo do valor que for calculado, o sistema primeiramente consulta os parâmetros de qualidade de serviço e depois executa a adaptação necessária.

5.2.1 Bateria

Para efeito de testes parte-se do pressuposto de que o uso da carga de bateria em dispositivos geralmente segue um comportamento linear decrescente. Para tentar representar este comportamento o algoritmo de adaptação atribui um decréscimo de x em $\log_t x = 0,5$, onde t é o tempo gasto com o processamento. Para obter o tempo gasto o cliente captura o tempo de sistema consumido com a requisição, o processamento no servidor, compactação, transmissão e decodificação das imagens no cliente. A faixa padrão de variação de quantidade de bateria vai de 0 a 100 de onde é decrementado x unidades.

A restrição feita ao sistema em relação a este parâmetro é que o tempo mínimo definido pelo usuário não ultrapasse a quantidade de bateria capaz de operar durante este período.

Neste caso, se o usuário estiver com 4% de bateria disponível, por exemplo, o tempo definido pelo usuário não poderá exceder 16 segundos. Se o dispositivo estiver com esta quantidade de bateria disponível e o usuário escolher o tempo maior do que 15 segundos para resposta o sistema emitirá um alerta como mostra a figura 5.6.

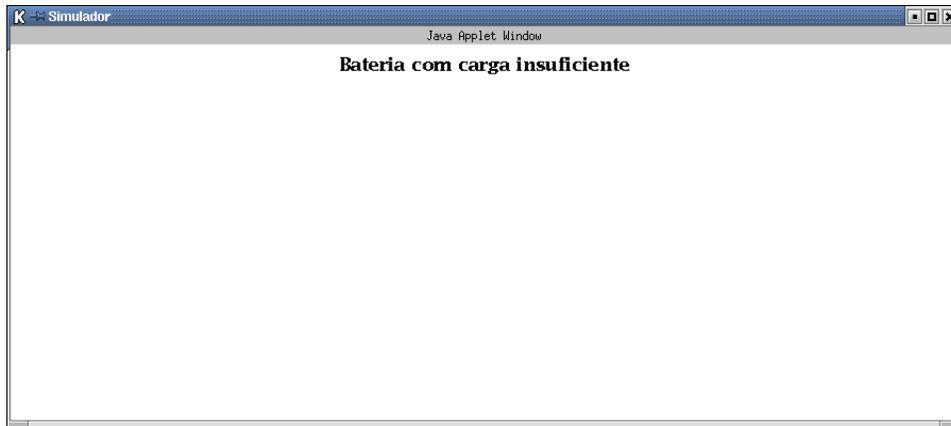


Figura 5.6: Tela para indicação de quantidade insuficiente de carga da bateria

5.2.2 Largura de banda

A largura de banda disponível é um parâmetro que define a quantidade de informação que pode ser transmitida de acordo com o tempo definido pelo usuário. Sendo assim, se o usuário definir um tempo mínimo de 5 segundos, por exemplo, e a largura de banda disponível for de 300 Kbps, como mostra a figura 5.7, então o sistema poderá transmitir até 183,1 Kbytes de dados aproximadamente.

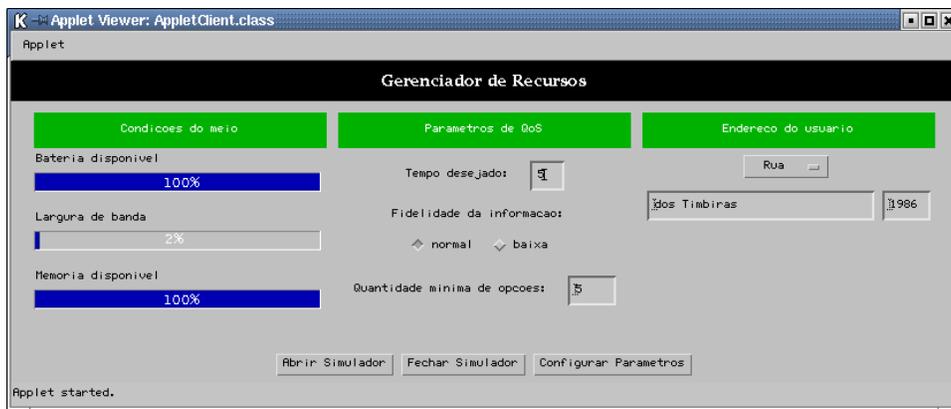
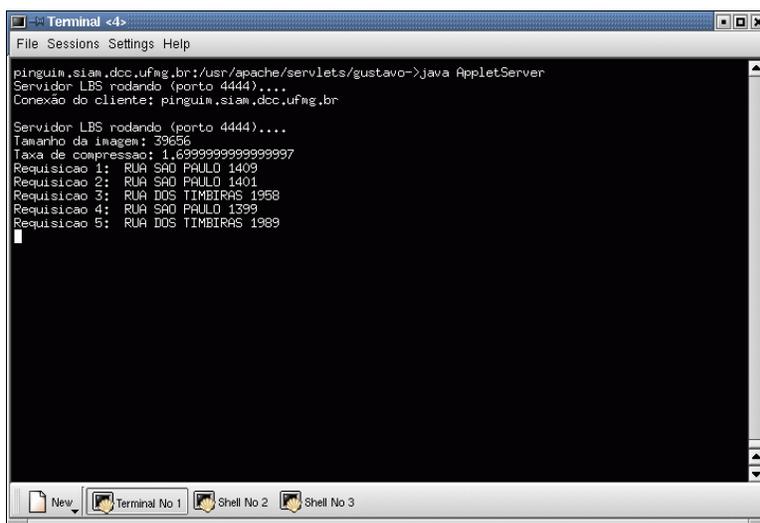


