

Sumarização baseada em contexto ambiental como fator de adaptação de objetos multimídia em dispositivos de computação móvel

Luís Fernando Fortes Garcia¹, Prof. Dr. José Valdeni de Lima¹

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{lgarcia, valdeni}@inf.ufrgs.br

***Abstract.** This article proposes an adaptive system architecture capable of summarizing multimedia objects, particularly textual objects, for Internet mobile devices. The proposed architecture provides text adaptation based on summarization techniques considering time and location aspects. This solution provides the right information to the right user based on its location.*

1. Introdução

A crescente migração da computação *desktop-based* para a computação móvel, disponibilizada por equipamentos como *palmtops*, *handwelds* e celulares com capacidade de integração com a Internet, representa uma importante mudança de paradigma. As informações, inicialmente desenvolvidas para uma plataforma computacional tradicional devem agora ser adaptadas para as características e as diferentes possibilidades contextuais geradas pela utilização de dispositivos móveis, considerando (i) diferentes localizações geográficas – oriundas da intrínseca mobilidade dos dispositivos e horários – decorrentes da possibilidade de utilização ubíqua, (ii) características técnicas como capacidades de processamento, armazenamento e transmissão de dados e (iii) possibilidades de interação homem-máquina através de dispositivos de entrada de dados diferenciados – baseados em escrita manual, reconhecimento de voz ou pequenos teclados e saída, representados normalmente por dispositivos com dimensões e capacidades reduzidas.

Consolida-se, então, a necessidade de desenvolvimento de modelos de adaptação de informações e recursos computacionais que possibilitem melhorias significativas na interação homem-informação no âmbito da computação móvel.

Para o tratamento da sobrecarga cognitiva e desorientação, gerada pelo excesso de informação, podemos apresentar como fator de adaptação a utilização de sumarização automática dos objetos envolvidos, como textos, imagens, vídeos e sons. A geração de versões sumarizadas dos documentos originais, além de permitir um acesso mais universal em dispositivos limitados, pode, também, contribuir para a redução da sobrecarga das informações pela objetividade da informação.

Pode-se estabelecer uma relação entre os conceitos de adaptação, sumarização e contextualização aplicados sobre a plataforma computacional móvel. Através da sumarização dos objetos, com informações e características providas pelo levantamento e aplicação do contexto, especialmente em fatores ambientais – localização física e

temporal - podemos aprimorar o processo de geração de sumários adequados ao contexto global do usuário de serviços de computação móvel.

O objetivo deste trabalho contempla a definição de um modelo de sumarização baseado em contexto ambiental –utilizando fatores de localização geográfica e temporal - que possa ser utilizado como fator de adaptação em dispositivos de computação móvel com capacidades de comunicação sem fio.

2. Contextualização

Um mesmo objeto textual pode conter diferentes assuntos que podem interessar ou não a diferentes perfis de usuários. Este interesse pode ser diretamente proporcional ao local e ao horário referentes ao usuário. A filtragem destes assuntos poderia ser uma técnica coerente na construção de algoritmos de sumarização. A filtragem de diferentes assuntos, levando em conta os aspectos do local e horário de acesso à informação, não são explorados nos trabalhos da bibliografia da área. O trabalho aqui apresentado pretende a adaptação de textos através da sumarização dos mesmos, levando em conta local e horário de acesso às informações para dispositivos móveis conectáveis à Internet.

Assim, é perfeitamente possível a identificação de cenários em função de local e horário capazes de permitir a filtragem de informações a serem providas. Por exemplo, quando um usuário se encontra fisicamente localizado em um pregão de bolsa de valores é natural receber sumários tratando de assuntos do mercado financeiro tais como cotações de moedas, taxas de juros e valorização de ações. Mas, quando o mesmo usuário desloca-se para um *shopping center* o novo contexto ambiental altera o foco da sumarização para assuntos relacionados a produtos, liquidações e filmes em cartaz.

3. Sumarização baseada em contexto ambiental como fator de adaptação de objetos multimídia em dispositivos de computação móvel

A arquitetura contempla a adaptação de um método de sumarização automática extrativo – especificamente baseado em palavras-chave – através da inclusão da característica de obtenção dinâmica da relevância de palavras-chaves provenientes do contexto ambiental do usuário e de seu perfil.

A abordagem proposta pode ser aplicada a diversos tipos de objetos multimídia – textos, imagens, fotos, vídeos e sons. A integração destes na arquitetura depende primordialmente da (i) utilização de técnicas de sumarização correspondentes aliada a (ii) implementação do mapeamento dos significados semânticos dos objetos multimídia com a relevância das palavras-chave associadas.

Outra característica da arquitetura apresentada dá-se pela possibilidade de utilização de diversas técnicas de sumarização automática de textos; A escolha deu-se por técnicas baseadas em palavras-chave visto que, na abordagem proposta não é imprescindível o aprofundamento no tratamento semântico do texto, devido ao forte enfoque em relevância de palavras-chaves.

Neste primeiro momento optou-se pelo desenvolvimento apenas relacionado a objetos textuais pelo fato de existirem várias iniciativas nesta direção [Mani 2001], [Martins e Pardo 2001] e [Gonzalez e Lima 2001]. As técnicas utilizadas de sumarização não são plenamente eficazes, resultando muitas vezes em sumários incompreensíveis.

desenvolvidas e por questões de compatibilidade desta mídia com todas as plataformas computacionais móveis.

