

1

Introdução

Nos últimos anos, a evolução das tecnologias sem fio tem impulsionado e motivado o desenvolvimento de aplicações capazes de explorar de forma vantajosa as características da mobilidade do usuário. Além das características primárias que fazem parte da lógica de negócio, essas aplicações levam em conta o contexto corrente do usuário, do dispositivo ou da rede sem fio para prover serviços e conteúdo mais adequados ao usuário final. Como exemplo de uso de contexto, essas aplicações poderiam considerar a localização do usuário ou a vazão do enlace da rede sem fio para personalizar o seu comportamento, como, por exemplo, enviar uma oferta de um produto de uma loja próxima ao usuário ou implementar uma compressão dos dados transmitidos para amenizar algum problema da rede sem fio.

Serviços e middleware de provisão de contexto [5, 6, 7, 8] têm sido extensamente pesquisados e propostos com o objetivo de auxiliar o desenvolvimento de aplicações sensíveis ao contexto, especialmente para redes móveis. Diferentes autores [9, 8, 10, 11] propuseram duas categorias principais de informações de contexto: fatores físico e humano. A primeira foi dividida dentro de três sub-grupos: contexto do ambiente (por exemplo, luz, temperatura, hora, etc.), infra-estrutura computacional (por exemplo, recursos do dispositivo e da comunicação, etc.), e localização (por exemplo, coordenadas relativas e absolutas, etc.). Fatores humanos também foram classificados em sub-grupos: informação pessoal (por exemplo, preferências, interesses, habilidades, etc.), relacionamento social (grupo, família), e atividade (por exemplo, ocupado, descansando, trabalhando, etc.).

De fato, o acesso às informações contextuais dos dispositivos e usuários abre uma grande variedade de novas possibilidades para implementar aplicações distribuídas para usuários móveis. Entretanto, nas redes móveis, um fator de complexidade adicional diz respeito à heterogeneidade dos dispositivos usados e das tecnologias e protocolos de comunicação usados nas redes nas quais eles estão conectados. As aplicações precisam tratar dispositivos clientes com características distintas e em contextos variados. Desta forma, é interessante que o conteúdo transmitido pela aplicação, por exemplo, uma página HTML,

uma mensagem ou uma imagem, seja personalizado de acordo com as condições e estado atual dos recursos do dispositivo. Assim, há a necessidade de efetuar adaptações às condições dinâmicas dos dispositivos cliente, como por exemplo, a formatação/codificação adequada ao tamanho da tela ou ao número de cores disponíveis.

Na abordagem em que o conteúdo é pré-adaptado (adaptação estática de conteúdo) [12, 13], são mantidas várias cópias do mesmo conteúdo em diferentes formatos e precisão para atender a cada tipo de cliente, e o servidor apenas seleciona a versão mais adequada para cada cliente, de acordo com o contexto do dispositivo. Outra alternativa é adaptação dinâmica de conteúdo [14, 15, 16, 17]. Nesta abordagem há apenas uma versão do conteúdo (a versão original) e, no momento do envio deste para o cliente, uma adaptação é efetuada de acordo com o contexto corrente do cliente. Esta solução não requer a autoria de novas versões, sendo mais flexível para atender um conjunto mais amplo e dinâmico de condições de contexto (relacionados à rede, ao dispositivo, ao usuário, etc) que necessitem de adaptações. Para a adaptação de conteúdo, várias técnicas podem ser utilizadas, dentre elas destilação, refinamento, sumarização [18], filtragem [19, 20] e *transcoding* [21].

