

CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA
FACULDADE DE TECNOLOGIA DE INDAIATUBA

Gustavo Nunes Queirós

**REDUÇÃO DE BANCO DE DADOS PARA UTILIZAÇÃO EM
DISPOSITIVOS MÓVEIS**

INDAIATUBA
Novembro/2010

Gustavo Nunes Queirós

REDUÇÃO DE BANCO DE DADOS PARA UTILIZAÇÃO EM DISPOSITIVOS MÓVEIS

Trabalho de Conclusão de Curso apresentado como pré-requisito para conclusão do Curso Superior de Tecnologia em Informática da Faculdade de Tecnologia de Indaiatuba, sob a orientação do Prof. Sérgio Furgeri

INDAIATUBA

Novembro/2010

SUMÁRIO

Introdução	8
1.1 Objetivos.....	11
1.1.1 Objetivo geral.....	11
1.1.2 Objetivos específicos	11
1.2 Justificativa	12
1.3 Metodologia	12
1.4 Visão Geral.....	13
1.5 Problematização:	14
1.6 Hipótese	14
Capítulo I	16
1. Redução de banco de dados	16
1.1 Protocolo baseado em prioridades.....	18
2. Técnicas de Redução de Banco de Dados	22
2.1 Técnicas modificadoras de domínio	22
2.1.1 Camadas de Redução de dados.....	22
2.1.2 Abstração através da agregação	23
2.1.3 Hierarquia de Conceitos.....	24
2.2 Técnicas modificadoras do esquema.....	26
2.2.1 Redução vertical por projeção de tuplas	26
2.2.2 Redução horizontal por seleção de tuplas.....	27
Capítulo II	28
1 A Família de SGBDs SQL Server	28
2.1 Características gerais do SQLSC	29
2.2 Manipulação de bancos de dados.....	31
2.3 Uso da linguagem SQL	31
2.4 Acesso multiusuário.....	32
2.5 Concorrência	32
2.6 Transações.....	33

2.7 Transferência de Dados – Cliente – Servidor	34
2.8 RDA	35
Capítulo III	37
1 Introdução.....	37
2 O software	38
2.1 Arquitetura:	38
2.2 Visão Geral:	39
2.3 Funcionamento:	40
3.1 Detalhamento da redução dos bancos de dados	43
4 Ambiente de Desenvolvimento.....	50
Considerações finais.....	51
Referências Bibliográficas	55
Apêndice A - Principais tipos de dados presentes no SQLSC	56

Lista de Figuras

Figura 1- Representação da fragmentação vertical.....	26
Figura 2- Representação da fragmentação horizontal.....	27
Figura 3- Arquitetura de funcionamento do sistema.....	39
Figura 4-Tela apresentada ao vendedor.....	41
Figura 5-Tela apresentada ao gerente.....	42
Figura 6.1- Modelo Entidade Relacionamento (MER) principal da empresa.....	44
Figura 6.2- Modelo Entidade Relacionamento (MER) principal da empresa.....	45
Figura 7- Modelo Entidade Relacionamento (MER) para o cargo de gerente.....	46

Lista de Tabelas

Tabela 1- Representação da relação Empregados_1.....	25
Tabela 2- Representação da relação Empregados_2.....	25
Tabela 3- Principais diferenças entre as duas edições de SGBDs.....	52

Lista de Fórmulas

Fórmula 1- Cálculo da eficácia de uma técnica de redução de banco de dados (Chan e Roddick, 2003, pg.4).....	17
--	----

Resumo

O objetivo geral do presente trabalho é a exposição das características e a criação de um ambiente adequado para a redução de banco de dados, proporcionando o seu uso em um dispositivo móvel. Primeiramente, define-se o conceito de dispositivos móveis considerado neste trabalho e a sua utilização na sociedade atual, principalmente no ambiente de trabalho. Discute-se então os possíveis meios de acesso às informações através dos dispositivos móveis e é proposta, como suposta alternativa viável, a utilização de bancos de dados reduzidos contendo apenas as informações necessárias às tarefas específicas dos usuários dentro do contexto corporativo. São analisadas as técnicas, desenvolvidas por diversos autores, para o descobrimento das informações relativas ao contexto móvel do usuário e os procedimentos existentes para redução de bancos de dados. É feita uma análise das características funcionais de um representante dos SGBDs específicos para utilização em dispositivos móveis, o SQL Server Compact. Finalmente, apresenta-se um software que cria um ambiente adequado para a redução de um banco de dados, utilizando algumas técnicas de redução apresentadas e o SGBD analisado previamente. Para a demonstração do funcionamento deste software utiliza-se um banco de dados com informações fictícias mas completas e bem formatadas, que simulam uma situação real e assim validam o seu uso. Conclui-se então, a viabilidade do uso de reduções de banco de dados para o acesso às informações, utilizando um dispositivo móvel em um ambiente corporativo.

Palavras-chaves: redução de banco de dados, dispositivos móveis, contexto móvel, SGBD

Introdução

Dispositivos móveis são aparelhos que possuem funcionalidades similares a computadores convencionais mas que apresentam tamanho reduzido, sendo portanto, como o próprio nome sugere, de fácil portabilidade, móveis. Normalmente são formados por uma tela para apresentação de informações e para a inserção de dados possuem um mini-teclado ou mesmo uma tela sensível ao toque (*touch screen*). Para exemplificar, pode-se citar alguns aparelhos encontrados facilmente como: telefones celulares, *smartphones*, *PDA*s (*Personal Digital Assistant*) e outros similares. Vale ressaltar que para o presente trabalho não se enquadram nesta categoria *laptops*, *notebooks*, *netbooks* ou qualquer outro tipo de computador que apresente recursos de armazenamento e processamento compatíveis com os computadores *desktops* convencionais.

Dispositivos móveis representam um segmento de produtos tecnológicos em constante expansão. Segundo WinterGreen Research (2005), em 2004 foram vendidos 882 milhões destes aparelhos e estima-se que em 2011 sejam vendidos 2.1 bilhões.

Todos esses dispositivos, atualmente, possuem capacidade de comunicação *wireless* (sem fio) possibilitando a conexão com internet e/ou redes locais. E apesar de serem concebidos com o intuito de serem portáteis e, portanto, com um tamanho reduzido, nos últimos anos tais aparelhos têm evoluído em ritmo acelerado em vários aspectos, tanto em desempenho, processamento, armazenagem de dados e duração de baterias.

Atualmente, o tema da mobilidade de profissionais dentro do ambiente de trabalho tem sido muito discutido. De fato, muitas empresas utilizam os dispositivos móveis para operações que requisitam o deslocamento de profissionais para locais distantes de suas instalações físicas. Muitos desses dispositivos utilizam softwares que dependem de informações armazenadas no banco de dados das empresas para o seu correto funcionamento. Portanto, para um software desse tipo, pode-se considerar que o seu valor é tão alto quanto a quantidade de informações necessária que é capaz de acessar.

Para a obtenção das informações que necessitam, estes softwares podem utilizar as tecnologias de comunicação sem fio presentes nos dispositivos móveis para acessarem os bancos de dados das empresas. Desse modo, seria necessário garantir que a conexão dos dispositivos com os computadores das empresas ocorresse independente de localidade, com uma banda de transferência de dados mínima e estável, além de não apresentar falhas de conexão.

O cenário apresentado acima, atualmente não é encontrado nem mesmo em países com redes de comunicação mais avançadas que a nacional, como mostra o trabalho de Roddick e Chan (2003, pg. 2). No cenário nacional, de acordo com o “Relatório de acompanhamento setorial para a tecnologia da informação e comunicação (2008)”, apenas 54,6 % da população brasileira está na área de cobertura da tecnologia 3G (principal tecnologia de acesso à internet móvel utilizada no Brasil).

Uma possível solução, para garantir o acesso de tais dispositivos às informações necessárias a qualquer momento, é a utilização de banco de dados reduzidos para serem instalados em dispositivos com recursos limitados.

Atualmente, existem vários programas capazes de criar e gerenciar esse tipo de banco de dados especial para a aparelhos móveis. Tais programas fazem parte de uma categoria de sistemas conhecidos como SGBD, sigla que significa Sistema Gerenciador de Banco de Dados. Segundo Date (2003) SGBDs “(...)são sistemas cujo objetivo principal é gerenciar o acesso e a correta manutenção dos dados armazenados em um banco de dados(...)” .

Não podemos instalar qualquer SGBD em um dispositivo móvel, isso devido a restrição de recursos disponíveis já mencionada. No entanto, os principais desenvolvedores desse tipo de software já oferecem SGBDs compatíveis com tais aparelhos, podemos citar como exemplo: *SQL Server Compact Edition*, produzido pela *Microsoft*; *Oracle lite 10g*, produzido pela *Oracle*; *SQL Anywhere*, produzido pela *Sybase*, apenas para citar os principais representantes do setor.

A presença de bancos de dados instalados localmente em tais dispositivos, no entanto não soluciona por completo o problema dos aplicativos que necessitam das informações das bases de dados das empresas independente do local ou momento. Segundo Roddick e Chan (2003, pg. 1), mesmo com um crescimento expressível da capacidade de armazenamento de dados apresentado pelos dispositivos móveis e que tende a continuar a crescer em ritmo acelerado, a relação entre a quantidade de dados armazenados por estes e a capacidade de dados armazenados nos bancos de dados fixos das empresas é alta e tende a ser manter estável.

Assim, parte da solução para esse problema está em selecionar a partir do total de informações presente na base de dados das empresas, uma determinada quantidade dados para a posterior persistência no banco de dados do aparelho,

sendo que essa quantidade é restrita pela capacidade de armazenamento disponível no dispositivo.

Um profissional que necessita realizar uma tarefa distante das instalações físicas da sua empresa, seja qual for essa tarefa, estará interessado na maioria das vezes em um sub conjunto definido de informações presente no total de informações do banco de dados dessa empresa. Segundo Lubinski(2000), o processo de redução de banco de dados é um processo sensível ao contexto.

1.1 Objetivos

1.1.1 Objetivo geral

Criar um ambiente adequado para permitir a redução de um banco de dados e, com isso, adequar seu uso para uma aplicação móvel.

1.1.2 Objetivos específicos

- Disponibilizar um software que tenha a capacidade de gerar um banco de dados para dispositivos móveis a partir de um banco de dados de maior capacidade.
- Mostrar e analisar as principais técnicas existentes utilizadas para a redução de bancos de dados.
- Realizar uma análise das principais características e funcionalidades do SGBD para dispositivos móveis: SQL Server Compact Edition.