Figura 5.7: Gerenciador configurado para uma largura de banda de 300 Kbps e com tempo de resposta de 5 segundos

Se o usuário estiver interessado em imagens de alta resolução e em uma quantidade padrão de 5 opções como resposta, por exemplo, deve-se considerar que a soma dos tamanhos das imagens estáticas para 5 opções corresponde a 5 vezes o tamanho da imagem comprimida a uma taxa de alta resolução (por exemplo, 2 bpp), totalizando aproximadamente 193,6

Kbytes. Observa-se claramente que com estas condições o sistema fica impossibilitado de atender o usuário. Entretanto, com a capacidade de adaptação, o sistema e a aplicação podem contornar este problema aplicando a compressão nas imagens da resposta, como pode ser visto nas figuras 5.8 e 5.9.

A figura 5.8, que representa a tela de *log* do servidor da aplicação, mostra o tamanho inicial da imagem utilizando a taxa de 2 bpp e logo abaixo a taxa de compressão de 1,7 bpp que foi aplicada para que a requisição pudesse ser atendida. Deve-se ressaltar que todas as imagens utilizadas na aplicação são inicialmente compactadas com resolução máxima. Este valor inicial de compressão serve de referência para a aplicação de outras taxas de compressão (valor base para a predição). Na figura 5.9, que representa a tela de *log* do *proxy*, são exibidos os tamanhos das imagens (33328 bytes) após a compressão sugerida pela aplicação. Para as 5 opções de resposta totalizaram-se 162,73 Kbytes, satisfazendo desta maneira uma quantidade menor ou igual a quantidade máxima permitida.



```

Terminal <4>
File Sessions Settings Help

pinguin.siam.dcc.ufmg.br:/usr/apache/servlets/gustavo->java AppletServer
Servidor LBS rodando (porto 4444)....
Conexão do cliente: pinguin.siam.dcc.ufmg.br

Servidor LBS rodando (porto 4444)....
Tamanho da imagem: 39656
Taxa de compressão: 1.6999999999999997
Requisicao 1: RUA SAO PAULO 1409
Requisicao 2: RUA SAO PAULO 1401
Requisicao 3: RUA DOS TIMBIRAS 1958
Requisicao 4: RUA SAO PAULO 1399
Requisicao 5: RUA DOS TIMBIRAS 1989

```

Figura 5.8: Tela de log do servidor

Mesmo que uma imagem tenha sido comprimida a uma taxa menor do que a taxa considerada de alta resolução, pode-se considerar que imagens comprimidas até 1 bpp são de alta resolução [11].

5.2.3 Memória no terminal

A quantidade de memória disponível no terminal é um parâmetro crítico pois caso o cliente não tenha condições de armazenar os dados provenientes das consultas, o sistema

```

Shell - Terminal <2>
File Sessions Settings Help

pinguin.siam.dcc.ufmg.br:/usr/apache/servlets/gustavo->.java AppletListener
ProxyAdapter rodando (porto 2222)....
Conexão de cliente.

Requisicao 1
  Target bitrate = 1.7 bpp (i.e. 33429 bytes)
  Achieved bitrate = 1.6948524 bpp (i.e. 33328 bytes)
  Tamanho da imagem depois da compressao: 33328

Requisicao 2
  Target bitrate = 1.7 bpp (i.e. 33429 bytes)
  Achieved bitrate = 1.6948524 bpp (i.e. 33328 bytes)
  Tamanho da imagem depois da compressao: 33328

Requisicao 3
  Target bitrate = 1.7 bpp (i.e. 33429 bytes)
  Achieved bitrate = 1.6948524 bpp (i.e. 33328 bytes)
  Tamanho da imagem depois da compressao: 33328

Requisicao 4
  Target bitrate = 1.7 bpp (i.e. 33429 bytes)
  Achieved bitrate = 1.6948524 bpp (i.e. 33328 bytes)
  Tamanho da imagem depois da compressao: 33328

Requisicao 5
  Target bitrate = 1.7 bpp (i.e. 33429 bytes)
  Achieved bitrate = 1.6948524 bpp (i.e. 33328 bytes)
  Tamanho da imagem depois da compressao: 33328

```

Figura 5.9: Tela de *log* do *proxy*

deverá reagir com alguma ação. Assim como ocorre com os outros parâmetros citados anteriormente, se houverem condições de se aplicar a adaptação, o sistema e a aplicação tomarão as devidas providências, mas caso isso não seja possível o cliente recebe uma resposta negativa do servidor.