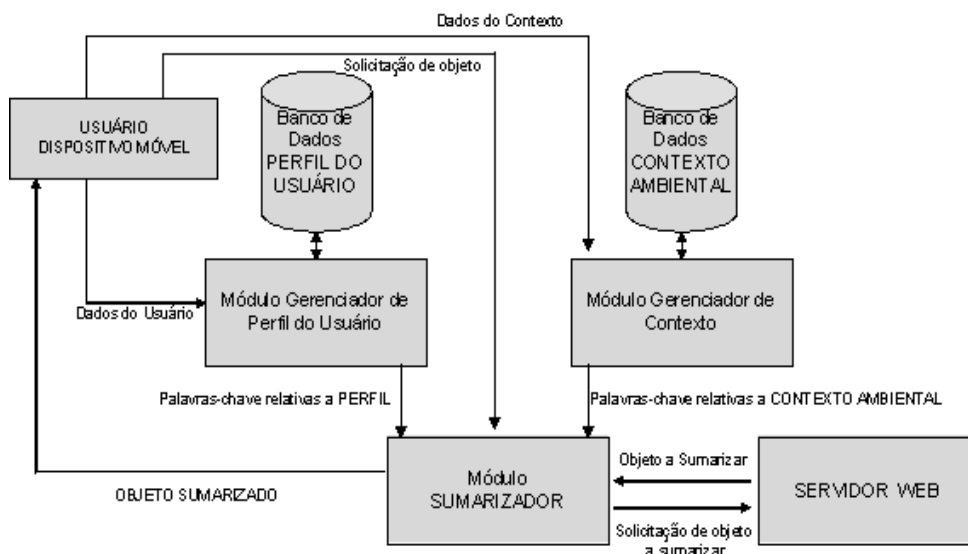


Figura 1 – Arquitetura proposta

A arquitetura, ilustrada na figura 1, contempla três módulos principais: (i) Módulo Gerenciador do Perfil do Usuário; (ii) Módulo Gerenciador de Contexto e (iii) Módulo Sumarizador.

O Módulo Gerenciador do Perfil do Usuário é responsável pela aquisição – através de monitoração da interação usuário-sistema ou fornecimento explícito, usando mecanismos como formulários, do perfil do usuário. Entende-se por perfil informações referentes desde a identificação do usuário e suas preferências a assuntos e palavras-chaves relevantes por assunto. O refinamento do perfil do usuário pode ser considerado dinâmico e é constantemente atualizado pelo Módulo Gerenciador do Perfil do Usuário.

O Módulo Gerenciador de Contexto, por sua vez, é responsável pela aquisição, conversão e tratamento dos dados ambientais correntes do usuário. Os dados de contexto ambiental tratado na arquitetura incluem localização geográfica e temporalidade.

O Módulo Sumarizador implementa um algoritmo de sumarização automática de textos adaptado para trabalhar com relevância de palavras-chaves fornecidas pelos Módulos Gerenciadores de Perfil do Usuário e Gerenciador de Contexto.

O modelo proposto está sendo implementado em um sistema que integra em sua parte cliente equipamentos de Internet móvel baseado em tecnologia GSM e/ou emuladores de celulares GSM e na parte servidor um servidor Web Apache rodando preferencialmente em GNU/Linux. Para o desenvolvimento do *gateway* de sumarização adaptativa utiliza-se de linguagem Java acessando um banco de dados relacional

MySQL ou PostgreSQL. Podemos ressaltar o enfoque em utilização, na medida do possível, de ferramentas baseadas em Software Livre.

Como trabalhos correlatos – parcialmente relacionados a este modelo - podemos destacar o trabalho do Buyukkokten especificamente no que se trata de sumarização específica para dispositivos de computação móvel através da utilização de cinco técnicas de sumarização. Entretanto a solução apresentada por Buyukkokten não considera o contexto ambiental nem o perfil do usuário no processo de sumarização, focando-se na geração de sumário genérico. [Buyukkokten 2001].

5. Conclusão

Este artigo apresentou um novo modelo de adaptação em dispositivos de Internet móvel através da implementação de uma arquitetura de sumarização automática de objetos multimídia – notadamente de objetos textuais - baseada em contexto ambiental – localização e temporalidade.

Acreditamos que suas principais contribuições envolvam a adaptação de algoritmos de sumarização automática de textos para um funcionamento mais refinado através da utilização de relevância de palavras-chaves relacionadas ao perfil do usuário e do contexto ambiental em que o usuário dinamicamente se encontra como também a preocupação com a adaptação dos objetos envolvidos no ambiente de computação móvel. como possíveis expansões do modelo destacamos o suporte aos demais tipos de objetos – imagem, som e vídeo.

Referências Bibliográficas

Atalato, T., Peraaho, J. (2001) “Designing mobile-aware adaptive hypermedia”, <http://www.wis.win.tue.nl/ah2001/papers/alatalo-2.pdf>. Agosto.

Buyukkokten, O., Garcia-Molina, H., Paepcke, A. (2001) “Seeing the whole in parts: text summarization for web browsing on handheld devices”, <http://www.www10.org/cdrom/papers/594/index.html>.

Dey, A., Abowd, G., Salber, D. (1999) “A context-based infrastructure for smart environments”, <http://www.cc.gatech.edu/fce/ctk/pubs/MANSE99.pdf>.

Gonzalez, Marco.; Lima, Vera L.S. de (2001). Semantic Thesaurus for Automatic Expanded Query in Information Retrieval 8 th Symposium on String Processing and Information Retrieval (SPIRE_01), Chile, IEEE Computer Society Publications, ISBN 0-7695-1192-9. Proceedings, p.68-75.

Mani, I. (2001) “Automatic Summarization”, Jonh Benjamins, USA.

Mani, I., Maybury, M. (1999) “Advances in automatic text summarization”, MIT Press, USA.

Martins, C., Pardo, T., Espina, A., Rino, L. (2001) “Introdução à sumarização automática”, <http://www.nilc.icmc.usp.br/~thiago/RTDC00201-CMartinsEtAl.pdf>.