1.2 Justificativa

Como já demonstrado anteriormente, os dispositivos móveis não podem confiar inteiramente nas redes atuais de comunicação *wireless* (sem fio) para a recuperação das informações que necessitam.

Soma-se a esse cenário a constante evolução dos recursos de armazenamento e processamento dos dispositivos móveis e o surgimento dos SGBDs específicos para esses aparelhos.

Assim, aplicando-se técnicas de redução de banco de dados que levam em conta o contexto de utilização do usuário, torna-se possível a persistência e a recuperação das informações necessárias ao usuário no próprio dispositivo móvel, justificando o desenvolvimento deste estudo.

1.3 Metodologia

O tema da redução de banco de dados é discutido, predominantemente, em publicações acadêmicas internacionais como monografias e dissertações, sendo estes tipos de publicações as principais fontes utilizadas para a construção do presente trabalho.

Dentre as publicações acadêmicas utilizadas, procurou-se pelas técnicas de redução de banco de dados mais comumente utilizadas e que pudessem ser utilizadas em um maior número de bancos de dados sem grandes modificações. Outra característica das técnicas procuradas foi a importância que davam ao contexto do usuário no momento de definir o conteúdo que seria gerado pela redução.

Para o estudo do SGBD SQL Server Compact Edition foram utilizadas informações oficiais fornecidas pelo seu fabricante, a Microsoft, em forma de livros eletrônicos e planilhas de funcionalidades.

Finalmente, para a construção do software foram utilizadas informações fictícias pertencentes à uma empresa fictícia que atua no ramo de autopeças. Tais informações apesar de fictícias são organizadas de modo a simular o funcionamento de uma empresa real e, portanto, demonstrar a utilidade do software para o mercado.

1.4 Visão Geral

O presente estudo é dividido em três capítulos que serão descritos a seguir:

- O primeiro capítulo apresenta e analisa algumas das principais técnicas, propostas por autores diversos, a respeito da redução de banco de dados. Sendo que as técnicas que foram utilizadas para a construção do software (apresentado no último capítulo) são especificadas com maior detalhamento.
- O segundo capítulo realiza um breve estudo sobre as principais características e funcionalidades do SGBD para dispositivos móveis da Microsoft mais recente, o SQL Server Compact Edition 3.5, que também será utilizado na construção do software do último capítulo.
- O terceiro e último capítulo apresenta um sistema simples que reduz um banco de dados para utilização em dispositivos móveis, com o objetivo de demonstrar na prática algumas das técnicas de redução de banco de dados apresentadas no capítulo um.

1.5 Problematização:

A seguir são elencadas algumas das questões que o presente trabalho pretende responder ao longo de seu desenvolvimento:

Quais as principais técnicas de redução de banco de dados presentes nos estudos referentes à este assunto?

A literatura apresentada nesse trabalho sobre redução de banco de dados apresenta as técnicas existentes, geralmente, de uma perspectiva teórica, com poucos exemplos práticos, muito menos com aplicações voltadas ao mundo corporativo. Seria possível então utilizar algumas destas teorias apresentadas e transformá-las em algo com aplicação viável dentro de um sistema empresarial? Se a resposta à pergunta anterior for afirmativa, como seria feita essa aplicação?

Quais as principais características estruturais e funcionais que diferenciam um SGBD convencional de um SGBD utilizado em dispositivos móveis, como o SQL Server Compact Edition?

1.6 Hipótese

Sendo um dos objetivos desse trabalho o desenvolvimento de um sistema que permita a redução de um banco de dados para a utilização em dispositivos móveis e dado que dentre as publicações acadêmicas e livros utilizados na bibliografia não existem exemplos práticos de métodos desse tipo, o atual trabalho, portanto, considera como hipótese a viabilidade e possibilidade de implementação de um programa que permita essa redução. Considera-se ainda como hipótese que

estas reduções ao serem realizadas possam servir para aplicação similar em outros sistemas que dêem suporte à empresas de diversos ramos de atividade.

Ao longo do desenvolvimento desse trabalho, que se dará a partir do próximo capítulo, as hipóteses apresentadas serão discutidas suas validades colocadas à prova.

Capítulo I

1. Redução de banco de dados

O processo de redução de banco de dados está ligado fortemente ao conceito de adaptação. A adaptação, por sua vez, é influenciada pelo ambiente ou contexto no qual está imerso o usuário. Nesse sentido, a escolha dos dados que irão compor o banco de dados reduzido se relaciona não apenas ao interesse que um usuário tem em um subconjunto de informações que se relacionam à tarefa que realizará, como já mencionado anteriormente, mas a um conjunto mais amplo de aspectos. Segundo Lubinski (2000), esse contexto mais amplo no qual está inserido o usuário é chamado “contexto móvel”, e é composto por:

- Pessoas (cargo do profissional dentro da empresa, tarefa específica a ser realizada)
- Recursos (capacidade de processamento, armazenamento e sistemas do dispositivo móvel, disponibilidade de redes de conexão sem-fio)
- Informação (quantidade de dados necessários, tipos de dados, meta-dados)
- Outros contextos (localidade, horário)

Uma técnica eficiente de redução de banco de dados deve levar em consideração todas as características citadas acima e desse modo selecionar as informações que melhor atenderão as necessidades do usuário dentro das restrições existentes. Segundo Roddick, Mohania & Madria (1999), a medida da

utilidade e usabilidade da informação gerada para o usuário é chamada de “capacidade de informação”.

Assim, segundo Chan e Roddick (2003, pg.4):

“(…) a medida de eficácia de uma determinada técnica de redução de banco de dados pode ser vista como a relação entre a redução física da quantidade de dados e a perda de *capacidade de informação* como resultado. Claramente, essa relação não é linear, já que alguns atributos em uma relação serão mais críticos que outros e isso irá variar com o contexto”.

Em termos matemáticos podemos ver essa relação representada pela fórmula:

$$W = \frac{\zeta_{DB}}{\zeta_{SDB}} \cdot \frac{\Upsilon_{SDB}}{\Upsilon_{DB}}$$

Fórmula 1 - Cálculo da eficácia de uma técnica de redução

Onde ζ_{SDB} e ζ_{DB} representam a capacidade de armazenamento do banco de dados reduzido e do banco de dados original, respectivamente. Υ_{SDB} e Υ_{DB} representam a “capacidade de informação” do banco de dados reduzido e do banco de dados original, respectivamente, e W representa a medida da eficácia de determinada técnica de redução.

Considerando que para um determinado banco de dados os valores de ζ_{DB} e Υ_{DB} são fixos e que ζ_{SDB} tem um limite determinado, o objetivo da técnica analisada é aumentar o Υ_{SDB} , ou seja, a “capacidade de informação” resultante.

O restante deste capítulo será organizado da seguinte maneira: primeiramente será apresentada uma técnica (protocolo) proposta por Roddick, Mohania & Madria

(1999) que tem o objetivo de formalizar um procedimento para definir mais precisamente alguns aspectos do “contexto móvel” no qual está inserido o usuário. Em seguida, será feita uma análise das principais técnicas de redução de banco de dados encontradas na literatura relacionada ao tema deste trabalho.

1.1 Protocolo baseado em prioridades

Grande parte das técnicas presentes na literatura relacionada ao assunto de redução de banco de dados, levam pouco em consideração grande parte dos aspectos que compõe o “contexto móvel”, definido anteriormente. O trabalho de Roddick, Mohania & Madria (1999), no entanto, busca retomar esses aspectos falhos em outras publicações, apresentando uma técnica que define de maneira abrangente o contexto dos interesses de um usuário.

A técnica proposta, no entanto, não representa propriamente uma técnica de redução, com uma série de procedimentos para diminuir o tamanho de um banco de dados. Esta, tem o objetivo de dar prioridades ou pesos para as informações do banco baseado em uma série de critérios estabelecidos, servindo assim de subsídio para que outras técnicas possam definir mais facilmente quais dados deverão permanecer, quais dados podem ser excluídos, quais a mudanças nos domínio dos atributos entre outros procedimentos utilizados para reduzir banco de dados.

Este protocolo divide os aspectos que devem ser analisados para a formação das prioridades ou pesos em três categorias para uma melhor organização do procedimento. As prioridades ou pesos são valores numéricos gerados a partir das próprias escolhas do usuário ou geradas automaticamente (por observação ou experimentação) e são atribuídos a cada item de cada tabela, sendo calculados

através de uma fórmula matemática. Esta fórmula, no entanto, não será analisada neste trabalho, para maiores detalhes consultar a obra de Roddick, Mohania & Madria (1999).

A primeira categoria utilizada pelo protocolo é chamada de critérios primários e é formada por procedimentos que utilizam um agente externo, normalmente o usuário interessado nos dados, para selecionar intencionalmente as informações. Sendo formada por:

- *Enumeração*: neste procedimento um agente externo, normalmente o usuário final do banco de dados ou talvez um especialista da área em questão, escolhe intencionalmente os dados que acredita serem essenciais que estejam presentes no banco de dados reduzido. Normalmente a escolha ocorre a nível de tuplas, que podem estar em várias relações. Assim, os dados selecionados através desse procedimento são, na sua totalidade, fortes candidatos a estarem presentes no banco de dados final e portanto os pesos atribuídos serão claramente elevados.
- *Contextuação*: este critério leva em consideração o contexto pessoal, citado anteriormente, e que é formado por informações relativas ao usuário, como a sua especialização profissional, as tarefas que realiza dentro da empresa e outros. Assim, a partir destes conhecimentos prévios pode-se inferir quais dados podem ser interessantes de serem incluídos e são atribuídas prioridades maiores.
- *Usos anteriores*: este critério tem o objetivo de determinar alguns dados a serem incluídos baseado no histórico de uso do banco de dados por parte do usuário. Analisa-se as queries que um determinado usuário utiliza ao longo

de um período de tempo e atribui-se valores às informações baseado na frequência de acesso. Assim, as informações mais acessadas recebem uma prioridade maior e conseqüentemente uma terão uma probabilidade maior de serem escolhidas.

Na segunda categoria chamada de *critérios secundários*, os dados selecionados a partir da primeira categoria são utilizados para que outros dados relacionados a esses sejam incluídos. Ou seja, os dados selecionados nessa categoria são todos derivados de relações com os dados da primeira. Essa categoria é formada pelos seguintes critérios:

- *Modelo ou Baseado no Esquema*: este critério utiliza a visão do modelo conceitual utilizado para representar o banco de dados, atribuindo valores aos dados relacionados ao dados selecionados na primeira categoria, utilizando o critério de quanto menor a “distância” do dado em relação ao dados já selecionado, maior é o valor da sua prioridade. A distância, portanto, é a medida do numero de relações (tabelas) que se deve percorrer, através do menor caminho, para ligar o dado previamente selecionado ao dado que se relaciona com esse.
- *Indução*: este critério utiliza técnicas de indução, normalmente através do processo de *data mining*, para identificar relações entre os atributos já escolhidos e outros dados. Assim, pode-se identificar, por exemplo, que dois itens aparecem frequentemente juntos em varias *queries* utilizadas e portanto se um destes for previamente escolhido pela primeira categoria, o outro receberá uma prioridade de valor alto.