Para definir um cenário que exemplifique a necessidade de adaptação diante de uma quantidade escassa de memória basta que se reduza o seu valor. Considerando que este valor seja de 45 Kbytes, por exemplo, e que o usuário deseja visualizar 5 opções como resposta, ele deverá mudar o nível de fidelidade para o valor “baixo” para conseguir obter a resposta. Neste momento o usuário deverá optar entre os parâmetros que sejam mais adequados às suas necessidades pois cada escolha acarreta em um *trade-off* em relação às outras características existentes.

Como já mencionado neste capítulo, para transmitir 5 opções de resposta com imagens de alta qualidade são necessários 193,6 Kbytes. No entanto, mais uma vez fica claro que a quantidade de informação gerada não deve ser maior do que 45 Kbytes, ou seja, o sistema de adaptação deverá aplicar uma taxa de compressão sobre as imagens para não ultrapassar o limite permitido.

Sabendo que para fidelidade normal as taxas de compressão só podem ir até o valor de 0,7 bpp, tem-se então que para 5 opções são gerados 67,1 Kbytes. Este valor ainda não satisfaz as condições existentes e por isso a escolha da fidelidade “baixa”, com valores que vão até 0,2 bpp, deve ser feita pois para 5 opções tem-se como resultado uma quantidade de 19,1 Kbytes. Neste caso o sistema pode aplicar uma taxa um pouco maior do que a

mínima de 0,2 bpp. O resultado da aplicação deste cenário pode ser visualizado na figura 5.10.



Figura 5.10: Requisição de um dispositivo com pouca memória disponível

5.3 Considerações sobre os testes

Os testes que foram realizados utilizando o simulador tiveram como principal objetivo testar a arquitetura em relação à disponibilização do serviço desenvolvido levando-se em consideração a qualidade das respostas assim como o comportamento do mesmo diante de alguns cenários de restrições. É bastante relevante o avanço que se tem com a possibilidade de definição de parâmetros de QoS realimentando o sistema pois, assim como proposto, o processo de adaptação é executado obedecendo sempre os parâmetros mínimos definidos pelos usuários.

Os resultados obtidos com os testes trazem uma importante contribuição pois a partir destes pode-se destacar que a necessidade de uma arquitetura diferenciada como a que foi proposta é realmente um diferencial na oferta de serviços para o ambiente móvel. As preocupações com as restrições existentes deverão ainda permanecer por muito tempo pois, como pode ser visto pelas tecnologias desenvolvidas mais recentemente ou mesmo pelas tendências das próximas gerações, os principais fatores limitantes para o uso dos serviços são as características ligadas ao meio sem fio.

Capítulo 6

Conclusões e Trabalhos Futuros

Neste capítulo são apresentadas as considerações finais relacionadas a este trabalho. São feitos também comentários a respeito do estudo e do processo de desenvolvimento que levaram às conclusões aqui apresentadas.

6.1 Conclusões

Com o atual desenvolvimento do mercado de serviços para dispositivos móveis de acesso a dados a demanda por aplicações úteis vem aumentando cada vez mais. Aplicações que apresentam informações simples, que são entregues de forma rápida e segura e que levem em consideração a localização física do usuário terão um grande diferencial pois poderão personalizar cada vez mais certos eventos que, outrora, não poderiam ser concebidos.

Entretanto, para que estas aplicações tenham sucesso entre os usuários, os projetistas deverão se preocupar agora com o surgimento de novos requisitos relacionados a problemas do meio e do dispositivo. Este é um dos principais fatores que tem sido discutido pela academia pois traz à tona o problema de responsabilidade pelo processo de adaptação, fator essencial para a concepção dos novos projetos. Cada uma das abordagens existentes apresentam vantagens e desvantagens mas a estratégia do tipo *Application-aware adaptation* revela maior flexibilidade pois permite a cooperação entre o sistema e a aplicação.

Levando-se em consideração o ponto levantado anteriormente, este trabalho procurou descrever uma arquitetura que fosse colaborativa e conseqüentemente flexível para as possíveis aplicações. Para validar o funcionamento do arcabouço proposto, o desenvolvimento do protótipo da aplicação baseada na localização do usuário foi de fundamental importân-

cia para o estudo feito neste trabalho. Apesar da investigação feita ter sido iniciada com o estudo de serviços baseados na localização do usuário e suas características, vários fatores levaram à consideração das arquiteturas adaptativas pois a natureza do ambiente móvel influencia profundamente o projeto de qualquer aplicação.

A conclusão de que a adaptação é extremamente necessária em ambientes onde as características são tão variáveis quanto no ambiente móvel não é mais novidade. Resta agora que se estabeleçam modelos de arquitetura adaptativa para a viabilização das futuras aplicações. Este trabalho pode ser considerado como uma contribuição em ambos os sentidos. Os testes executados mostram que a eficácia do sistema em relação a sua capacidade de reação a parâmetros de restrições é bastante atraente pois o usuário pode ser atendido mesmo em condições desfavoráveis de uso do serviço.

6.2 Trabalhos futuros

Como trabalhos futuros podem ser citadas várias características da arquitetura e do protótipo que poderiam ser desenvolvidas. Em relação a aplicação, a adição de uma tabela à base existente com dados de provedores de serviço como restaurantes, bares, hospitais, etc., seria inevitável pois a busca especializada por provedores de serviços poderia sofisticar ainda mais a aplicação.

