

# 動的なライツマネジメントのための WebCrawler型DRM情報半自動付加機能の実装実験 An experimental study of a dynamic content rights management system in web crawler model

瀬高昌弘<sup>†</sup>      天野直紀<sup>†</sup>      嶋津恵子<sup>‡</sup>  
Masahiro Sedaka<sup>†</sup>      Naoki Amano<sup>†</sup>      Keiko Shimazu<sup>‡</sup>

<sup>†</sup> 東京工科大学 大学院 バイオ・情報メディア研究科  
<sup>‡</sup> 慶應義塾大学 デジタルメディア・コンテンツ統合機構

<sup>†</sup> Graduate School of Bionics, Computer and Media Sciences, Tokyo Univ. of Technology

<sup>‡</sup> Research Institute for Digital Media and Content, Keio Univ.

## 要旨

現在我々はコンテンツの利用促進を目指し、新たなアクセスコントロールの方法を検討している。これまでのCMS(Content Management System)の機能に依存した、画一的な操作制限方式ではなく、“宣言ファイル”に記述された操作制御情報をコンテンツに付与することで、コンテンツ発信者の意図や目的を直接反映した操作制御方式である。分散サーバ上にコンテンツとともに併載された“宣言ファイル”を検索エンジン用クローラで収集することで実現させる。利用者からのアクセス要求が発生する都度、別の専用サーバ上で個々のコンテンツに対するダイナミックな権限付与を行う。今回の権限付与部分を先行してシステム構築した。この方法は、ダウンロード開始までの処理に時間がかかることが懸念されるが、実証検証の結果、実利用に耐えられることを確認した。

## 1. はじめに

Web 上で配信されるコンテンツは参照権や複製権について利用者ごとに制御されていることが多い。現状では著作権法に基づき、不正使用の抑止や権利の侵害の事前回避を重視するあまりに、コンテンツの流通そのものが妨げられ、場合によってはコンテンツを公開しないことがある[1]。一般的にアナログコンテンツがデジタル化され、ネットワーク上で共有されると付加価値が高まると考えられるが、実際にはコンテンツが死蔵されている[2]。また、公開されたコンテンツに対する利用者の操作制御はCMS<sup>1</sup>の機能に依存した画一的なものであり、コンテンツの有効利用を狙ったものは存在しない。利用者に応じた細かい操作制御を行うことができないと、コンテンツ提供者はコンテンツを公開することよりも危険回避を優先し、コンテンツの公開を諦める。そこで我々は、コンテンツ発信者の意図や目的を直接反映し、きめ細かい操作制御を実現できる新たなライツマネジメントのフレームワークを設計中である。ここで言うライツマネジメントとはデジタルコンテンツ共有用の DRM システムの基本仕様である。我々は今回、コンテンツに対する利用者ごとの権限付与を DRM システム上のどの部分で実現させるかを再考した。これまではそれぞれの CMS 自身がこの機能を持つことを前提としており、利用者からのアクセス要求が発生した都度、権限等を確認し設定に従って承諾する仕組みを提供している。この方法は、安定した制御を実現できるが、行える操作制御内容は利用するシステムに依存するため、コンテンツ提供者が実現したい制御切り分けが行えないことがある。これに対し、我々はコンテンツの発信者の意図や目的を直接反映することを目指した。そこで、コンテンツ発信者が操作制御宣言ファイルに操作制御内容を記述し、このファイルを検索エンジン用クローラで収集する WebCrawler 型を考案した。収集したこれらのファイルを基にして、利用者からのアクセス要求が発生するごとにコンテンツに対して操作制御情報を付与する。

本来はこの WebCrawler 型の操作制御宣言ファイル収集とアクセス要求ごとのコンテンツに対して操

<sup>1</sup> Contents Management System デジタルコンテンツの管理を行うシステムのこと。

作制御情報付与の有用性を検証すべきだが、今回はアクセス要求ごとに操作制御情報を付与する際の問題点を取り上げる。操作制御情報を付与する時間がアクセス要求発生の際にかかるため、ダウンロード開始までに時間がかかることが懸念される。我々は、このダウンロード開始までの処理時間が許容範囲内に収まっていることを確認したことを報告する。