A terceira categoria, chamada de critérios terciários, trabalha exclusivamente com os dados selecionados pela primeira e segunda categorias, aumentando ou diminuindo as prioridades desses dados. É formada pelos seguintes critérios:

- *Inferência baseada no tempo*: este critério leva em consideração que eventos que aconteceram mais recentemente devem ganhar um peso ou prioridade maior do que eventos que ocorreram há mais tempo. Neste caso, a variável tempo não se refere à data em que foi inserido determinada informação no banco de dados. Se refere à data em que foi realizada determinada ação a qual os dados estão referenciando. Como exemplo, considere um usuário que tenha acabado de inserir dados em um banco de dados e que por algum motivo se referem a acontecimentos que ocorreram há duas semanas. Dessa forma, dados que tenham sido inseridos ontem e se referem a eventos que ocorreram também ontem devem receber prioridades maiores do que os dados apenas mencionados, que acabaram de ser inseridos.
- *Inferência baseada na localidade*: este critério, similarmente ao critério acima, assume que algumas informações inseridas no banco de dados fazem referência à localidade, como por exemplo, a cidade do registro de um cliente de uma determinada cidade ou o local onde ocorreu um determinado evento. Assim, através de algum meio o usuário pode determinar a localidade que mais lhe interessa e assim os dados que se referem à informações que ocorreram em locais mais próximos receberão um peso maior.

Dentro da literatura acadêmica relacionada à redução de banco de dados, os autores dividem as técnicas de redução de várias maneiras, de acordo com critérios diversos, não existindo uma padronização. Neste texto, será utilizada a divisão

presente no trabalho de Roddick, Mohania & Madria (1999), que divide as técnicas em dois grandes grupos. No primeiro grupo, estão presentes os procedimentos que modificam somente o domínio relativo aos atributos, ou seja, os valores que podem ser inseridos nos atributos, mantendo o esquema que formam as tabelas intacto. No segundo grupo, estão presentes os procedimentos que modificam o esquema das tabelas, ou seja, excluem, adicionam, ou modificam os atributos e/ou as tuplas que formam as tabelas.

2. Técnicas de Redução de Banco de Dados

2.1 Técnicas modificadoras de domínio

2.1.1 Camadas de Redução de dados

Essa técnica foi descrita em Lubinski (2000) e se baseia na idéia, relacionada ao contexto móvel, de que o usuário está interessado em um conjunto determinado de dados presentes no banco de dados, chamado pelo autor de foco do usuário. Esse conjunto de dados, por sua vez, é delimitado por parâmetros específicos a cada situação e interesse do usuário durante o processo de redução.

Assim, à medida que as informações representadas no banco se distanciam do foco do usuário, diminuem conseqüentemente a sua importância. Desse modo, a técnica descrita propõe que se criem domínios dividindo as informações, de acordo com a distância que apresentam em relação ao foco do usuário.

Supõe-se que um certo conjunto de dados tenha sido dividido em N domínios (domínio $D(1)$ até $D(N)$). O domínio $D(1)$ representa as informações mais relevantes para o usuário e estas serão representadas no banco de dados reduzido da mesma

forma que aparecem no original. Já o último domínio, $D(N)$, representa as informações menos importantes para o usuário e assim serão representadas de uma maneira muito abstrata, como por exemplo, identificando apenas o número de tuplas que compõe este domínio.

O passo seguinte do processo é atribuir para cada domínio criado anteriormente, uma camada contendo uma série de técnicas de redução. Como regra geral, podemos definir que na camada $D(1)$ nenhum método de redução será aplicado e no domínio $D(N)$, ou seja, o último domínio, serão aplicadas as técnicas mais extremas de redução. Para os domínios intermediários ao domínio $D(1)$ e $D(N)$, não existe uma regra de técnicas de redução a serem aplicadas. Essas dependerão, de maneira mais abrangente, dos aspectos que compõe o contexto móvel no qual está inserido o usuário.

2.1.2 Abstração através da agregação

Esta técnica utiliza meta-dados referentes aos valores dos atributos para a realização das reduções. De maneira simplificada, meta-dados são dados sobre outros dados, nesse caso, são dados que trazem informações extras relacionadas aos valores de um conjunto de atributos.

Os meta-dados são criados a partir dos valores de atributos de um certo número de tuplas. Estas tuplas são substituídas por uma nova tupla (normalmente utilizando a técnica: hierarquia de conceitos, ver próximo tópico) e os meta-dados calculados são inseridos no campo que deu origem a estes meta-dados. Para um melhor entendimento deste processo, será apresentado um exemplo simples após a explicação da próxima técnica de redução.

Assim, a aplicação desta técnica a um banco de dados reduz rapidamente a quantidade de informações presente, no entanto, como faz uso de generalizações de tuplas, uma quantidade relativamente grande de informações pode ser perdida. Desse modo a utilização desta técnica deve ser bem analisada pois pode levar a uma diminuição da *capacidade de informação* do banco de dados resultante.

2.1.3 Hierarquia de Conceitos

Este procedimento cria uma hierarquia de conceitos a partir do domínio de um determinado atributo. Na prática, ocorre a substituição de alguns valores do domínio de um atributo para um valor, de nível de abstração mais elevado, que os representa de maneira mais abrangente.

Como exemplo, considere um atributo, pertencente a uma relação qualquer, chamado *esporte*. Esse atributo possui um domínio representado pelo conjunto de valores {futebol, vôlei, basquete, tênis, automobilismo}. Assim, pode-se criar uma hierarquia de conceitos com esse domínio, em que existe um domínio mais elevado de valores que generaliza alguns valores do domínio original. O novo domínio pode ser composto por {esportes individuais, esportes coletivos}. O valor do novo domínio: esporte individual representaria os valores do antigo domínio: tênis e automobilismo, analogamente o valor esportes coletivos representaria os valores: futebol, vôlei e basquete

Pode-se criar vários níveis de hierarquias de acordo com a necessidade, e após a aplicação dos novos valores aos atributos podem ocorrer duplicação de tuplas, que são eliminadas e conseqüentemente diminuem a quantidade de dados do novo banco de dados.

Esta técnica apresenta o mesmo problema presente na técnica acima citada, pois faz uso de generalizações diminui rapidamente a quantidade de informações. Portanto, a sua aplicação deve ser cuidadosamente analisada para que a capacidade de informação não se degenere na mesma proporção da quantidade de dados reduzida.

Através das tabelas 1 e 2 será apresentado um simples exemplo da aplicação das duas técnicas anteriormente definidas, para um melhor entendimento dos conceitos. Considere uma tabela, pertencente a um banco de dados qualquer, contendo dados sobre funcionários de uma universidade. Estes dados são: nome do cargo que ocupam (*cargo*) e a respectiva faixa de salário (*faixa_salario*). Assim, pode-se transformar a relação chamada *Empregados_1* na relação chamada *Empregados_2*, através das técnicas duas técnicas descritas anteriormente.

cargo	faixa_salario
Professor Titular	R\$ 15.000
Professor Associado	R\$ 6.000 – R\$ 8.000
Professor Auxiliar	R\$ 4.000
Professor Assistente	R\$ 2.500 – R\$ 3.000

Tabela 1 – representação da relação *Empregados_1*

cargo	faixa_salario
Equipe acadêmica sênior	R\$ 6.000 – R\$ 15.000
Equipe acadêmica júnior	R\$ 2.500 – R\$ 4.000

Tabela 2 – representação da relação *Empregados_2*

Utilizando a hierarquia de conceitos, os valores dos atributos cargo da relação *Empregados_1* foram reagrupados em dois novos grupos na relação *Empregados_2*. Assim, *Equipe acadêmica sênior* compreende *Professor titular* e

Professor associado, e *Equipe acadêmica júnior* compreende *Professor auxiliar* e *Professor assistente*. Desse modo, deve-se modificar também os valores dos salários para se adequarem às novas tuplas. Usa-se então a abstração por agregação, que leva em consideração os meta-dados sobre os salários da primeira tabela para criar novos valores para o atributo *faixa_salario* da relação *Empregados_2*.

2.2 Técnicas modificadoras do esquema

2.2.1 Redução vertical por projeção de tuplas

É uma das técnicas mais comumente usadas para redução de banco de dados. É realizada através da eliminação de atributos de uma dada relação. Funciona similarmente à resposta que se obtém quando se restringe os atributos que serão mostrados, a partir de uma tabela, ao se utilizar o comando *select*, da linguagem SQL.

Alguns autores utilizam o termo *fragmentação vertical* para indicar o resultado da aplicação desta técnica. A figura 1 exemplifica esta fragmentação da tabela:

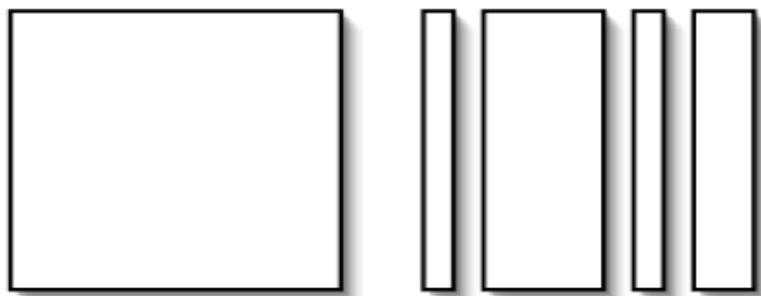


Figura 1- Representação da fragmentação vertical

A única ressalva que se faz para a aplicação deste procedimento, é o fato de que alguns atributos que participam, direta ou indiretamente, da estrutura funcional das tabelas do banco de dados não poderão ser excluídos. Como é o caso de um atributo que representa uma chave estrangeira e que serve para relacionar duas tabelas.

2.2.2 Redução horizontal por seleção de tuplas

Assim como a redução vertical por projeção de tuplas, esta técnica é uma das mais usadas para redução de banco de dados. Consiste na retenção ou eliminação de tuplas de acordo com certos critérios de seleção. Seu funcionamento é similar ao funcionamento da cláusula `where` utilizada em comandos `select` na linguagem SQL. No caso da cláusula `where`, a resolução dos termos que a seguem, para verdadeiro ou falso, indicam se uma tupla irá ser adicionada ou não.