Outro elemento muito importante para tornar a aplicação mais viável seria estudar uma forma de se agregar uma base de dados vetorial para a geração dinâmica dos mapas de cada trecho consultado. Da forma como o protótipo se apresenta atualmente, os mapas não refletem a região consultada pois ele utiliza imagens estáticas. Uma outra investigação muito importante que poderia ser feita seria o estudo da possibilidade de utilizar transmissões vetoriais progressivas [8]. A transmissão progressiva de imagens no tipo ráster pela *web*, por exemplo, tem sido muito utilizada para oferecer aos usuários representações mais grosseiras da imagem original antes que esta última consiga ser totalmente carregada no cliente. Da mesma forma, o estudo de transmissões vetoriais progressivas estão sendo objeto de estudo pois apresentam vários desafios em relação a imagens vetoriais, como é o caso dos mapas.

Testar a utilização de outros algoritmos de compressão de imagens além do formato utilizado neste trabalho. Devido a um conjunto de fatores decidiu-se por utilizar o padrão JPEG2000. O estudo mais aprofundado sobre a utilização de outros formatos de imagem a partir de considerações técnicas, pode trazer vantagens ao processo de adaptação desde

que se leve em consideração os requisitos do meio sem fio.

Especificar e implementar gerenciadores para outros tipos de dados no cliente e no *proxy*. A possibilidade da adaptação de informações como vídeo, por exemplo, é muito importante para garantir a robustez da arquitetura. Além disso, com o futuro advento das redes sem fio de banda larga, o uso de recursos multimídia em dispositivos móveis será muito comum, mesmo sob as restrições já conhecidas.

Para este trabalho em específico decidiu-se por utilizar o padrão JPEG

Outro caminho para trabalhos futuros seria o desenvolvimento de novos serviços baseados na localização adaptativos para a arquitetura aqui proposta. Por ser uma arquitetura aberta torna-se bastante viável a utilização de outros serviços baseados na localização. Seguem abaixo algumas sugestões que podem ser desenvolvidas.

- Vídeos baseados na localização

Vídeos informativos sobre qualquer endereço que esteja mais próximo da localização do usuário. O serviço de vídeo pode ser requisitado para visualização de clipes sobre algum endereço que seja do interesse do usuário. A partir do serviço de procura por endereços mais próximos o usuário obterá uma lista de endereços que satisfaçam a sua requisição e irá então selecionar a opção de visualização caso esta esteja disponível para o endereço escolhido.

- Acesso a servidores *web* baseado na localização

Esse serviço é implementado através da utilização da informação de localização no contexto de sites e portais da *web*. Cada site ou portal poderá ser referenciado por uma área de localização geográfica e cada uma das áreas poderá conter diversas categorias de serviços que poderão ser escolhidos pelo usuário. Informações contidas no site e que estão disponíveis para o acesso através de dispositivos móveis podem ser oferecidos na forma de texto, imagens ou mesmo através de vídeos.

Bibliografia

- [1] Nokia Corporation. <http://www.nokia.com>, May 2001.
- [2] 3GPP. Stage 2 Functional Specification of Location Services in UTRAN. Tech. Rep. 3G TS 25.305, December 1999.
- [3] ACHARYA, A., BADRINATH, B., IMIELINSKI, T., AND NAVAS, J. A WWW-based Location-dependent Information Service for Mobile Clients. Tech. rep., Rutgers University, July 1995.
- [4] ANDRÉ, F., AND SEGARRA, M. T. A Generic Approach to Satisfy Adaptability Needs in Mobile Environments. In *Proceedings of the 33rd Hawaii International Conference on System Sciences* (June 1999), no. 3723.
- [5] ANGIN, O., CAMPBELL, A. T., KOUNAVIS, M. E., AND LIAO, R. R.-F. The Mobile Toolkit: Programmable Support for Adaptive Mobile Networking. In *IEEE Personal Communications, Special Issue on Adapting to Network and Client Variability* (August 1998), vol. 5, pp. 32–44.
- [6] American National Standards Institute. <http://www.ansi.org/>, May 2001.
- [7] BAUMANN, S., COLLOMB, F., DIEN, H. P., PARIS, G., PILLONI, P., AND CASAL, C. R. LOCUS Project - Overview of Location Services. Tech. Rep. IST-1999-14093 LOCUS, Information Society Technologies - The European Commission, February 2001.
- [8] BERTOLOTTO, M., AND EGENHOFER, M. J. Progressive vector transmission. In *ACM-GIS* (1999), pp. 152–157.
- [9] The Official Bluetooth SIG Website. <http://www.bluetooth.com>, March 2001.

