

WWW ページのディレクトリ構造に対応したレーティング、フィルタリングに関する提案

西塚 一樹* 西埜 覚* 苗村 憲司**
 通信・放送機構 横浜コンテンツリサーチセンター* 慶應義塾大学環境情報学部**

1. はじめに

レーティング、フィルタリングシステムの国際標準化を目指している ICRA(Internet Content Rating Association) [1] は「フィルタリングにはセルフレイティングによるデータを利用するのが適切であり、必要である。」とし、反面、「セルフレイティングには相応のコストが必要で営利目的の大企業以外は難しい。」としている。コストが必要な理由は、A. レーティングの判断に時間がかかる B. ページ単位でレーティングデータの記述が必要であることが考えられる。

ICRA は、新しいシステムのためにアンケート形式でレーティングデータの作成をするツールを HTML 作成ツールに組み込む形で提供する。これによって、レーティングの判断にかかる時間は削減できると考えられる。

2. 現状と課題

(1) ホームページ作成ツールの利用状況

WWW ページの HTML 構文に META タグの generator で記述されている HTML 作成ツールを調査した。利用状況は 81,038 件中 24,317 件(30.0%)である(図 1)。HTML 作成ツールの利用が今後大きく拡大されるか判断が難しい。

ツール利用ページ 24,317(30.0%)	ツール非利用ページ 56,721(70.0%)
---------------------------	----------------------------

図 1 HTML 作成ツールの利用状況

(2) ラベルビューロの課題

W3C(WWW コンソーシアム)の勧告する PICS ルール [2] に基づいたラベルビューロは問い合わせのあった URL で登録されているレーティングデータが存在しなければ、そのディレクトリを代表するページ(「/」で終わる URL に対応するページ)の URL から上位へ対応するレーティングデータがあるまで問い合わせを繰り返し(例 A)、その結果により PICS ラベルを応答する。

例A レーティングデータの問い合わせ順

www.domain.com/dir1/dir2/dir3/page.html
 ↑ V ↑ N ↑ III ↑ II ↑ I

問い合わせは先頭から↑の指す文字列まで行われます。

ラベルビューロはディレクトリ構造に対応しているが、コンテンツ提供者の WWW サーバにあるページへセルフレイティングデータの問い合わせには利用できない。

(3) インターネットエクスプローラー 5.0 の課題
 一般的なブラウザのひとつインターネットエクスプローラーの Ver5.0 では、コンテンツアダプタの機能を有効にするとアクセスする WWW ページに PICS ラベルの記述がない場合、ドメインを代表するページへアクセスしそのページの PICS ラベルをフィルタリングに利用する。(図 2)

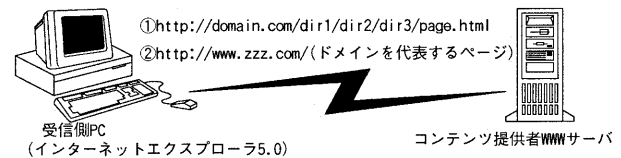


図 2 インターネットエクスプローラー 5.0 での応答

しかし、これはサーバを提供するときのようにドメインを共有している場合、下位のディレクトリを利用するコンテンツ提供者のページのフィルタリングに利用できない。

(4) ICRA の新しいシステムの場合

受信者は、アクセスするページにレーティングデータの記述がない場合、そのディレクトリを代表するページへアクセスしそのレーティングデータをフィルタリングに利用する。代表するページにレーティングデータの記述がない場合、ドメインを代表するページまで上位へ向かってアクセスを繰り返す。アクセスの繰り返しには各ディレクトリを代表するページが存在することが必須条件である。1 ページの表示のために数ページへのアクセスを必要とする可能性があり、コンテンツ提供者の WWW サーバへの負荷と通信コストが高くなる懸念される。

上述した課題を解決するために、レーティングデータを記述したファイルを利用したディレクトリ構造に対応する方式を提案する。

3. ディレクトリ構造対応の方式

レーティングデータを記述したファイル(仮に index.BB.html とする)を利用したディレクトリ構造に対応する方式とは、① index.BB.html をコ

*A proposal of approaching to rating and filtering for the directory tree of contents on WWW, Kazuki NISHIDUKA, Satoru NISHINO, Telecommunications Advancement Organization of Japan, and Kenji NAEMURA, Keio University