Alguns autores utilizam o termo fragmentação horizontal para indicar o resultado da aplicação desta técnica. A figura 2 exemplifica esta fragmentação da tabela:

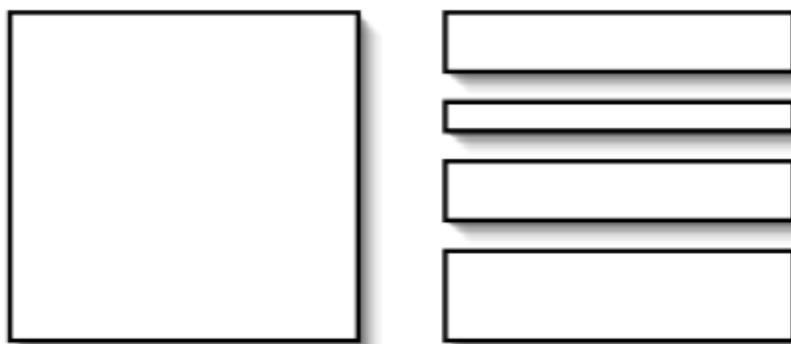


Figura 2- Representação da fragmentação horizontal

Capítulo II

1 A Família de SGBDs SQL Server

A empresa de tecnologia Microsoft oferece uma família de edições de SGBDs do tipo relacional, cada edição é feita especialmente para um público específico, contendo assim características próprias. Entre estas edições encontramos o SQL Server Standard edition e o *SQL Server Compact 3.5*. Ambos os softwares serão utilizados para na criação do software que será descrito no próximo capítulo do trabalho e foram escolhidos pois apresentam as características necessárias para a implementação das funcionalidades desejadas do sistema.

Segundo o site da Microsoft, a edição *SQL Server 2008 Standard R2*, a mais recente até o momento da escrita deste trabalho, representa um SGBD com capacidade para o compartilhamento e gerenciamento de dados em empresas e departamentos de pequeno até grande porte. Permite o gerenciamento de um número ilimitado de conexões simultâneas, possibilidade de criação de databases de até 524PB, utilização de até 64GB de memória RAM, até 4 CPUs para processamento e uma série de outras funcionalidades necessárias para cumprir as tarefas requisitadas para uma aplicação destinada ao seu público alvo, além de ser um software pago. Já a edição *SQL Server Compact 3.5 SP2* representa um SGBD, que pode ser utilizado sem o pagamento de licença, e possui capacidades mais limitadas em relação a edição previamente descrita. É voltada principalmente para a utilização em dispositivos móveis que utilizam *Windows ME* e *Windows Pocket PC* mas que pode ser utilizado também em computadores desktop, rodando qualquer tipo de plataforma Windows. O SGBD *SQL Server Compact 3.5*, é o foco deste

capítulo e a seguir serão apresentadas suas principais características, funcionalidades e arquitetura de funcionamento.

A versão analisada será a mais recente, até o momento de escrita deste trabalho, chamada: *SQL Server Compact 3.5*. Para fins de diminuição da quantidade de texto utilizado e repetição, será considerado deste momento em diante a abreviação do nome *SQL Server Compact 3.5*, para SQLSC.

Para esta versão que foi lançada em 2008, a documentação mais completa encontrada até esta data é um livro online distribuído livremente pela própria fabricante do software, intitulado: *SQL Server Compact Book Online*. Este livro além de apresentar as características do SGBD e fazer comparações com as outras versões disponíveis, contém uma série de tutoriais práticos, cobrindo todas as funcionalidades necessárias para o desenvolvimento do software proposto do próximo capítulo. Assim, este livro será a base de referência para o conteúdo que será mostrado nos próximos parágrafos desse capítulo.

2.1 Características gerais do SQLSC

As funções básicas do SQLSC, assim como em grande parte dos SGBDs do tipo relacional disponíveis atualmente, podem ser resumidas em:

- Criação, eliminação e edição de tabelas e seus dados.
- Criação, gestão e eliminação de índices.
- Visualização dos esquemas e dos tipos de dados.
- Replicação e acesso remoto a dados para sincronização.

Sendo um SGBD voltado para ser usado em dispositivos móveis, o SQLSC, apresenta características compatíveis com as restrições, já citadas anteriormente, impostas por esses aparelhos. Leva em consideração assim a limitação de armazenamento e de processamento, além da possibilidade de falta de conexão destes dispositivos, disponibilizando um sistema que condiz com estas restrições.

O SQLSC ocupa somente de 2,2 MB a 2,6 MB de espaço em disco, dependendo do processador utilizado. Isso representa um pequeno espaço de memória necessário e pode dessa forma ser utilizado em dispositivos com capacidade de armazenamento bem limitado. Todo conteúdo de um banco de dados gerado é armazenado em um arquivo simples com extensão *.sdf* e pode conter no máximo 4GB de espaço utilizado. A necessidade de apenas um arquivo para cada banco de dados torna a exportação e importação de dados simples e flexível.

Este SGBD, diferentemente das outras edições da família SQL Server, não funciona como um serviço do sistema operacional Windows, e os arquivos que o compõe devem ser instalados juntamente com a aplicação que faz uso deste.

Apesar do tamanho reduzido deste SGBD, ele apresenta funcionalidades bastante avançadas. Em linhas gerais, o SQLSC fornece até 256 conexões simultâneas a um banco dados, até 249 índices em cada tabela, suporte para transação, integridade referencial com exclusão em cascata, criptografia de 128 bits, senha de segurança, otimizador de consultas, além da possibilidade de sincronização com uma instância remota do SQL Server através de duas técnicas: *merge replication* ou RDA.

2.2 Manipulação de bancos de dados

A gestão e modificação de banco de dados pode ser feita através da utilização da API de programação disponível, ou com a linguagem de consulta SQL. A manipulação através da API é feita em conjunto com componentes de programação como o OLE DB (com a linguagem Visual C++) ou ADO.NET (com as linguagens C++ e Visual Basic). Já a utilização direta de SQL utiliza programas de administração de bancos de dados como o Management Studio ou Management Studio Express, ambas ferramentas da Microsoft.

2.3 Uso da linguagem SQL

A linguagem de consulta utilizada pelo SQLSC é um subconjunto da Transact SQL, que é uma linguagem de consulta derivada de SQL, criada pela empresa Microsoft e utilizada por completo no SQL Server Standard. Portanto, grande parte das consultas que utilizam a linguagem SQL e que são válidas para o SQL Server Standard Edition também são válidas para a edição compacta. No entanto, uma grande diferença, é a impossibilidade de criação de procedimentos, views e funções.

No final deste trabalho, em anexo I, será mostrada uma tabela apresentando os principais tipos de dados suportados pelo SQLSC.

2.4 Acesso multiusuário

É possível existir múltiplas conexões simultâneas a um mesmo banco de dados, SQLSC limita esse número de conexões a 256, mas é recomendado que não se ultrapasse 100. A gestão destas conexões é feita automaticamente pelo programa sem que haja a necessidade de interferências do desenvolvedor. O acesso multiusuário pode ocorrer de duas maneiras: a primeira se dá quando vários programas utilizam o mesmo banco de dados ao mesmo tempo. Já a segunda, representa o caso em que um mesmo programa pode ter mais de uma conexão com um mesmo banco. Este caso é comum quando um determinado software está em sendo executado pelo usuário e ao mesmo tempo está realizando a sincronização com uma instância remota de um banco de dados. Sem o acesso multiusuário estas tarefas seriam impossíveis de serem realizadas simultaneamente.

2.5 Concorrência

Com a possibilidade de acesso multiusuário, ocorre o que se chama na informática de concorrência. Concorrência diz respeito ao acesso de múltiplos usuários a mesma informação ao mesmo tempo. Com isso torna-se necessário proteger os dados acessados para que não ocorram erros e inconsistências durante as atualizações e visualizações de dados. SQLSC utiliza o mecanismo de controle de concorrência pessimístico, em que os dados que serão atualizados durante uma transação são “travados” (outro termo muito utilizado, em inglês, é *lock*), até que a transação seja encerrada.

2.6 Transações

Segundo Date 2003, transações representam unidades lógicas de trabalho que são formadas por uma série de operações de atualização em um banco de dados e devem seguir algumas regras para que sejam consideradas transações válidas.

- Garantia de atomicidade das operações, ou seja, ou todas as operações que formam a transação são executadas com sucesso ou nenhuma é executada.
- Garantia de durabilidade das modificações, ou seja, os dados alterados devem persistir depois do fim da transação.
- Transações devem ser isoladas, ou seja, as alterações realizadas por uma transação só devem visíveis à outras transações depois do término com sucesso da primeira.

SQLSC suporta transações, estas podem ser definidas de três maneiras: programaticamente usando funções da ADO.NET e OLE DB, através de programas específicos da Microsoft para administração de banco de dados: o SQL Server Compact 3.5 Query Analyzer no dispositivo móvel e o SQL Server Management Studio no ambiente desktop. Algumas restrições no entanto estão presentes, não é possível realizar transações distribuídas, ou seja, uma transação não pode ser realizada utilizando várias instâncias do SQLSC, além de não serem suportadas transações aninhadas, ou seja, uma transação que se inicia dentro de outra.

2.7 Transferência de Dados – Cliente – Servidor

Para qualquer aplicação em que o banco de dados utilizado no dispositivo móvel seja algum tipo de replicação, mesmo que não completa, de outro banco de dados presente, por exemplo, em um servidor de uma dada empresa, a capacidade de comunicação entre esses dois repositórios de dados se torna muito importante. Esta capacidade de comunicação, mesmo que possível apenas por alguns intervalos de tempo, é que permite a sincronização de atualizações, garantindo assim a consistência entre os dados presentes no cliente (SQLSC) e o servidor (SQL Server). O SQLSC torna possível esta transferência de dados de três maneiras:

- Utilizando a técnica Merge Replication, que permite a sincronização automática de atualizações entre o SQLSC e SQL Server e vice – versa, além de possuir funções prontas e configuráveis para resolver conflitos de atualização de dados.

- Utilizando as funções chamadas SyncServices para ADO.Net, que permitem a sincronização entre o SQLSC e qualquer outro banco de dados que seja compatível com a interface ADO.Net.

- Utilizando a técnica RDA (Acesso Remoto a Dados), que possibilita um modo simples de troca de dados entre SQLSC e SQL Server.

Dentre as três maneiras citadas acima, a mais completa e recomendada pela própria Microsoft quando se trata de uma aplicação com muitos dispositivos móveis que acessam os mesmos dados simultaneamente é a Merge Replication. No entanto, esta técnica é a que exige a configuração e manutenção mais complexa e

longa, podendo não ser vantajosa para sistemas mais simples em que a concorrência não seja tão presente.