- [10] Cellpoint Corporation. <http://www.cellpoint.com>, March 2001.
- [11] CHRISTOPOULOS, C., SKODRAS, A., AND EBRAHIM, T. The JPEG2000 Still Image Coding System: an Overview. *IEEE Transactions on Consumer Electronic* 46, 4 (November 2000), 1103–1127.
- [12] DAVIES, N., BLAIR, G., CHEVERST, K., AND FRIDAY, A. Supporting Adaptive Services in a Heterogeneous Mobile Environment. In *Proceedings of the 1st Workshop on Mobile Computing Systems and Applications (MCSA '94)* (Santa Cruz, California, U.S., December 1994), pp. 153–157.
- [13] DAVIS JR., C.A. AND FONSECA, F.T. Endereços: A base de um Projeto de Geoprocessamento Urbano. Notas de Aula, 1995.
- [14] DAVIS JR., C.A. AND FONSECA, F.T. Conceitos e Aplicações em GIS. Notas de Aula, 1999.
- [15] Federal Communications Commission - Enhanced 911 Home Page. <http://www.fcc.gov/e911/>, May 2001.
- [16] EFSTRATIOU, C., CHEVERST, K., DAVIES, N., AND FRIDAY, A. Architectural requirements for the effective suport of adaptive mobile applications. In *To appear in the MDM2001 (Mobile Data Management)* (July 2001).
- [17] EIA/TIA. Cellular Radio-Telecommunications Intersystem Operations. Revision C IS-41, Electronics Industry Alliance/Telecommunications Industry Association, 1995.
- [18] ETSI. Digital cellular telecommunications system (Phase 2+); Location Services (LCS); Service description, Stage 1. Tech. Rep. GSM 02.71, European Telecommunications Standards Institute, 1998.
- [19] ETSI. Digital Cellular Telecommunications System (Phase 2+); Location Services (LCS); (Functional Description) - Stage 2. Tech. Rep. GSM 03.71, European Telecommunications Standards Institute, 1999.
- [20] ETSI. Digital cellular telecommunications system (Phase 2+); Technical realization of Cell Broadcast Service (CBS). Tech. Rep. GSM 03.41, European Telecommunications Standards Institute, 1999.

- [21] ETSI. Digital cellular telecommunications system (Phase 2+)/Specification of the Subscriber Identity Module -Mobile Equipment (SIM-ME) interface. Tech. Rep. GSM 11.11, European Telecommunications Standards Institute, 1999.
- [22] European Telecommunications Standards Institute. <http://www.etsi.org>, May 2001.
- [23] FORMAN, G., AND ZAHORJAN, J. The Challenges of Mobile Computing. *IEEE Computer* (Apr. 1994), 39–47. <ftp://ftp.cs.washington.edu/tr/1993/11/UW-CSE-93-11-03.PS.Z>.
- [24] FOX, A., BREWER, E., GRIBBLE, S., AND AMIR, E. Adapting to Network and Client Variability via On-Demand Dynamic Distillation. In *Proceedings of the Seventh International ACM Conference on Architectural Support for Programming Languages and Operating Systems*. (Cambridge, MA USA, October 1996).
- [25] FRIDAY, A., DAVIES, N., BLAIR, G., AND CHEVERST, K. Developing Adaptive Applications: The MOST Experience. *Journal of Integrated Computer-Aided Engineering* 6, 2 (1999), 143–157.
- [26] LBS Requirements Document. <http://www.gsmworld.com/technology>, January 2000. North American Interest Group of the GSM Mou Association.
- [27] NTT Docomo i-Mode Technology. <http://www.nttdocomo.com>, March 2001.
- [28] International Mobile Telecommunications-2000 (IMT-2000). <http://www.itu.int/imt>, February 2001.
- [29] ISO/IEC JTC1 SC29 /WG 1. JPEG 2000 Part 1 Final Draft International Standard Release. Tech. rep., ISO/IEC, August 2000.
- [30] SUN Site - The Source for Java Technology. <http://www.java.sun.com>.
- [31] JJ2000 - An implementation of the JPEG2000 standard in Java. <http://jj2000.epfl.ch/>, April 2001.
- [32] KATZ, R. H. Adaptation and Mobility in Wireless Information Systems. *Personal Communications* 1, 1 (1994), 6–17. <http://daedalus.cs.berkeley.edu/publications/Personal.ps.gz>.

- [33] KUNZ, T., AND BLACK, J. P. An architecture for adaptive mobile applications. In *Proceedings of Wireless 99, the 11th International Conference on Wireless Communications* (Calgary, Alberta, Canada, July 1999), pp. 27–38.
- [34] LOPEZ, X. R. GeoJava for Internet and Mobile Location Services. <http://www.jlocationsservices.com/company/Oracle/>, 2000.
- [35] LU, S., LEE, K.-W., AND BHARGHAVAN, V. Adaptive Service in Mobile Computing Environments. In *Proceedings of the 5th International Workshop on Quality of Service (IWQOS'97)* (Columbia University, New York, USA, 1997), pp. 25–36.
- [36] MATEUS, G. R., AND LOUREIRO, A. A. F. Introdução à computação móvel. 11a. Escola de Computação, Julho 1998.
- [37] MOEGLEIN, M., AND KRASNER, N. An introduction do SnapTrack Server-Aided GPS Technology. In *Proc. of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS-98)* (September 1998), p. 333.
- [38] MOULY, M., AND PAUTET, M. B. *The GSM System for Mobile Communications*. No. ISBN 2-9507190-0-7. Cell and Sys, Palaiseau, France, 1992.
- [39] NAUGHTON, P., AND SCHILDT, H. *Java 2: The Complete Reference*, third ed. McGraw-Hill, Berkeley, California, 1999.
- [40] NOBLE, B. System support for mobile, adaptive applications. *IEEE Personal Computing Systems* 7, 1 (February 2000), 44–9.
- [41] NOBLE, B. D., AND SATYANARAYANAN, M. Experience with Adaptive Mobile Applications in Odyssey. *Mobile Networks and Applications* (1999).
- [42] NOBLE, B. D., SATYANARAYANAN, M., NARAYANAN, D., TILTON, J. E., FLINN, J., AND WALKER, K. R. Agile Application-Aware Adaptation for Mobility. In *Sixteen ACM Symposium on Operating Systems Principles* (Saint Malo, France, Oct. 1997), pp. 276–287. <http://www.cs.cmu.edu/afs/cs/project/coda/Web/docdir/s16-reprint.ps.Z>.
- [43] Qualcomm and Lucent Technologies announce Wireless Location Development Project. <http://www.lucent.com/press/0799/990715.nsa.html>, March 2001.