Os sistemas sensíveis a contexto, em geral, utilizam a técnica de adaptação dinâmica de conteúdo para melhor se adequar às necessidades dos clientes móveis. Adaptações dinâmicas de conteúdo geralmente são operações custosas e demandam um alto poder de processamento, principalmente para conteúdo multimídia. Efetuar todas estas operações no servidor, para cada cliente móvel, torna-se pouco eficiente e escalável. Assim, uma abordagem comum na arquitetura de aplicações para redes sem fio é a utilização de proxies [22, 23, 24], que intermedeiam toda a comunicação entre clientes e servidores. Estes proxies são responsáveis por realizar transformações, adaptações ou funções de gerência em benefício de um ou mais clientes, tais como adaptação de conteúdo, tradução de protocolos, *buffering* de mensagens. Em geral, eles estão localizados em nós da rede fixa (na fronteira com a rede sem fio, próximo aos clientes) e permitem o desenvolvimento de “clientes leves”. Uma das principais vantagens da utilização de proxies é compensar as diferenças (por exemplo, na vazão, latência, confiabilidade) entre as redes com e sem fio, além de permitir uma integração menos custosa com sistemas existentes, já que as questões relativas ao enlace sem fio, à adaptação sensível a contexto e à gerência de mobilidade são tratadas pelos proxies, não sendo necessárias mudanças nos servidores.

Esses sistemas utilizam o modelo tradicional de comunicação requisição/resposta que suporta um estilo de interação guiado por demanda, no

qual usuários requisitam explicitamente informações de servidores específicos e aguardam sua resposta (estilo *pull*). Este modo de interação não é apropriado para muitas aplicações móveis, que se baseiam na disseminação de conteúdo, como serviços de notícias, sistemas de informação de tráfego, de boletins meteorológicos, ou de acompanhamento de ações (bolsa de valores).

No paradigma de comunicação *push*, a iniciativa de comunicação é efetuada pelas fontes de informação. Os servidores entregam automaticamente as informações atualizadas aos clientes, após o registro de interesse prévio destes com os servidores. A característica mais atrativa desse modelo de comunicação é que os clientes não precisam enviar requisições repetidamente ao servidor (como no estilo *pull*), reduzindo drasticamente o número de pedidos e consumo de largura de banda. Esse paradigma apresenta uma solução viável ao problema de disseminação de conteúdo escalável para clientes móveis. Uma aplicação de disseminação de conteúdo permite a entrega de conteúdo de diversas fontes a inúmeros clientes interessados, como por exemplo, uma aplicação de agência de notícias *on-line* que distribui conteúdo multimídia contendo textos, imagens e vídeos. O tempo de publicação e o tempo da mudança de conteúdo são eventos probabilísticos que não podem ser preditos *a priori*.

O modelo de interação publish/subscribe (Pub/Sub) reflete o estilo da interação da disseminação de conteúdo. As fontes de conteúdo são os publicadores (*publishers*) que definem e fornecem o conteúdo a um serviço de disseminação de conteúdo que, por sua vez, o distribui aos assinantes interessados (*subscribers*), destinatários do conteúdo. Os assinantes podem filtrar o conteúdo a ser notificado, definindo no registro de interesse (*subscription*) o tipo e as propriedades do conteúdo de seu interesse. Além disso, devido às suas características inerentes, tais como, assincronismo, fraco acoplamento entre as partes e a extensibilidade do sistema, a comunicação publish/subscribe é considerada bastante vantajosa para ambientes móveis [25, 26], onde os dispositivos podem estar freqüentemente indisponíveis ou desconectados. Por último, a natureza baseada em eventos do modelo de interação Pub/Sub oferece meios para projetar aplicações que sejam cientes dos estados dos clientes e adaptáveis às mudanças no contexto destes.

1.1

Cenário de aplicação

O foco desta tese está em aplicações de disseminação de conteúdo para clientes móveis, tais como, agência de notícias (com conteúdo multimídia) e boletins de tráfego aéreo. As principais características deste tipo de aplicação são:

(i) conteúdo difundido para muitos clientes ao mesmo tempo; (ii) conteúdo dinâmico, que sofre freqüentes alterações e atualizações; (iii) clientes com demandas específicas de adaptação de conteúdo, de acordo com os recursos do dispositivo e rede; e (iv) comunicação publish/subscribe. Além disso, diferentemente de outras aplicações, as aplicações aqui abordadas utilizam a comunicação publish/subscribe para a **disseminação do conteúdo** de interesse do cliente, e não apenas para distribuição de mensagens de controle e de eventos de mudança de contexto (do dispositivo e rede) ou perfis dos clientes.