A técnica RDA, por outro lado, apresenta simples configuração, mas quando se trata de um sistema em que muitos dispositivos compartilham e atualizam os mesmos dados esta técnica torna-se pouco adequada e, portanto não recomendada. Como o foco do presente trabalho não é um aprofundamento nas questões de sincronização de dados, mas esta se faz necessária para que o programa apresentado no próximo capítulo seja pertinente com seus objetivos, será utilizada a técnica RDA para a construção deste sistema.

A seguir será apresentado de maneira mais detalhada o funcionamento da técnica de transferência de dados RDA.

2.8 RDA

RDA ou Acesso Remoto a Dados representa uma das maneiras do SQLSC trocar informações com SQL Server. No entanto, para que ocorra essa comunicação é necessário um intermediador, o Microsoft Internet Information Services, ou IIS, como é mais conhecido. O IIS representa um servidor de aplicações Web e, portanto permite a comunicação entre SQLSC (cliente) e SQL Server (servidor) através de qualquer rede que utilize o protocolo HTTPS ou HTTP.

Na nomenclatura utilizada por RDA, o ato do recebimento de dados pela instância do SQLSCE provenientes do SQL Server chama-se “Pull” e a operação inversa chama-se “Push”. RDA não requer configurações no SQL Server, as operações de Push e Pull são sempre inicializadas por iniciativa da aplicação

instalada no dispositivo móvel. A operação de Pull retorna sempre um conjunto de dados que devem ser armazenadas no SLQSCE em uma nova tabela, portanto, para que dada tabela seja atualizada, devemos deletá-la e criá-la novamente com o novo conjunto de informações provenientes do SQL Server. Para cada tabela criada através do método Pull, pode-se habilitar uma função para rastreamento das atualizações realizadas nesta tabela no dispositivo móvel. Assim, cada inserção (insert), atualização (update) e deleção (delete) realizada na tabela que está sendo rastreada será salva e posteriormente, se requisitada pelo programa do dispositivo móvel, essas atualizações serão enviadas ao SQL Server, para manter a consistência dos dados, essa é a chamada operação de Push.

A operação de Push, no entanto, não realiza nenhum controle em relação as atualizações que são mandadas ao SQL Server. Se uma aplicação enviar uma atualização de uma linha da tabela e logo em seguida outra aplicação enviar uma atualização da mesma linha, o SQL Server refletirá sempre a última atualização realizada, sem nenhuma regra pré-definida.

O único controle realizado pela operação de Push é a construção de uma tabela que contém os erros ocorridos na tentativa de propagação das atualizações em direção ao SQL Server. Tais erros, também chamados de conflitos, são limitados às operações mal sucedidas de inserção, deleção e atualização de linhas de uma tabela e não compreendem o caso de atualizações consecutivas de uma mesma linha. Esta tabela com os erros ocorridos, é o único mecanismo presente em RDA para resolução de conflitos e qualquer resposta a esses conflitos deve ser implementado pelo desenvolvedor, não existe nenhuma operação previamente programada para ser escolhida.

Capítulo III

1 Introdução

O software que será descrito no atual capítulo tem o objetivo de mostrar alguns exemplos de aplicações das técnicas de redução de banco de dados analisadas no capítulo 1, demonstrando assim a utilidade destas técnicas principalmente em ambientes corporativos.

O ambiente empresarial fictício no qual este software se situa, consiste de uma revendedora de autopeças para lojas. Mais precisamente, esta revendedora negocia as peças com suas respectivas fabricantes, as armazena e depois as revende para lojas do ramo. Assim, supondo ser uma empresa de médio a grande porte com um número considerável de fornecedores e clientes, existe a necessidade de gerir muitos produtos diferentes e também muitos pedidos de compra e venda. Além disso, como se trata de uma empresa que tem seu lucro baseado em vendas, grande parte de seus funcionários são vendedores e esses devem se deslocar constantemente a procura de pedidos de venda para clientes e outros clientes em potencial.

A soma de uma base de dados com muitos registros e a necessidade constante de deslocamento de grande parte dos funcionários representa um ambiente propício para a aplicação das técnicas de redução de banco de dados utilizando dispositivos móveis, que é o tema deste trabalho

2 O software

2.1 Arquitetura:

Os componentes necessários para o funcionamento deste sistema são os mesmos necessários para a execução da técnica RDA, descrita no capítulo anterior. Assim, em um servidor na sede da empresa, contendo o banco de dados que será o alvo da redução, deverá ser instalado o SQL Server, o componente SQL Server Compact Server Agent e o servidor IIS (Internet Information Systems da Microsoft) de internet. Já no dispositivo móvel, deverá ser instalado o SQL Server Compact Edition, juntamente com o componente SQL Server Compact Client Agent e a aplicação desenvolvida. O servidor IIS, terá a função possibilitar o transporte de dados entre o dispositivo móvel e o servidor através dos protocolos utilizados na internet: HTTP ou HTTPS. Os componentes SQL Server Compact Server Agent e o SQL Server Compact Client Agent são os responsáveis por fornecer funções que habilitam a comunicação entre o SQL Server do servidor e o SQL Server Compact Edition do dispositivo. Este fluxo de utilização das camadas que compõe a arquitetura deste sistema pode ser melhor visualizado pela figura 3.

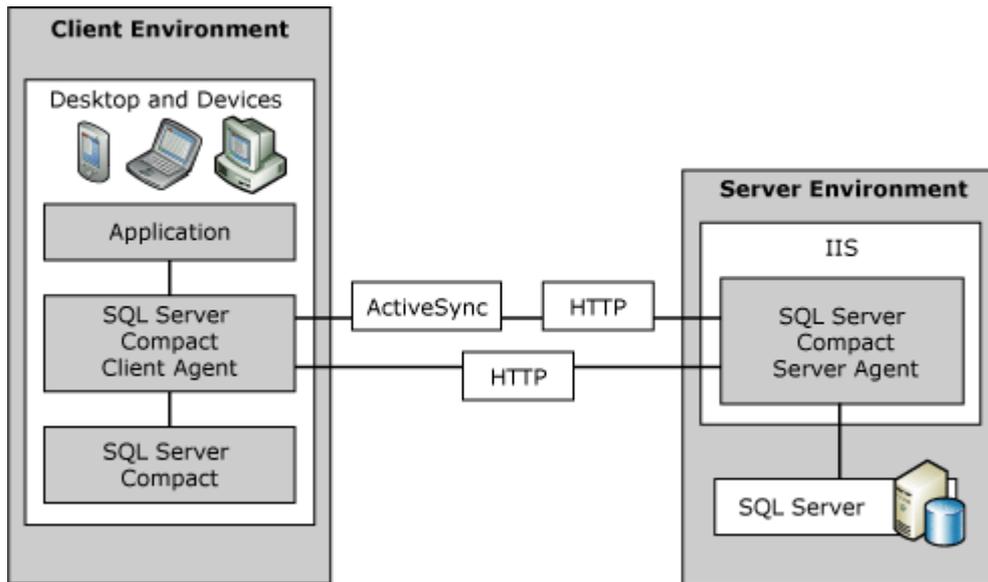


Figura 3: Arquitetura de funcionamento do sistema

No software proposto, a interação entre o usuário e o sistema é feita apenas através de um dispositivo móvel. Como este utiliza a internet para trocar informações com o banco de dados central da empresa, assim, funções como o carregamento inicial das tabelas do banco de dados, presente no dispositivo, e eventuais operações de sincronização podem ser feitas à distância, sendo necessário apenas a disponibilidade de alguma conexão com a internet. Portanto, essa característica traz certa flexibilidade para a utilização do sistema.

2.2 Visão Geral:

Cada usuário do sistema pertencerá a um dos dois tipos existentes: vendedor ou gerente. Como já citado no capítulo I, a redução de banco de dados é um processo influenciado pelo “contexto móvel” no qual o usuário está inserido, que é composto por fatores como: o cargo do usuário, as funções a este atribuídas, localidade e distância geográfica, distância temporal entre outros. Assim, um

usuário do tipo vendedor, representa um cargo com funções mais operacionais e está situado, na hierarquia dos cargos da empresa, em um nível inferior, mais próximo à base.

Como eventualmente é necessário o deslocamento do vendedor para realizar vendas junto a clientes, durante estas visitas ele estaria interessado em informações relativas à estes cliente, como: um histórico recente (de 3 a 6 meses anteriores) dos produtos comprados, representando uma maneira de achar um padrão de comportamento de compras e assim prever o conteúdo dos próximos pedidos, possibilitando um melhor atendimento; e informações relativas à localização (endereço), nomes e telefones para contato com os clientes a serem visitados.

Já o usuário do tipo gerente representa um usuário com um cargo acima dos vendedores na hierarquia dos cargos da empresa, e apresenta funções mais relacionadas à gestão de outros funcionários, contendo maiores responsabilidades com o desempenho da empresa. Assim, supondo que o gerente seja responsável por um determinado número de vendedores, ele estaria interessado em informações como: o valor vendido e a meta de vendas de cada mês em um passado recente (entre 2 e 4 meses) para cada um de seus vendedores, com um maior detalhamento dos pedidos do mês atual e dos próximos meses.

2.3 Funcionamento:

Ao ser inicializado o software no dispositivo móvel, este requisitará um login e uma senha para acessá-lo. Ao serem inseridos dados válidos, o sistema identificará automaticamente se o usuário é do tipo vendedor ou gerente. Para o usuário do tipo vendedor será mostrada a tela representada pela figura 4.



Figura 4 – Tela apresentada ao vendedor

O vendedor deverá então selecionar quais os clientes serão visitados e para qual período de tempo passado ele gostaria de receber informações sobre os produtos vendidos para estes clientes selecionados. Após o clique no botão “Carrega Dados”, o sistema requisitará ao SQL Server presente no servidor central, utilizando funções do RDA, as tabelas já reduzidas, sendo em seguida inseridas no banco de dados local ao dispositivo móvel. Estas tabelas poderão assim ser utilizadas por outros softwares, disponibilizando informações úteis ao usuário

conforme a necessidade. Esta função, como já citado anteriormente, pode ser realizada em qualquer lugar desde que exista uma conexão disponível com a internet.

Analogamente ao procedimento anterior para o usuário do tipo vendedor, o usuário do tipo gerente, ao ser identificado, receberá uma tela como a figura 5.