- [44] SANTOS, S. M. S. A. O uso de técnicas de compressão no projeto de protocolos para comunicação móvel. Master's thesis, Universidade Federal de Minas Gerais, Março 1999.
- [45] SATYANARAYANAN, M. Mobile Information Access. *IEEE Personal Communications* 3, 1 (Feb. 1996).
<http://www.cs.cmu.edu/afs/cs/project/coda/Web/coda.html>.
- [46] SATYANARAYANAN, M., AND NARAYANAN, D. Multi-fidelity Algorithms for Interactive Mobile Applications. In *Proceedings of the 3rd international workshop on Discrete algorithms and methods for mobile computing and communications* (Seattle, WA USA, August 1999), pp. 1–6.
- [47] Signalsoft Wireless Location Services. <http://www.signalsoftcorp.com>, March 2001.
- [48] SILVA, F. B. Páginas Amarelas na WWW: Apresentação de Resultados de Acordo com a Localização Geográfica. Master's thesis, Universidade Federal de Minas Gerais, Março 1999.
- [49] SWEDBERG, G. Ericsson's Mobile Location Solution.
http://www.ericsson.com/review/1999_04/article93.shtml, 1999.
- [50] Symbian - Wireless Technology. <http://www.symbian.com>, April 2001.
- [51] TANENBAUM, A. S. *Computer Networks*, third ed. Prentice Hall Inc., Upper Saddle River, New Jersey, 1996.
- [52] VOELKER, G. M., AND BERSHAD, B. N. Mobisaic, An Information System for a Mobile Wireless Computing Environment & Engineering. In *IEEE Workshop on Mobile Computing Systems and Applications* (Santa Cruz, CA, US, Dec. 1994).
<http://snapple.cs.washington.edu/mobile/mcsa94.html>.
- [53] WAP - Wireless Application Protocol. <http://www.wapforum.org>, March 2001.
- [54] Webraska Wireless Navigation Worldwide. <http://www.webraska.com>, March 2001.

Glossário

3G Third Generation

A-GPS Assisted GPS

AOA Angle of Arrival

BTS Base Transceiver Station

CGI Cell Global Identity

Cell-ID Cell Identification

DGPS Differential GPS

E-OTD Enhanced Observed Time Difference

GPS Global Positioning System

GSM Global System for Mobile Communication

HLR Home Location Register

HTML Hypertext Markup Language

HTTP Hypertext Transfer Protocol

IN Intelligent Network

IS-41 Interim Standard 41

JDBC Java DataBase Connectivity

JPEG Joint Photographic Experts Group

LAC Location Area Code

LBS Location Based Services

LCS LoCation Services

LSIF Location System Information Function

MAP Mobile Application Part

MCC Mobile Country Code

MNC Mobile Network Code

MSISDN Mobile Station ISDN Number

NMR Network Measurement Results

OTD Observed Time Difference

PCF Position Calculation Function

PCS Personal Communications Service

PLMN Public Land Mobile Network

QoS Quality of Service

RTD Relative Time Difference

SAI Service Area Identifier

SIG Sistemas de Informação Geográfica

SIM Subscriber Identity Module

SMLC Servinig Mobile Location Centre

SMS-C Short Message Service Center

SMS Short Message System

SRNC Serving Radio Network Controller

STK Subscriber Identity Module Application Toolkit

TA Timing Advance

TCP/IP Transport Control Protocol/Internet Protocol

TDOA Time Difference Of Arrival

TOA Uplink Time Of Arrival

UE User Equipment

UMTS Universal Mobile Telecommunications System

URL Uniform Resource Locator

VLR Visitor Location Register

WAP Wireless Application Protocol

WML Wireless Markup Language

XML Extensible Markup Language

Apêndice A

Padronização dos serviços de localização

Com o atual estágio de desenvolvimento de tecnologias para serviços baseados na localização do usuário algumas padronizações foram e ainda estão sendo definidas. Seguem a seguir algumas definições de tecnologias e equipamentos que foram citados no texto desta dissertação.

