

Q1: サーバーが送りたい情報とユーザがほしい情報のORは、本当の通知すべき情報と思うが、XML文書の差分を考慮する必要がないではないか? 想定しているアプリケーションは?

A: ご指摘の通りです。ただし、サーバー側の意図とユーザ側の意図は、必ずしも一致しているとはいえないと思います。2者の調整が必要であるので、第3者からWebページの変更監視を行うべきであるが(静的Webページに対応しているWebSCANはそうになっているが)、サーバー内部のデータベースと変換プログラムの監視を行えないかもしれないという問題があります。ユーザの意図に関しては、プレゼンであまり説明しなかったが、ユーザプロファイルを用いてユーザがほしい情報のみを受信することが可能となっています。今後、サーバー意図のルール記述、ユーザプロファイルについて検討していきたいと思います。

また、本研究では、頻繁に変更される情報(ニュース, etc.)から、どうやって重要な情報を検出するかをフォーカスしている。そのため、我々は、変更前後の情報を比較して変更した情報は今までにない情報であるか(新鮮度)、最近よくある情報か(流行度)といった尺度を用いて重要と思われる情報をピックアップしてユーザに通知するというアプローチを取っています。

Q2: XML文書の差分に関しては、一般的にはedit-distanceという手法を利用するが、提案手法との違いは?

A: 本研究では、文書の構造的差については、XPathを用いてましたが、edit-distanceを利用することも考えられます。我々は、変更後の文書の構造と類似する構造を持っている文書数を用いて、変更後の文書の構造がどの程度新しいかについて調べています。また、内容評価(新鮮度・流行度)のための比較対象を特定するためにXPathを利用しています。今後、エージディスタンスの利用を検討したいと思います。