Figura 5 – Tela apresentada ao gerente

O gerente deverá escolher o período para o qual gostaria de receber informações sobre o valor vendido e a meta de venda mensal para cada vendedor sob sua responsabilidade. Além de escolher se quer também receber informações detalhadas dos pedidos realizados por seus vendedores no mês atual e os pedidos agendados para os meses seguintes. O processo que se segue após o clique no botão “Carrega Dados” é o mesmo que foi descrito acima para o usuário do tipo vendedor.

Na próxima seção serão mostrados os modelos entidade relacionamento dos bancos de dados envolvidos neste sistema: o banco de dados central da empresa, com todas as informações necessárias para a continuação do negócio e os bancos de dados reduzidos derivados do banco central e que são inseridos no dispositivo móvel de acordo com o tipo de usuário, vendedor ou gerente.

3.1 Detalhamento da redução dos bancos de dados

Nas figuras 6.1 e 6.2 será mostrado o modelo conceitual que representa como estão estruturados os dados presentes no banco de dados da empresa. Apesar de ser um exemplo fictício, devido a dificuldade de se encontrar exemplos de bancos de dados utilizados por empresas reais, tais dados muito provavelmente estariam presentes em um banco de dados de uma empresa deste ramo, pois, os dados utilizados representam informações indispensáveis para o funcionamento deste tipo de negócio.

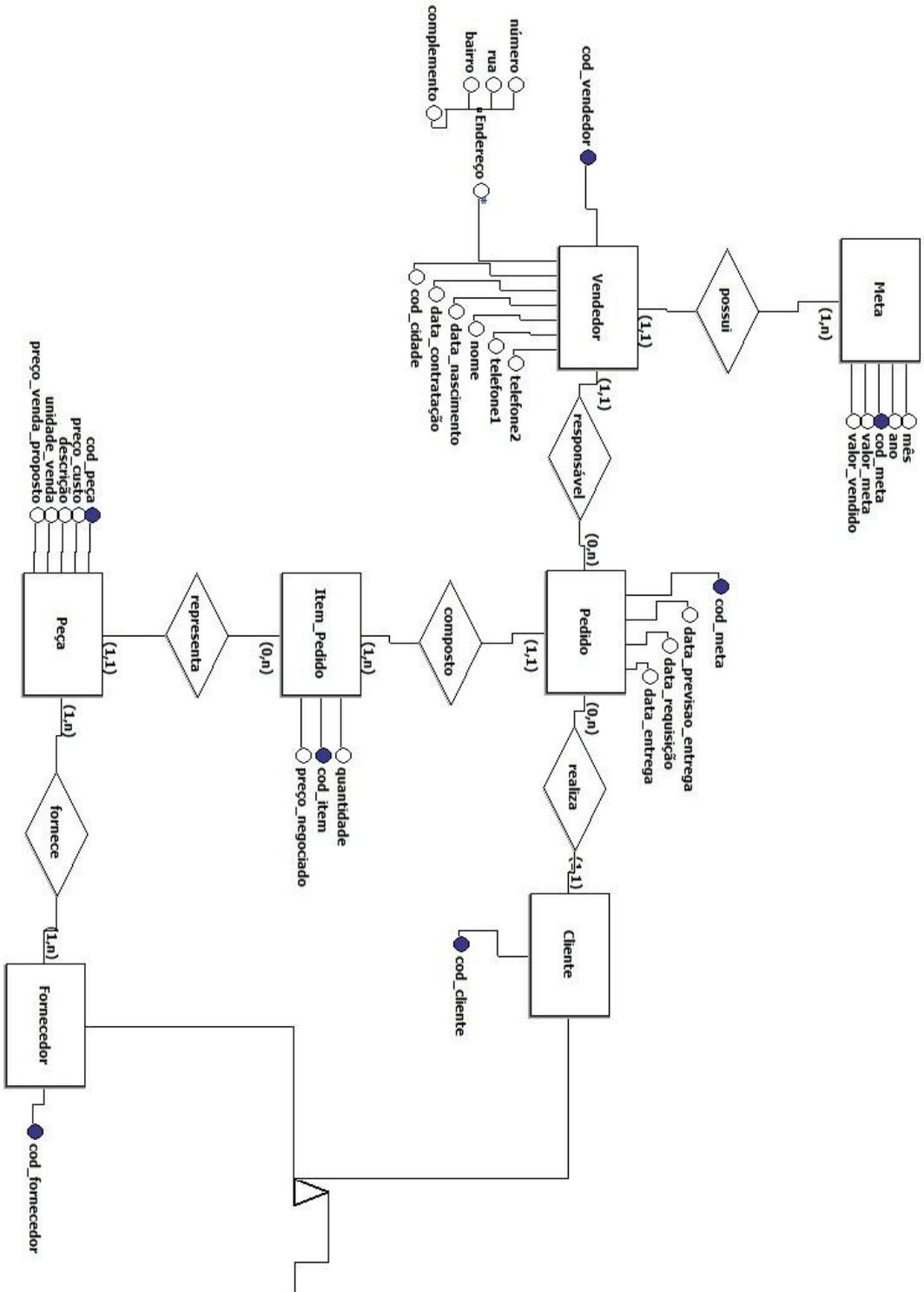


Figura 6.1 – Modelo Entidade Relacionamento (MER) principal da empresa (continua na próxima página)

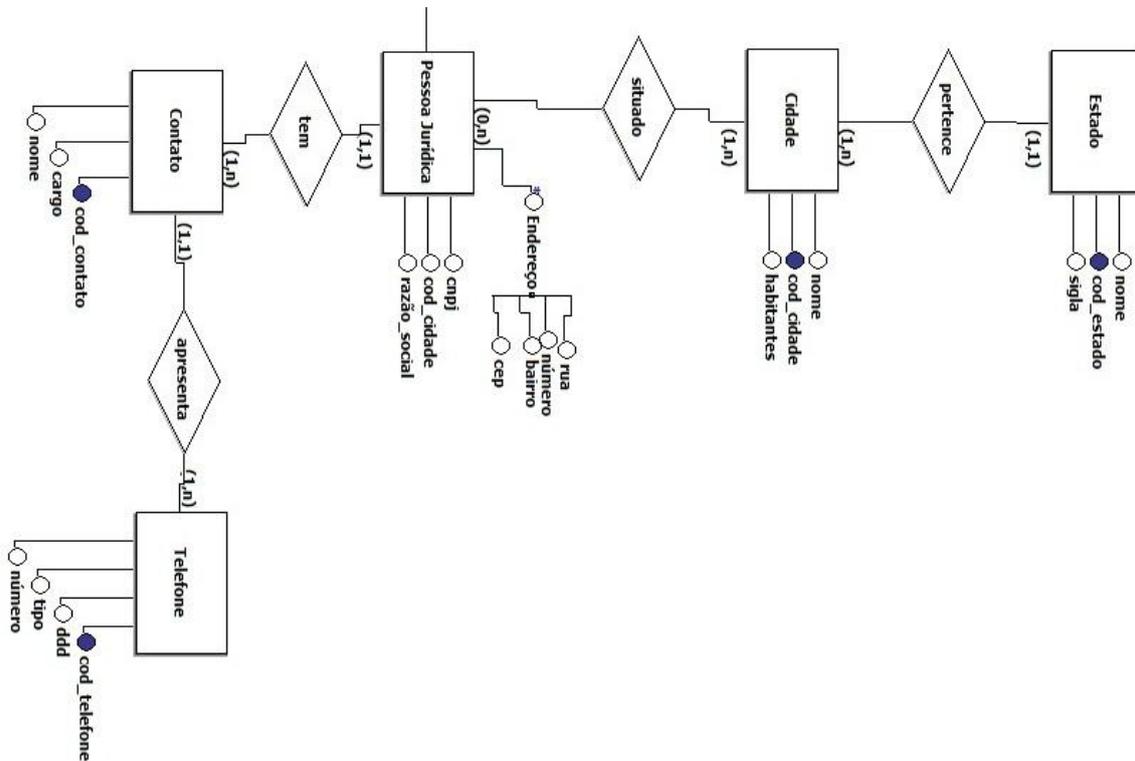


Figura 6.2 – Modelo Entidade Relacionamento (MER) principal da empresa

Obviamente, esse banco de dados, sendo o principal da empresa, contém dados detalhados sobre cada setor e cada operação que ocorre no dia – a – dia desta corporação. Não seria conveniente, nem possível, replicar este banco em um banco de dados dentro de um dispositivo móvel, assim, utilizando o “contexto móvel”, já citado no início deste capítulo, foram identificadas as informações mais importantes para os dois tipos de usuário do sistema e aplicadas as técnicas de redução de banco de dados para se chegar aos modelos conceituais que serão mostrados na figura 6 e 7.