A.1 Definições de padrões

A.1.1 Location estimate

É a localização geográfica do dispositivo móvel, expressada em dados de latitude e longitude. A estimativa *Location estimate* deve ser representada em um formato universal bem definido. A tradução deste formato universal para outro sistema de localização geográfico deve ser suportada.

A.1.2 Mobile Assisted positioning

Qualquer método de localização baseado no dispositivo (como por exemplo E-OTD e GPS) no qual a unidade móvel disponibiliza as medidas de posicionamento para a rede para o cálculo da localização estimada pela rede. A rede pode oferecer dados adicionais para a unidade móvel para possibilitar que este possa medir sua posição ou mesmo para melhorar o desempenho destas medidas.

A.1.3 Mobile Based Positioning

Qualquer método de localização baseado no dispositivo onde a unidade móvel faz ambas as medidas de posicionamento e de cálculo de uma estimativa de localização.

A.1.4 Location request

- Mobile Originating Location Request (MO-LR): a partir de uma unidade móvel para um servidor LCS. Um MO-LR também pode ser usado para a requisição da localização de outra unidade móvel (outro cliente LCS).
- Mobile Terminating Location Request (MT-LR): a partir de um cliente externo à PLMN.
- Network Induced Location Request (NI-LR): requisição de uma unidade móvel a partir de um cliente que pode ser considerado interno a qualquer entidade da PLMN que serve esta unidade móvel.

A.1.5 LoCation Services (LCS)

O LCS é um conceito de serviço na padronização dos sistemas, como por exemplo os sistemas GSM e UMTS. Ele especifica todos os elementos de rede e entidades necessários, suas funcionalidades, interfaces, tão bem quanto as mensagens de comunicação, para a implementação das funcionalidades dos serviços de localização em uma rede celular. O LCS não especifica nenhum serviço baseado na localização (como por exemplo serviços de valor agregado), exceto aqueles relacionados às chamadas de emergência.

A.1.6 Padronização de tecnologias de localização para GSM

Segue a seguir uma lista das padronizações das tecnologias de localização utilizadas em redes GSM:

- Timing Advance - TA [19]
- Time Of Arrival - TOA [19]
- Enhanced Observed Time Difference - E-OTD [19]
- Global Positioning System - GPS [19]

A.1.7 Padronização da arquitetura GSM

A descrição dos elementos de rede e suas interface/pontos de referência são listadas a seguir. Deve-se ressaltar que há somente um ponto de referência entre o servidor LCS PLMN e o cliente LCS chamado Le. O Le permite o *roaming* se uma interface apropriada for definida. Sua descrição se encontra em [18].

Pode haver mais de uma simples interface de rede LCS para diferentes clientes LCS ou outras redes. Há uma interface inter-LCS PLMN chamada Lg que conecta duas redes LCS independentes para troca de mensagens.

Base Station Subsystem (BSS)

O BSS é um elemento envolvido na manipulação de vários procedimentos de localização.

Gateway Mobile Location Centre (GMLC)

O GMLC é o primeiro nodo que um cliente LCS externo acessa o GSM PLMN. O GMLC pode requisitar informação de roteamento do HLR através da interface Lh. Depois de executar o registro de autorização, ele envia requisições de localização para e recebe estimativas finais de localização do VMSC através da interface Lg.

Serving Mobile Location Centre (SMLC)

O SMLC tem a funcionalidade necessária para dar suporte ao LCS. Em uma PLMN, onde pode haver mais de um SMLC, a interface Lp permite o acesso a informação e recursos entre mais de um SMLC. O SMLC gerencia toda a coordenação e escalonamento de recursos necessários para se localizar uma unidade móvel. Ele também calcula a estimativa e precisão da localização final. Dois tipos de SMLC são possíveis:

- NSS (Network Software Solution) baseado no SMLC suporta a localização de uma MS através da sinalização pela interface Ls da MSC visitada.
- BSS baseado no SMLC suporta a localização através da sinalização na interface Lb com o BSC que serve à MS.

O SMLC controla um número de LMUs (para o propósito de se obter as medidas de interface de rádio para localizar assinantes MS). As funcionalidades do SMLC e do GMLC podem ser combinadas no mesmo nodo físico, combinadas em nodos físicos já existentes ou

estar em diferentes nodos. Para serviços de localização, quando um Cell Broadcast (CBC) é associado com um BSC, o SMLC pode ter uma interface com o CBC para difundir dados de assistência [20].

Location Measurement Unit (LMU)

Uma LMU faz medidas do sinal de rádio para suportar um ou mais métodos de localização. Há dois tipos de medidas, localização e medidas de assistência. Todas as medidas obtidas por um LMU são fornecidos para um SMLC em particular. Dois tipos de LMU são definidos:

- LMU do tipo A: acessado exclusivamente sobre interface GSM. Não há conexão física para qualquer elemento de rede.
- LMU do tipo B: acessado sobre uma interface (Abis) de um BSC. A LMU pode ser um elemento de rede *standalone* endereçado usando somente uma pseudo- cell-ID, conectado ou integrado em um BTS.