本書は次の構成をとる。2章では動的なコンテンツ操作制御情報付与機能について説明する。3章では実装したシステムについて検証を行い、4章でその考察を述べる。5章でまとめと今後の課題を記す。

## 2. アクセス要求ごとのコンテンツ操作制御情報付与機能

我々の提案するコンテンツに対する操作制御の方法は、(1)CMS に依存しないコンテンツ発信者の自由なコンテンツ利用制限と促進の宣言と、(2)利用者のプロフィール取得、(3)コンテンツにアクセス要求が発生した都度、(1)と(2)からコンテンツへ操作制御情報付与を行う機能から成る。

これらの3つの機能は図 1 に示すとおり、2つのサーバと1つのコンテンツ作成環境で実現される。

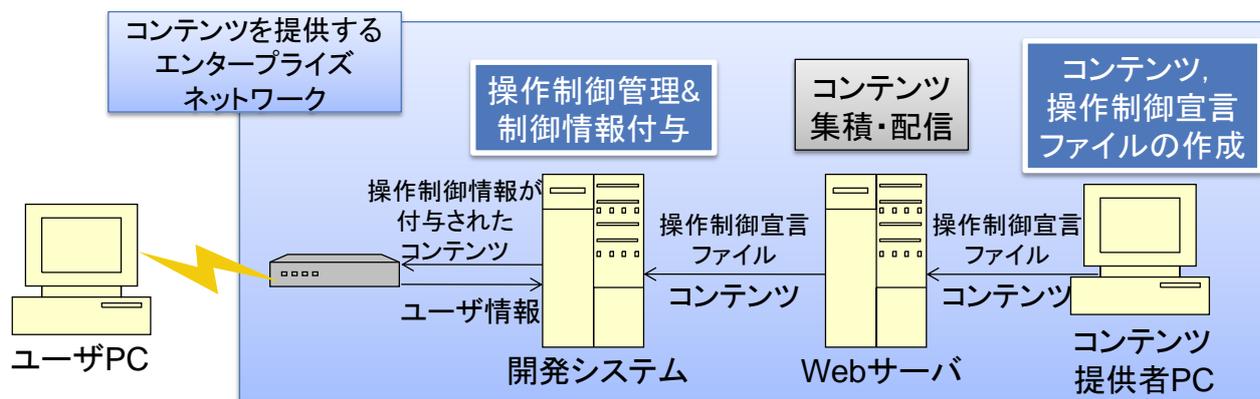


図 1 3つの機能の配置

### 2.1. CMS に依存しないコンテンツ利用制限と促進宣言

コンテンツ発信者は操作制御宣言ファイルに操作制御内容を記述し、Webサーバ上の公開したいコンテンツと同じディレクトリに置く(図 1 の「コンテンツと操作制御宣言ファイルの作成」)。操作制御宣言ファイルは、図 2 に例を示すとおり、誰に権限を与えるのかと何の権限を与えるかの情報の組で構成される。前者は特定個人(や企業や学校といった社会的に認められたグループ)を指定し、後者はコンテンツに対する(1)テキスト選択・コピーの許可、(2)印刷の許可、(3)閲覧可能期間の3種類である。また、構造化形式ではなく、一般の利用者が直感的に作成できるようにカンマ(,)で区切る形式を採用した。

```

東京工科大学, 印刷許可
瀬高昌弘, テキストコピー許可
:
    
```

図 2 操作制御宣言ファイルの例

今回の実験の目的は、コンテンツに対する操作制御を付与する際に発生する負荷が、コンテンツダウンロード開始までの時間ほどの程度影響するかの確認である。この負荷はコンテンツに対するアクセス要求ごとに毎回発生するので、軽視できない。目的に対し、的確に検証するために、操作制御宣言ファイルは、最低限の情報の記載にとどめた。具体的には図 3 に示すとおり、誰に権限を与えるのかは全ての利用者に対してとし、全ての利用者に対する許可を\*(アスタリスク)と記した。また、アクセスしてきた利用者全てに対し、どの操作を許可するかをコンテンツごとに指定した。許可する操作は、(ア)コンテンツ内のテキスト選択・コピーの許可と、(イ)ファイル印刷の許可の、いずれかもしくは(ウ)両方であ