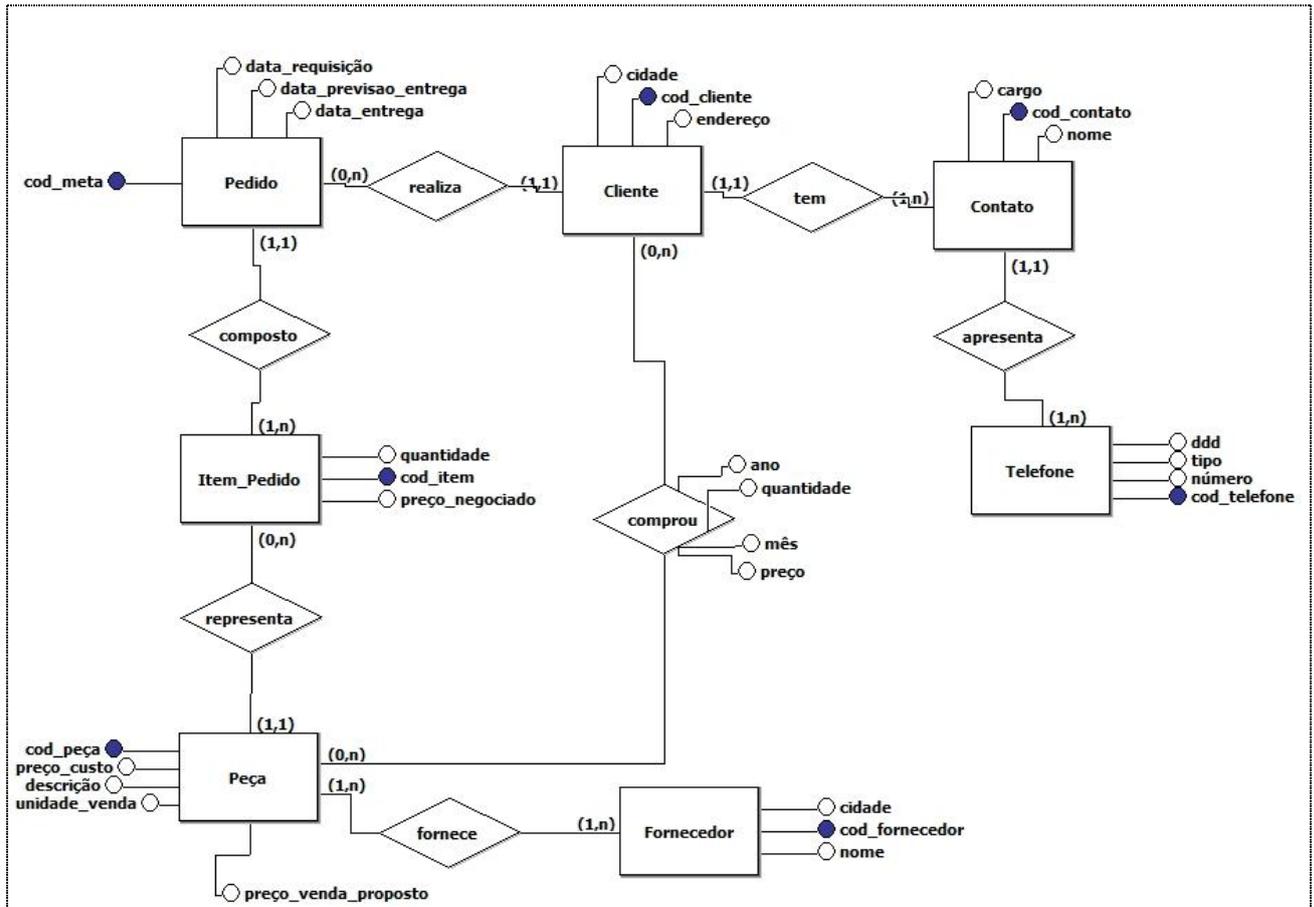


Figura 7 – Modelo Entidade Relacionamento (MER) para o cargo de vendedor

Os aspectos que compõe o “contexto móvel” do vendedor são: a tarefa de visitar clientes e potenciais clientes em busca de pedidos de venda; o nível operacional do cargo que exercem, o que os faz se interessarem por detalhes dos pedidos; os locais que se encontram os clientes que devem visitar; a importância maior apresentada por dados de pedidos recentes dos clientes visitados.

Considerando os aspectos expostos acima, foram aplicadas ao modelo conceitual do banco de dados principal as seguintes técnicas de redução para se chegar ao modelo conceitual que será usado para o dispositivo do vendedor:

- **Fragmentação Vertical:** os atributos que não agregavam nenhum valor e interesse à tarefa que guia esta redução, a visita aos clientes, foram retirados das suas respectivas tabelas. Como exemplo, o

endereço completo e o número do CNPJ dos fornecedores. Algumas tabelas, no entanto, foram excluídas pois nenhum de seus atributos tinha valor naquele contexto, a tabela do próprio vendedor, a tabela de metas e a tabela de estados, são alguns exemplos. Outras tabelas, como apresentavam apenas um atributo de interesse foram também excluídas e esse atributo foi realocado nas tabelas que referenciavam essa tabela excluída. Como exemplo, a tabela Cidade foi excluída e o atributo nome da cidade foi realocado na tabelas que faziam referência a essa tabela.

- **Fragmentação Horizontal:** a escolha dos clientes que serão visitados e a escolha do número de meses anteriores correspondente ao histórico de compras do cliente serão as informações escolhidas pelo vendedor que permitirão a fragmentação horizontal de todas as tabelas do banco, pois se uma tabela não for afetada diretamente por essa seleção, será afetada indiretamente através das referências entre as tabelas do esquema.
- **Abstração através da agregação:** o histórico de compras dos meses anteriores representado pela relação nomeada Comprou, é realizada utilizando meta-dados provenientes das tabelas Pedido e Item Pedido. Como é interessante para o vendedor apenas saber para cada produto quantas unidades foram vendidas e o preço médio por cada mês selecionado, tais informações foram extraídas das tabelas citadas acima, reorganizadas e inseridas na tabela Comprou, diminuindo significativamente a quantidade de dados utilizada nas tabelas originais.

Como se pode notar na figura do modelo conceitual do vendedor, as tabela Pedido e Item Pedido apesar de terem sido reduzidas e transformadas na tabela intitulada Comprou, ainda estão presentes no modelo. Isso ocorre pois, como a tarefa dos vendedores é realizar pedidos junto aos clientes, a definição dessas tabelas deve existir no modelo para que sejam inseridos novos pedidos. Portanto, ao carregar tais tabelas no dispositivo, estas virão vazias e servirão apenas para conter os futuros pedidos.

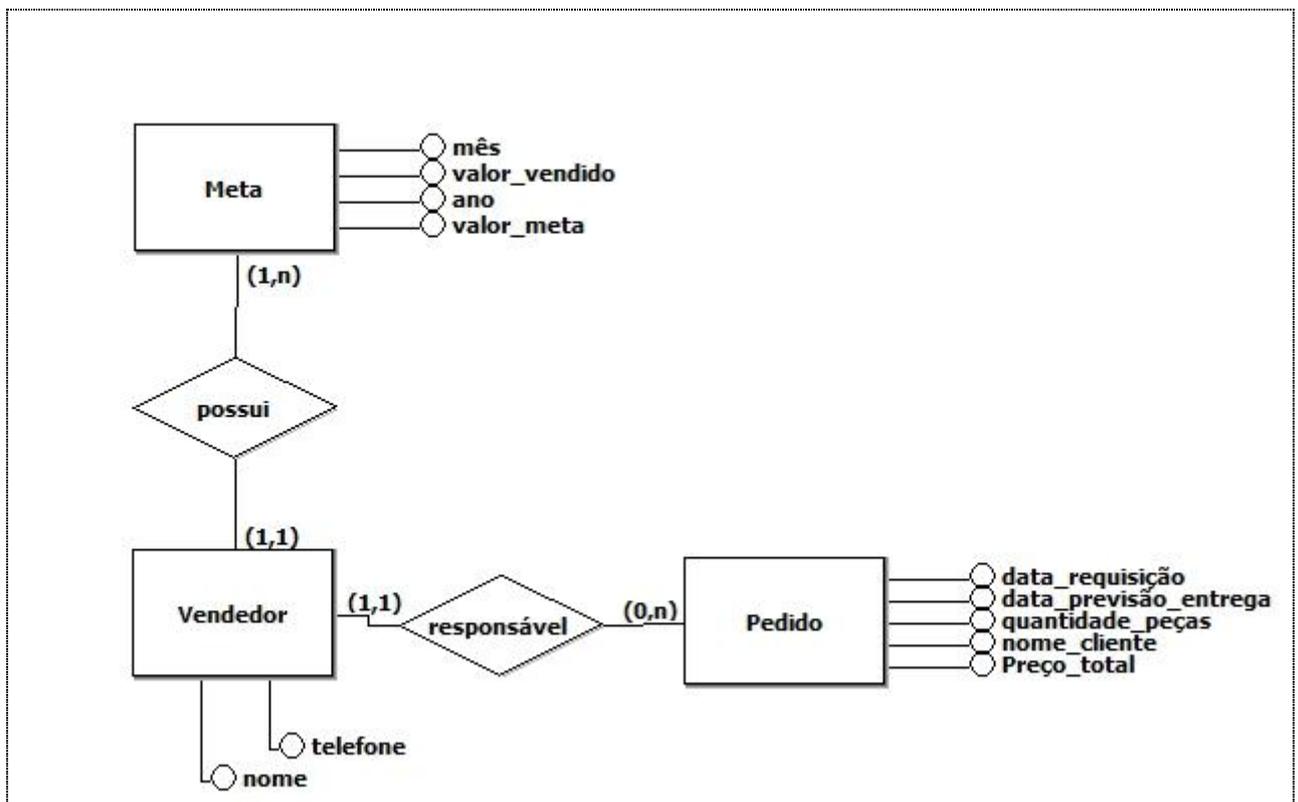


Figura 7 – Modelo Entidade Relacionamento (MER) para o cargo de gerente

Os aspectos que compõe o “contexto móvel” do gerente são: a tarefa de supervisionar e gerenciar os vendedores sob sua responsabilidade; o nível

administrativo do cargo que exercem; a importância maior apresentada por dados de pedidos recentes de seus vendedores.

Considerando os aspectos expostos acima, foram aplicadas ao modelo conceitual do banco de dados principal as seguintes técnicas de redução para se chegar ao modelo conceitual que será usado para o dispositivo do gerente:

- **Fragmentação Vertical:** os atributos que não agregavam nenhum valor e interesse à tarefa que guia esta redução, o gerenciamento dos vendedores, foram retirados das suas respectivas tabelas. Como exemplo, todos os dados da tabela Vendedor exceto seu nome e telefone de contato. Algumas tabelas, no entanto, foram excluídas pois nenhum de seus atributos tinha valor naquele contexto. Assim, foram mantidas apenas as tabelas Vendedor, Pedido e Meta, o restante foi excluído.
- **Fragmentação Horizontal:** a escolha do número de meses anteriores correspondente ao histórico de metas e vendas dos vendedores será a informação escolhida pelo gerente que permitirá a fragmentação horizontal de todas as tabelas do banco, pois se uma tabela não for afetada diretamente por essa seleção, será afetada indiretamente através das referências entre as tabelas do esquema.
- **Camadas de Redução de Dados:** essa técnica define que as informações mais próximas ao foco do usuário devem ser menos resumidas e as mais distantes mais resumidas. Assim, considerando que o foco do gerente é maior para pedidos de seus vendedores que foram realizados no mês atual da consulta e nos meses seguintes, pois a partir destes resultados podem ser tomadas decisões que

influenciem os resultados de vendas futuros. Portanto, a soma dos valores dos pedidos de venda realizados pelos vendedores nos meses passados aparecerá na tabela Meta, no atributo valor_vendido, para cada mês escolhido. No entanto, sendo o foco do gerente maior no presente e no futuro, os pedidos realizados no mês atual e os que serão entregues nos próximos meses são trazidos de maneira detalhada e não resumido como nos meses passados.

4 Ambiente de Desenvolvimento

Este tópico irá detalhar as ferramentas e linguagens utilizadas para a construção do software apresentado neste capítulo. O SGBD que serviu para simular o banco de dados principal da empresa foi o SQL Server 2005. Para a persistência dos dados no dispositivo móvel, como já bastante mencionado no capítulo anterior, foi utilizado o SQL Server Compact Edition 3.5. A linguagem de programação que permitiu a criação da programação do sistema que instalado no dispositivo móvel foi a linguagem Visual Basic. Net. E finalmente para dar suporte e facilitar a execução dos testes e o desenvolvimento geral do sistema, foi utilizado o ambiente integrado Visual Studio 2005. Todas as ferramentas e linguagens utilizadas pertencem a empresa de tecnologia Microsoft.