Medidas de assistência obtidas por uma LMU desde que elas sejam usadas em mais de um método de localização:

- Medidas de tempo de interfaces de rádio: ATDs (Absolute Time Differences) ou RTDs (Real Time Differences) dos sinais transmitidos pelas estações base, onde as diferenças de tempo são medidas em relação a alguma diferença de tempo absoluta (ATD) ou aos sinais de outra estação base (RTD).

Mobile Switching Centre (MSC)

O MSC tem funcionalidades que são responsáveis pela autorização de assinatura da MS ou pelo gerenciamento de requisições de localização relacionadas a chamadas ou não de GSM LCS. O MSC é acessível pelo GMLC através da interface Lg e pelo SMLC através da interface Ls.

Home Location Register (HLR)

O HLR contém os dados de assinaturas LCS e de roteamento. O HLR é acessível do GMLC através da interface Lh. A arquitetura LCS é planejada para suportar um alto grau de flexibilidade, onde qualquer SMLC físico pode suportar múltiplas interfaces Ls ou Lb

e onde uma mistura de diferentes tipos de SMLC podem servir uma simples rede ou uma simples área MSC.

A.1.8 Padronização de tecnologias de localização para UMTS

Segue a seguir uma lista das padronizações de tecnologias de localização utilizadas em redes UMTS:

- Cell-ID
- TDOA [2]
- Métodos Network Assisted GPS [2]

A.1.9 Padronização da arquitetura UMTS

Radio Network Controller (RNC) - Serving RNC

O servidor RNC (SRNC) é um elemento de rede do UTRAN (UMTS Radio Access Network) que contém funcionalidades necessárias para suportar o LCS em um PLMN. O SRNC gerencia toda a coordenação e escalonamento de recursos necessários para a localização de uma unidade móvel. Ele também calcula a estimativa de localização final e estima também a precisão alcançada. O SRNC pode controlar um número de LMUs para se obter medidas de interface de rádio com o objetivo de localizar os assinantes móveis na área que ele serve. O SRNC é administrado com as capacidades e tipos de medidas produzidos por cada um de suas LMUs. A sinalização entre um SRNC e uma LMU é transferida usando-se somente a interface Iub, algumas vezes a interface Iur (e também a interface Uu para possíveis LMUs *standalone*). A LMU retorna a um SRNC, a informação de tempo da interface de rádio. Esta informação pode ser usada para mais de um método de localização.

Location Measurement Unit (LMU)

A LMU faz algumas medidas (por exemplo, de sinais de rádio) e comunica estas medidas (por exemplo, funções) ao SRNC. A LMU contém a PSMF (Position Signal Measurement Function) [2] e pode também fazer cálculos associados com as medidas. Os detalhes das medidas a serem feitas pela LMU serão definidos pelo método LCS escolhido.

A LMU pode fazer suas medidas em resposta a requisições, de forma autônoma enviando a informação regularmente ou quando há alguma mudança significativa nas condições do sinal de rádio (por exemplo, mudanças no RTD). Podem haver um ou mais LMU associados com o UTRAN e a requisição LCS pode envolver medidas de uma ou mais LMU. Instruções relativas a tempo, a natureza e periodicidade destas medidas são também fornecidas pelo SRNC ou são pré-administradas no SRNC. A LMU pode ser de vários tipos e o SRNC irá selecionar as LMUs apropriadas dependendo do método LCS que está sendo usado.

User Equipment (UE)

O UE interage com as funções de coordenação de medidas para rotear os sinais necessários para o *uplink* baseado nas medidas do LCS e para fazer medidas de sinais de *downlink*. As medidas a serem feitas serão determinadas pelo método de localização escolhido. O UE pode também conter aplicações LCS, ou acessar uma aplicação LCS através da comunicação com a rede acessada pelo UE ou por uma aplicação do UE. Esta aplicação pode incluir as medidas necessárias e as funções de cálculo para determinar a localização do UE com ou sem assistência das entidades UTRAN LCS. O UE pode também, por exemplo, conter uma função de localização independente (por exemplo, GPS) e então ser capaz de informar sua localização, independente de qualquer transmissão UTRAN. O UE com a função de localização independente pode também fazer uso de informações difundidas pelo UTRAN que assiste a função.

Gateway Mobile Location Centre (GMLC)

O GMLC contém funcionalidades necessárias para suportar LCS. O GMLC é o primeiro nodo que um cliente LCS acessa em um PLMN GSM (por exemplo, o ponto de referência Le é suportado pelo GMLC). O GMLC pode requisitar informação de roteamento do HLR através da interface Lh. Depois de fazer a autorização do registro, ele envia uma requisição de localização para e recebe estimativas de localização do 3G-VMSC através da interface Lg.

3G-MSC/VLR

O 3G-MSC/VLR contém funcionalidades para autorização de registro de UEs e de gerência de requisições para LCS relacionadas ou não a chamadas. O 3G-MSC é acessível ao GMLC através da interface Lg. As funções LCS do 3G-MSC são relacionadas a carga e

cobrança, coordenação LCS, requisições de localização, autorização e operação de serviços LCS.

Home Location Register (HLR)

O HLR contém os dados de assinaturas de roteamento. O HLR é acessível pelo GMLC através da interface Lh. Para o *roaming* de UEs, o HLR pode ser de uma PLMN diferente da atual SRNC.