Entretanto, a grande maioria dos sistemas com suporte à adaptação de conteúdo para clientes móveis apenas pode ser usada por aplicações com comunicação síncrona (requisição/resposta) [18, 27, 24]. Por outro lado, alguns poucos sistemas para aplicações de disseminação de conteúdo, baseadas em comunicação assíncrona (do tipo Pub/Sub), provêm suporte a clientes móveis [28, 29]. Mas estes sistemas apenas tratam questões de conectividade e mobilidade dos clientes, focando em técnicas para roteamento eficiente e redução de perdas de conteúdo. Estes sistemas não abordam a questão de adaptação de conteúdo sensível às mudanças de contexto dos clientes. Para uma maior aceitação de serviços de disseminação de conteúdo a clientes móveis, é importante que o serviço seja flexível e capaz de entregar conteúdo customizado às necessidades (contexto corrente) dos clientes, mas sem provocar nenhum incômodo (sem intervenção) a eles.

1.2 Objetivos

O objetivo principal desta tese é prover o suporte ao desenvolvimento de aplicações de disseminação de conteúdo para clientes móveis, com adaptação de conteúdo sensível a contexto e utilizando comunicação publish/subscribe.

Um desafio em relação à comunicação publish/subscribe para aplicações cujo conteúdo precisa ser adaptado de acordo com o contexto dos clientes é o fato de que a comunicação é do tipo 1-para-muitos, ou seja, uma mesma mensagem é enviada ao mesmo tempo para um conjunto de clientes (assinantes do conteúdo). Entretanto, a adaptação do conteúdo depende do contexto dinâmico de cada cliente e, portanto, deve ser individualizada.

Uma alternativa é a utilização de proxies (posicionados próximos aos clientes) responsáveis por efetuar a interceptação das mensagens, e então adaptá-las conforme se faça necessário. Como as adaptações são específicas para cada cliente, o conteúdo precisaria ser em princípio replicado, para então sofrer as transformações necessárias e ser encaminhado ao cliente. Neste caso, para uma maior eficiência, é preciso também dispor de mecanismos para

otimizar a execução das adaptações, reduzindo a replicação de mensagens e o número de adaptações realizadas. A solução aqui proposta baseia-se na idéia de criação de grupos de clientes com contextos semelhantes, para os quais a replicação e adaptação do conteúdo da mensagem não é realizada individualmente para cada cliente, mas sim para cada grupo. Dessa forma, nosso objetivo é desenvolver um sistema capaz de atender a esta necessidade (adaptação de conteúdo dependente de contexto para comunicação Pub/Sub), tendo como foco questões de eficiência e escalabilidade das adaptações.

1.3 Contribuições

As contribuições principais desta tese estão relacionadas à disseminação de conteúdo para clientes móveis e compreendem: *i*) projeto de uma arquitetura de referência, *ii*) algoritmos para otimização da execução de adaptações de conteúdo, e *iii*) um framework para facilitar o desenvolvimento de proxies para aplicações com adaptação de textos e imagens para dispositivos móveis sensíveis ao contexto. A seguir, as contribuições são discutidas em mais profundidade.

Arquitetura de referência. Foi projetada uma arquitetura de referência para aplicações móveis de disseminação de conteúdo que permite a entrega customizada de conteúdo aos clientes móveis. Esta arquitetura é composta por elementos que encapsulam e tratam separadamente cada uma das diferentes funcionalidades necessárias ao suporte desse tipo de aplicação. A comunicação é baseada em um middleware publish/subscribe, que realiza a interação entre os usuários e provedores de conteúdo. O contexto dos clientes é provido por um serviço de contexto. A camada de adaptação, o núcleo da arquitetura, é responsável por selecionar e executar as adaptações adequadas a cada cliente de acordo com seu contexto. Por fim, a aplicação é responsável pela definição das regras de adaptação, i.e, regras que relacionam as situações de contextos e as adaptações adequadas. A separação das partes de comunicação e contexto da parte de adaptação permite, pelo menos teoricamente, uma integração com diversos middlewares de Pub/Sub e de contexto. Além disso, a estrutura modular da camada de adaptação permite instanciar apenas as partes necessárias para a aplicação de disseminação.