Considerações finais

A primeira questão proposta na problematização é referente às técnicas de redução de banco de dados existentes na literatura utilizada como bibliografia. Tal questão foi discutida ao longo do capítulo I, utilizou-se a divisão proposta por Roddick, Mohania & Madria (1999), que separa as técnicas em: técnicas modificadoras do domínio e técnicas modificadoras do esquema, para obter um efeito melhor de entendimento. Dentre as técnicas modificadoras do domínio foram citadas: camadas de redução de dados; abstração através da agregação; hierarquia de conceitos. Já entre as técnicas modificadoras do esquema, que estão entre as técnicas mais comumente utilizadas para redução de banco de dados, foram citadas: redução vertical por projeção de tuplas ou fragmentação vertical; redução horizontal por seleção de tuplas ou fragmentação horizontal.

A segunda questão presente na problematização é referente a possibilidade e viabilidade de se utilizar algumas das técnicas apresentadas no capítulo I para se criar aplicações que fossem úteis em processos dentro de certas empresas. E admitiu-se como hipótese que esse tipo de aplicação seria possível. Apesar do fato de que o banco de dados utilizado na construção da aplicação do capítulo III era fictício e a utilização do software por uma empresa real não aconteceu, pode-se considerar, pelo menos potencialmente, que a hipótese apresentada foi comprovada e a aplicação de tais técnicas no mundo corporativo é possível de ocorrer. Isso porque o tipo de negócio escolhido para se aplicar o sistema, uma revenda de autopeças, apresenta um modelo de negócios já bastante conhecido no mercado e replicado por várias empresas. Portanto, as informações contidas no banco de dados criado e os processos de negócio simulados na utilização sistema, mesmo que não representem todo o funcionamento de empresa real, muito provavelmente

estariam presentes em uma empresa deste ramo e assim a aplicação de tais técnicas pode-se considerar viável de ocorrer.

A terceira e última questão presente na problematização é referente às diferenças entre um SGBD convencional e um SGBD para dispositivos móveis. A resposta a esta questão foi discutida no capítulo II deste trabalho e utilizou para efeito de comparação o SGBD convencional, SQL Server Standard Edition 2008 e o SGBD para dispositivos móveis, o SQL Server Compact Edition 3.5 (SQLSC). Para se obter um melhor entendimento das principais diferenças encontradas nesse estudo, foi elaborada a tabela 3.

Características e Funcionalidades	SQL Server Standard Edition 2008	SQL Server Compact Edition 3.5
<i>Espaço necessário para instalação</i>	Aproximadamente 1GB, dependendo das funcionalidades e complementos	Entre 2,2 e 2,6 MB
<i>Formato</i>	Programa que funciona como serviço do sistema operacional	Arquivo com formato .sdf e uma série de arquivos .dll
<i>Tamanho máximo de armazenamento</i>	524 PB	4GB
<i>Possibilidade de criação de procedimentos, funções e visões.</i>	Possível	Não possível

Tabela 3- Principais diferenças entre as duas edições de SGBDs.

Ao longo do desenvolvimento deste trabalho foram encontradas algumas dificuldades, tanto durante a fase de dissertação, quanto durante a construção da aplicação. Após a definição do tema, a primeira dificuldade encontrada foi a pequena variedade e quantidade de publicações sobre o tema da redução de banco de dados. Neste mesmo período outra dificuldade encontrada foi que, mesmo dentro das publicações encontradas, existia uma carência de exemplos de aplicações

concretas das técnicas de redução, salvo o exemplo presente em Roddick, Mohania & Madria (1999), em que é mostrado um método para reduzir bancos de dados no contexto de um hospital com atendimentos médicos remotos. Outra dificuldade enfrentada foi a indisponibilidade de bancos de dados com dados reais de empresas, fato este que é facilmente entendido devido ao caráter sigiloso e importante destes dados para as empresas, mas que dificulta a implementação de um sistema que seja provado realmente funcional. Por último, apesar da fabricante dos SGBDs utilizados neste trabalho disponibilizar e-books que teoricamente contemplam todas as funcionalidades das últimas edições de seus softwares, os tutoriais para a implementação de alguns exemplos não eram totalmente detalhados, o que obrigou a utilização de outra referência para a implementação do sistema do último capítulo.

Levando em consideração o contexto nacional atual em que apenas 54,6 % da população brasileira está na área de cobertura da tecnologia 3G de acordo com o “Relatório de acompanhamento setorial para a tecnologia da informação e comunicação” (2008) e mesmo nas cidades em que a cobertura ocorre, na prática, o serviço é inconstante, pode-se considerar que a proposta de solução apresentada neste trabalho é viável de ser aplicada para dispositivos móveis atualmente. No entanto, ainda segundo o mesmo relatório citado anteriormente, a cobertura das redes 3G e outras tecnologias de transmissão de dados sem fio está em constante crescimento, além destes dados, de acordo com a pesquisa WinterGreen Research(2005), estima-se que em 2011 sejam vendidos 2,1 bilhões de dispositivos móveis, grande parte com possibilidade de acesso às redes 3G. Esta soma de fatores leva a crer que dentro de alguns anos o acesso à internet de praticamente qualquer local seja algo possível. Assim, provavelmente, quando este cenário for

verdadeiro, utilizar apenas os dados armazenados nos bancos de dados resumidos presente nos dispositivos móveis talvez não seja a melhor solução. Propõe-se então para trabalhos futuros, a utilização de um sistema híbrido que utilizaria os dados presentes nos bancos de dados resumidos como primeiro local de busca e se os dados necessários não forem encontrados utiliza-se o acesso à internet para recuperação dos dados direto nos bancos de dados centrais. Um sistema que funcionasse nestes moldes tiraria proveito da crescente capacidade de armazenamento dos dispositivos móveis e não congestionaria as redes de dados com o tráfego de dados redundantes e/ou fáceis de se prever a necessidade previamente, além de outras vantagens que podem ser discutidas em trabalhos futuros.

Referências Bibliográficas

- CHAN, D.; RODDICK, J. F.; **Context-Sensitive Mobile Database Summarisation**; Conferences in Research and Practice in Information Technology, Vol. 16, Australia, 2003.
- DATE, C. J; **Introdução a sistemas de banco de dados**; Rio de Janeiro, 2003.
- DUNHAM, M. H.; HALEL.; **Mobile computing and database, anything new?**, Estados Unidos, 1995.
- DHINGRA, P. ; SWANSON T; **Microsoft SQL Server 2005 Compact Edition, Estados Unidos**, Sams Publishing, 2008.
- LUBINSKI, A.; **Small database answers for small mobile resources**, Universidade de Rostock, Alemanha, 2000.
- Madria, S. K., Mohania, M. K. & Roddick, J. F.; **A query processing model for mobile computing using concept hierarchies and summary databases**, em 'Foundations of Database Organisation', Singapore, 1998
- Madria, S. K., Mohania, M. K. & Roddick, J. F.; **Methods and interpretation of database summarisation**, em 'Database and Expert Systems Applications', Australia, 1999.
- MICROSOFT; **Microsoft SQL Server 3.5 Compact Books Online**, Estados Unidos, 2008.
- WINTERGREEN RESEARCH; **Worldwide Wireless Handset Marke Opportunities, Strategies, and Forecasts, 2005 to 2011**, Estados Unidos, 2005.
- Rauen C. V;Hiratuka C.; **Relatório de acompanhamento setorial para a tecnologia da informação e comunicação**, 2008; In:
<http://www.abdi.com.br/?q=system/files/tic+dez.+2008.pdf>

Apêndice A - Principais tipos de dados presentes no SQLSC

Tipos de Dados	Descrição										
integer	Número inteiro varia de -2^{31} ($-2,147,483,648$) até $2^{31}-1$ ($2,147,483,647$). Tamanho de armazenamento 4 bytes										
bit	Inteiro com valor de 1 ou 0. Tamanho de armazenamento 1 bit.										
numeric (p, s)	Número com precisão fixa com escala numérica, varia de -10^{38+1} a $10^{38}-1$. A variável p especifica a precisão e pode variar de 1 a 38. A variável s especifica a escala e pode variar de 0 a p. Tamanho de armazenamento de 19 bytes.										
float	Número real (ponto flutuante) variando de $-1.79E+308$ a $1.79E+308$ Tamanho de armazenamento 8 bytes.										
datetime	Data e hora. Tamanho de armazenamento 8 bytes. Armazenado como dois inteiros com 4 bytes cada. Os primeiros 4 bytes armazenam o número de dias a partir de 1 de janeiro de 1900 e o restante dos bytes armazenam o número de milissegundos depois da meia noite. <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Formato</th> <th style="text-align: left;">Exemplo</th> </tr> </thead> <tbody> <tr> <td><i>yyyy/mm/dd hh:mm:ss</i></td> <td>1947/08/15 03:33:20</td> </tr> <tr> <td><i>mm/dd/yyyy hh:mm:ss</i></td> <td>04/15/1947 03:33:20</td> </tr> <tr> <td><i>dd mmm yyyy hh:mm:ss</i></td> <td>15 Jan 1947 03:33:20</td> </tr> <tr> <td><i>dd mmmm yyyy h:mm:ss</i></td> <td>15 January 1947 03:33:20</td> </tr> </tbody> </table>	Formato	Exemplo	<i>yyyy/mm/dd hh:mm:ss</i>	1947/08/15 03:33:20	<i>mm/dd/yyyy hh:mm:ss</i>	04/15/1947 03:33:20	<i>dd mmm yyyy hh:mm:ss</i>	15 Jan 1947 03:33:20	<i>dd mmmm yyyy h:mm:ss</i>	15 January 1947 03:33:20
Formato	Exemplo										
<i>yyyy/mm/dd hh:mm:ss</i>	1947/08/15 03:33:20										
<i>mm/dd/yyyy hh:mm:ss</i>	04/15/1947 03:33:20										
<i>dd mmm yyyy hh:mm:ss</i>	15 Jan 1947 03:33:20										
<i>dd mmmm yyyy h:mm:ss</i>	15 January 1947 03:33:20										
national character varying(n) Sinônimo:nvarchar(n)	Caracteres Unicode com tamanho que pode variável, de 1 a 4000 caracteres. Tamanho de armazenamento corresponde a duas vezes o número de caracteres armazenados, em bytes.										
binary(n)	Dados binários com tamanho, com máximo de armazenamento de 8000 bytes. O tamanho necessário para armazenamento corresponde ao número n em bytes.										

