

RADIALIZE: UMA FERRAMENTA PARA DESCOBERTA DE CONTEÚDO EM ÁUDIO NA WEB.

JANDEILSON ROLDAO BORGES (Autor), ALVARO RODRIGUES PEREIRA JUNIOR (Orientador)

Instituição de Ensino - Universidade Federal de Ouro Preto

Palavras Chaves:

Coleta, Web, Stream, Rádio, Áudio

Resumo:

O trabalho desenvolvido tem como objetivo o descobrimento de streams de áudio na web. Junto com o stream são coletadas as informações de sua origem (nome e página da rádio) de forma automática, que são dados importantes para o projeto Radialize, que se trata de uma ferramenta para recomendação e transmissão de streams de rádios pela web. Foi desenvolvida uma arquitetura que dada uma página da web, são extraídas todas as urls no código fonte e passam a ser candidatos a stream. Em seguida os candidatos são validados, utilizando-se um método heurístico de captura de fluxo na rede, o sistema se conecta ao possível stream, e por um tempo mede as alterações no tamanho do arquivo recebido, caso haja um crescimento constante no tamanho nesse intervalo, a url é considerada um stream. Na maioria dos casos o stream é encontrado no código fonte da página explorada. Quando essa informação não está explícita, o programa coleta os pacotes trocados pela rede do computador com a fonte da transmissão do áudio e dessa forma consegue obter o stream e validá-lo. Após o stream ser encontrado, é feita a recuperação do nome da rádio através do metadado disponibilizado no próprio stream. Foram realizados testes utilizando-se 100 rádios nacionais com site e streams já conhecidos previamente, o sistema apresentou como resultado uma revocação de 0.9 e precisão de 1, o que mostra que todos os streams encontrados são confiáveis. No entanto, o sistema não cobre todos os tipos de streams, pois existem protocolos diferentes que são de difícil identificação.

Publicado em:

- Evento: Encontro de Saberes 2015
- Área: CIÊNCIAS EXATAS E DA TERRA
- Subárea: CIÊNCIA DA COMPUTAÇÃO