る。コンテンツ内のテキスト選択・コピーの許可は”テキストコピー許可”，また印刷の許可は”印刷許可”と記した。

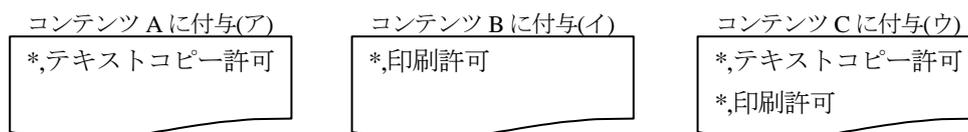


図 3 実験用操作制御宣言ファイルの例

## 2.2. 利用者のプロフィール取得

取得する情報から、特定個人（もしくは企業や団体等に所属していること）を確認する。この確認された値と 2.1 の方法で記された操作制御宣言ファイル上の値を比較し、許可する操作をシステムが決定する。取得する情報は、(A)コンテンツ利用者が何らかの方法で接続認証（ログイン）を行うことで収集もしくは参照用データベースから獲得するものと、(B)接続認証なしに得られるものがある。さらに後者は、一般の Web ページ参照時に得られるアクセス元 IP アドレスなどの情報とコンテンツにプログラムを組み込むことで得られる情報がある。コンテンツの参照などの操作を開始したタイミングでこのプログラムが起動し、クライアント環境に関する情報をサーバ側に返す方式である。

なお、今回の検証では 2.1 に述べたとおり、全ての利用者を想定しているので、実装から省いた。

## 2.3. コンテンツへの操作制御情報付与

この機能は、(i)我々の提案の主題である操作制御内容が記述された操作制御”宣言ファイル”を検索エンジン用クローラで収集する部分と、(ii)コンテンツに対するアクセス要求が発生したとき、”宣言ファイル”に従ってコンテンツに制御を加える部分からなる。前者は、コンテンツ提供者が Web サーバ上のコンテンツと同じ場所に置いた操作制御宣言ファイルを検索エンジン用クローラで収集し、蓄積する。このとき操作制御”宣言ファイル”のファイル名規則は、コンテンツのファイル名に拡張子”rinfo”を加える(e.g. sample.pdf → sample.pdf.rinfo)。後者は、コンテンツへのアクセス要求が発生したときに、収集された宣言ファイルの内容を参照し、操作制御情報をコンテンツに付与する。このとき前処理としてアクセス要求があったコンテンツを PDF ファイルに変換する<sup>2</sup>。操作制御情報の付与は、PDF ファイル操作のライブラリ (iText) を用いる[4]。この段階で、操作制御宣言ファイルが存在しない場合はコンテンツの操作そのものを中断する。

## 3. 実験

前章で述べた、これまでとは異なる操作制御方法を搭載した際にも、利用可能な時間内で処理されるかどうかを検証した。本来は WebCrawler 型の操作制御宣言ファイル収集とアクセス要求ごとのコンテンツに対する操作制御情報付与の有用性について検証すべきであるが、次の機会に行う。今回の実験は、ネットワークを介したファイルダウンロードにかかる所要時間という一般的な問題意識を背景にしている。我々の提案であるアクセス要求ごとに操作制御情報を付与する方式を採用すると、あらかじめ CMS 上で操作制御情報を付与しておく一元管理型と比べ、処理効率性が悪化する。これは操作制御情報を付与する時間がアクセス要求発生の際に発生することによる。そこで、従来手法と比較し、どのくらい、ダウンロード開始までの時間に差が発生するかを確認した。

