

Ovídio José Francisco

**Aplicação de técnicas de Mineração de Textos
para Organização e Extração de Históricos de
Decisões de Documentos de Reuniões**

Sorocaba, SP

13 de setembro de 2017

Sumário

1	INTRODUÇÃO	3
	Referências	5

1 Introdução

Reuniões são tarefas presentes em atividades corporativas, ambientes de gestão e organizações de um modo geral. Seu conteúdo é frequentemente registrado em texto na forma de atas para fins de documentação e consulta posterior.

As atas de reunião possuem características particulares, frequentemente apresentam um texto com poucas quebras de parágrafo e sem marcações de estrutura, como capítulos, seções ou quaisquer indicações sobre o tema do texto. Além disso, o estilo formal contribui para a escrita de textos com poucos detalhes, pois o ambiente dá preferência a textos sucintos.

Devido a fatores como a não estruturação e volume dos textos, a localização de um assunto em uma coleção de atas é uma tarefa custosa, especialmente considerando o seu crescimento em uma instituição. Usualmente, são feitas buscas manuais guiadas pela memória ou com uso de ferramentas computacionais baseadas em localização de palavras-chave. Normalmente esse tipo de busca exige a inserção de termos exatos e apresentam ao usuário um documento com as palavras buscadas em destaque, mantendo o texto longo, o que dificulta por exemplo o ranqueamento por relevância.

Uma vez que a ata registra a sucessão de assuntos discutidos na reunião, há interesse em um sistema que aponte trechos de uma ata que tratam de um assunto específico. Tal sistema tem duas principais tarefas: 1) Descobrir quando há uma mudança de assunto. 2) Descobrir quais são esses assuntos. Este trabalho tem como foco principal, a detecção de mudança de assuntos, que pode ser atendida pela segmentação automática de textos (??????). E a extração automática de tópicos.

A tarefa de segmentação textual consiste em dividir um texto em partes que contenham um significado relativamente independente. Em outras palavras, é identificar as posições nas quais há uma mudança significativa de assuntos.

A segmentação de textos é útil em aplicações que trabalham com textos sem indicações de quebras de assunto, ou seja, não apresentam seções ou capítulos, como transcrições automáticas de áudio, vídeos e grandes documentos que contêm vários assuntos como atas de reunião e notícias.

Pode ser usada para melhorar o acesso a informação solicitada por meio de uma consulta, onde é possível oferecer porções menores de texto mais relevantes ao invés de exibir um documento grande que pode conter informações menos pertinentes. A navegação pelo documento pode ser aprimorada, em especial na utilização por usuários com deficiência visual, os quais utilizam sintetizadores de texto como ferramenta de acessibilidade (??).

Além disso, encontrar pontos onde o texto muda de assunto, pode ser útil como etapa de pré-processamento em aplicações voltadas ao entendimento do texto, principalmente em textos longos.

Referências

LEE, J.-K. et al. Two-step sentence extraction for summarization of meeting minutes. In: . [S.l.: s.n.], 2011. v. 39, p. 614–619. Nenhuma citação no texto.

SCHWARTZ-ZIV, M.; WEISBACH, M. S. What do boards really do? evidence from minutes of board meetings. *Journal of Financial Economics*, v. 108, n. 13, p. 349–366, 2013. Nenhuma citação no texto.

TAKAHASHI, M.; KIDO, K.; HASHIMOTO, K. Towards the profitability trend extraction from the board meeting proceedings. *17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems - KES2013*, v. 78, 2013. Nenhuma citação no texto.