Algoritmos para agrupamento de clientes baseado em contexto. A execução das adaptações de conteúdo representa o ponto de maior consumo de recursos computacionais, podendo se tornar um “gargalo” do sistema. Para lidar com a questão de eficiência e escalabilidade do sistema, foram propostos

algoritmos que otimizam a execução de adaptações de conteúdo para o conjunto de clientes de uma difusão. Esses algoritmos utilizam as informações de contexto dos clientes para verificar os clientes com contextos similares e separá-los em grupos, para os quais serão realizadas as adaptações de conteúdo. Tais algoritmos apresentam ganho de desempenho a medida que o número de clientes é maior, e as situações de contexto são compartilhadas. Os algoritmos estão aptos a serem executados em sistemas multi-thread.

ProxyFramework. Foi desenvolvido também um framework, **ProxyFramework**, para facilitar a implementação do núcleo da arquitetura proposta, na forma de um proxy. Este framework implementa uma parte significativa da camada de adaptação, realizando funções de coleta de contexto dos clientes, determinação das adaptações adequadas e execução das adaptações selecionadas, além de permitir o armazenamento temporário de mensagens devido a desconexões. A tomada de decisão de como e quando adaptar, bem como a implementação dos adaptadores, são tratadas como pontos de extensão do framework. Além da comunicação Pub/Sub, o framework pode também ser usado para comunicação requisição/resposta.

1.4

Estrutura da tese

A tese está organizada em sete capítulos. Neste capítulo, foram apresentados a motivação, o objetivo e as contribuições desta tese.

O Capítulo 2 explica os conceitos básicos sobre adaptação de conteúdo, informações de contexto e sistemas de aquisição de modelagem dessas informações, e características da comunicação publish/subscribe. Esse capítulo também apresenta a arquitetura **MoCA**, desenvolvida no grupo de pesquisa do LAC da PUC-Rio, que foi usada como provedora de informações de contexto de dispositivos móveis.

O Capítulo 3 apresenta uma visão geral dos trabalhos correlatos sobre adaptação de conteúdo sensível a contexto em ambientes móveis e sistemas relacionados a disseminação de conteúdo utilizando comunicação publish/subscribe. São analisadas e comparadas as características das soluções adotadas nos diferentes sistemas para tratar de questões sobre obtenção do contexto dos clientes, definição de políticas de adaptação e execução das adaptações.

O Capítulo 4 descreve a arquitetura de referência para desenvolvimento de aplicações de disseminação de conteúdo com adaptação sensível ao contexto para clientes móveis. A arquitetura enfatiza a separação de responsabilidades em camadas e gerentes, incluindo elementos para obtenção de contexto,

adaptação de conteúdo, definição das regras de adaptação, comunicação publish/subscribe, entre outros. Além disso, são apresentados algoritmos para otimizar a execução das adaptações de conteúdo para grandes quantidades de clientes, melhorando assim a escalabilidade das adaptações. Esses algoritmos identificam clientes com contextos comuns, e permitem a customização eficiente de conteúdo. A análise de complexidade desses algoritmos também é apresentada e discutida.

O Capítulo 5 apresenta um framework, chamado *ProxyFramework*, para o desenvolvimento da camada de adaptação de conteúdo, proposta na arquitetura de referência (Capítulo 4). São discutidas as soluções adotadas para cada componente concreto ou extensível do framework. As partes concretas realizam, entre outras coisas, a coleta de contexto dos clientes e a execução das adaptações selecionadas, enquanto que a tomada de decisão de como e quando adaptar é tratada como um ponto de extensão do framework.

O Capítulo 6 contém uma avaliação de desempenho do *ProxyFramework*, implementado utilizando os algoritmos propostos no Capítulo 4. São descritos os cenários dos testes realizados, que visam avaliar a escalabilidade e o tempo de resposta do sistema, com o aumento do número de clientes e de adaptações. Neste capítulo, também são apresentados os resultados de uma enquête (através de um questionário), realizada com os desenvolvedores de protótipos de aplicações de disseminação de notícias (tipo RSS), sobre a usabilidade, isto é, grau de dificuldade e os benefícios do uso do framework para o desenvolvimento de aplicações adaptativas.

Finalmente, o Capítulo 7 apresenta as considerações finais da tese, salientando as contribuições do trabalho e as pesquisas futuras.