実験方法は、一元管理型を想定した、操作制御情報を付与しないで直接ダウンロードする方法と提案手法とで応答にかかった時間を比較した。提案手法は検索エンジン用クローラを用いず、すでにコンテンツを PDF に変換したものの配置先をあらかじめ指定しておく。アクセス要求ごとにそのコンテンツと同じディレクトリに存在する操作制御宣言ファイルを確認する。実験に用いる PDF ファイルはファイルサイズが約 1MB のファイルを用意し、100 回ずつプログラムで自動的にダウンロードした。

<sup>2</sup> PDF ファイルに変換したときのフロントページの作成は Java Servlet/JSP で構築した。

従来手法と提案手法でダウンロード開始までにかかった時間の平均時間・最小値・最大値は表 1 の通りである。提案手法は従来手法に比べておよそ 5.7 倍の時間がかかるものの、最大でも 0.7 秒以内に収まった。

表 1 応答にかかった時間(単位：msec)

方法	平均値	最小値	最大値
従来手法	80	79	83
提案手法	457	346	658

## 4. 考察

実験結果より、付与が従来手法に比べて提案手法はダウンロード開始までの時間は長くなる。一方、先行事例によると 10 秒以内であれば、Web コンテンツが表示されるまで待つことができると 8 割以上の人が答えている[5]。従って我々が提案する手法は実利用に耐えられると言える。さらに、コンテンツ発信者は、CMS が提供する操作制御機能に依存することなく、利用方法を宣言できる。検証したダウンロード時間の増加は、この有用性と比較すると十分耐えられるものであるといえる。

今回はアクセス要求ごとの操作制御情報付与部分のみ実装し、懸念される問題について検証したが、今後は操作制御宣言ファイルのフォーマットを決定や、コンテンツ作成者に操作制御宣言ファイルを確実にしてもらおう方法についても検討する必要がある。また、他のモジュールの実装を行い、フレームワーク全体としての有用性の確認を行う必要がある。

## 5. まとめ

我々はコンテンツ発信者の意図や目的を直接反映した操作制限と利用拡大を実現できる、新たなライツマネジメントのフレームワークを設計中であるが、今回、コンテンツに対する利用者ごとの権限付与を DRM システム上のどの部分で実現させるかを再考した。そして現状のシステムに対し、コンテンツ発信者が操作制御宣言ファイルに操作制御内容を記述し、このファイルを検索エンジン用クローラで収集する WebCrawler 型を考案した。収集した操作制御宣言ファイルを基にして、利用者からのアクセス要求が発生するごとにコンテンツに対して操作制御情報の付与を行う。WebCrawler 型とすることにより、すでに Web サーバを運用しているネットワークに提案したフレームワークを用いたシステムを追加し、コンテンツ作成者が操作制御情報を宣言するのみでシステムがコンテンツの存在を収集し、コンテンツに対する権利管理・運用を行うことができるという利点がある。本来はこの WebCrawler 型の操作制御宣言ファイル収集とアクセス要求ごとのコンテンツに対する操作制御情報付与の有用性を検証すべきだが、時間の都合上次の機会に行うこととし、今回はアクセス要求ごとに操作制御情報を付与する時間がかかるという問題を取り上げた。そして、操作制御情報付与部分を実装し、実利用に耐えられることを確認した。今後はフレームワーク全体の有用性を検証したい。

## 参考文献

- [1] 札幌市立大学 Ideas Agency 「第十回産学連携講座の再録:Creative Commons の潮流と iCommons Summit 2008」 [http://www.ideasagency.jp/2007/09/post\\_10.html](http://www.ideasagency.jp/2007/09/post_10.html)
- [2] 山形浩生, 永江朗, 白田秀彰, 論座「岐路に立つネット時代の知財戦略」朝日新聞社, 2007
- [3] 三好豊, 著作権法, 中央経済社, 2004
- [4] Bruno Lowagie, Paulo Soares 「iText, a Free Java-PDF Library」 <http://www.lowagie.com/iText/>
- [5] MDM ネットサーベイ 「'時感'意識実態調査の結果について」 [http://www2.mdb-net.com/netsurvey/w\\_report/report34.html](http://www2.mdb-net.com/netsurvey/w_report/report34.html)