コンテンツ提供者が用意し、②index.BB.htmlを受信者がフィルタリングに利用するものである。

(1) index.BB.html の用意

コンテンツ提供者は WWW サーバ上の管理するすべてのディレクトリに対して、index.BB.htmlを用意する(図3)。

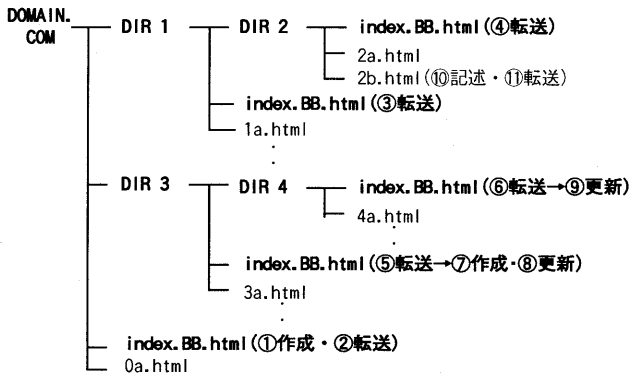


図3 コンテンツ提供者のディレクトリ構造

1) index.BB.html の作成とレイティングデータの記述

- ・管理する最上位のディレクトリ(図3①)
- ・上位とレイティングデータの異なるディレクトリ(図3⑦)
- ・ディレクトリとレイティングデータの異なるページ(図3⑩)

に対して行われる。すべてのページに同一のレイティングデータが適用できる場合、作成は1度で済む。

2) index.BB.html の転送と更新

- ・index.BB.html を最上位のディレクトリに対して転送する(図3②)。
- ・ディレクトリの作成時 index.BB.html を自動的に転送する(図3③~⑥)。
- ・レイティングデータの異なるディレクトリのindex.BB.html を更新する(図3⑧)。
- ・更新したディレクトリの下位のすべてのディレクトリのindex.BB.html を自動的にチェックし、更新する(図3⑨)。

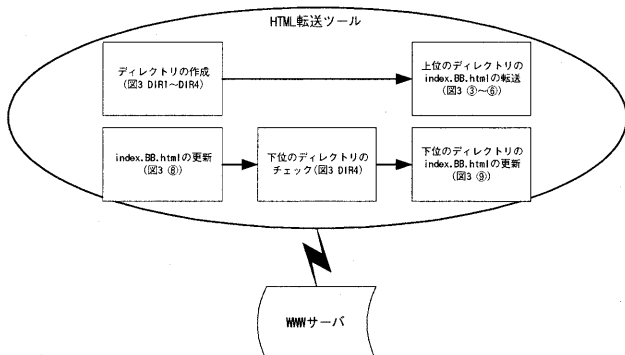


図4 index.BB.html の自動転送と更新

index.BB.html は、作成対象のディレクトリ以外には自動的に転送または更新され、すべてのディレクトリに生成される(図4)。

(2) index.BB.html を利用したフィルタリング

受信者はアクセスするページか同じディレクトリのindex.BB.htmlのレイティングデータでフィルタリングを行う。図3の場合、2a.htmlにレイティングデータの記述がなければ、DIR2のindex.BB.htmlへアクセスし、そのレイティングデータでフィルタリングを行う。(図5)



図5 index.BB.html を利用したフィルタリングの応答

コンテンツ提供者がindex.BB.htmlをすべてのディレクトリに用意するので、受信者がコンテンツ提供者のWWWサーバへのアクセスする回数は最大2回である。

(3) 実現のための機能の追加

1) HTML 転送ツールへの機能の追加

- ・新たなディレクトリを作成するとき、上位のディレクトリのindex.BB.htmlを転送する。
- ・index.BB.htmlを更新するとき、下位のディレクトリすべてのindex.BB.htmlをチェックし、更新する。

という機能を追加する。

2) ブラウザへの機能の追加

ブラウザにアクセスするページと同じディレクトリのindex.BB.htmlへアクセスする機能を追加する。

5. まとめ

ディレクトリ構造対応の方式を採用することで個別のページへレイティングデータを記述する時間を削減でき、アクセス数、通信コストの増加も最小限に抑制できる。また、コンテンツ提供者のディレクトリ単位で設定ができ、ドメインを共有している場合でも他のコンテンツ提供者に影響を与えない。この方式では、コンテンツ提供者がセルフレイティングしたデータによるフィルタリングを確実に行うことができる。

参考文献

- [1] <http://www.icra.org/>
- [2] <http://www.w3c.org/PICS/>
- [3] <http://www.nmda.or.jp/>