

## 低コスト会議録出版を志向する電子投稿・編集システム

### Submission and Edition System Orienting Low-Cost Publishing of Conference Proceedings

上田 哲史\*

Tetushi Ueta\*

\*徳島大学

Tokushima University

国際会議、研究集会などの会議録を生成するシステムをウェブアプリケーション等と組み合わせ実現した。投稿、査読、再投稿プロセスを、フリーウェアにパッチを施した投稿システムで運用する。内部で使用されるデータベースから必要な情報を取得し、各論文PDFへのリンクを埋め込んだ講演プログラム目次ページを自動的に生成する。採択論文はPDFlibを用いてページ番号やロゴ、様々なページへのしおりが埋め込まれる。いくつかの会議で実際に本システムを運用し、出版コストを金銭的にも人的にも低く抑えることができた。本稿では上記システムの概要を報告する。

**キーワード** : 会議録, 出版業務, 投稿システム, PDFlib, L<sup>A</sup>T<sub>E</sub>X

For managing paper submission, reviewing, and publishing proceedings in an international conference or workshop, we develop a publication system featuring a web-based interface submission server and an edition system by using PHP scripts, PDFlib, and L<sup>A</sup>T<sub>E</sub>X. The generated navigation PDF enables readers to browse conference papers easily. The submission server is provided by a patched freeware. Conference information, logos, page numbers, and a link to the navigation PDF are embedded in each paper. The navigation page is generated by L<sup>A</sup>T<sub>E</sub>X source files whose basic information are composed by extracted data from the submission server. In this paper, we describe the system in detail.

**Keywords**: proceedings, publication affairs, submission server, PDFlib, L<sup>A</sup>T<sub>E</sub>X

## 1. はじめに

昨今、国際会議や研究会における講演論文集は、急速に印刷冊子による提供形態は淘汰され、PDF論文を格納したCD-ROMなど、デジタルメディアによる供給に取って代わられた。その背景には、環境への配慮、印刷費用圧縮、印刷物運搬の負担軽減などが考えられるが、PDFブラウザで目次や索引のリンクを辿るだけの手軽な論文閲覧や、検索機能は、読者へ大いなる

利便性を提供している。

論文の投稿プロセスも、ハードコピーによるカメラレディ郵送からウェブやメールによるデジタル版下投稿へと変化した。論文収集、査読者への送付、カメラレディ受け付けを全て郵送で行っていた時代と比較すれば、編集効率は格段に上がっている。さらに近年、論文の収集からCD-ROMの発行を担う業者が現れ始めたため、十分に予算がある会議ではそれら業者の利用によって出版担当者の負担を減らすことが考えられる。

しかし、出版業務を全部業者に委託する場合であっても、採否の確定、著者の会議登録確認、講演プログラム編成などに関連する業者への指示、原稿のチェック、差し替えなどの作業は出

\*マルチメディア研究部門

〒770-8506 徳島市市南常三島町 2-1

Multimedia Research Division

〒770-8506 2-1, Minami-Josanjima, Tokushima-shi, Tokushima

E-mail : tetsushi@is.tokushima-u.ac.jp

版担当者が行わざるを得ない。業者と頻繁に連絡が必要となり、担当者が直接版下を編集出来ないだけに、むしろ泥縄の作業や連絡が増え、ミスが発生する原因にもなる。

ところで、PDF 版会議録の構成要素は単純である。講演者らが投稿してくる論文に適切にページ番号を入れ、講演プログラムに従って講演者情報や論文 PDF へのリンクを作成するのみである。論文の収集と査読さえ片付けば、会議録編集自体はルーチンワーク化することが可能である。

本報告では、UNIX サーバの PC 1 台で、論文投稿から査読、最終原稿収集、CD-ROM 会議録生成の過程を管理できるシステムについて詳細を述べる。本システムを使用し、出版業務にかかる時間的・金銭的・人的コストを大幅に抑えることができた。

## 2. 会議録の構成要素

### 2.1 会議録の条件

本論文で仮定する会議録の条件は以下とする：

1. Adobe Reader などの PDF ブラウザで閲覧・検索・印刷の操作が行えること
2. CD-ROM に収録する内容を全部印刷すれば、紙媒体の会議録が出来上がること

これらのためには、投稿論文が PDF であること、もしくは投稿サーバ内が、別形式の投稿論文フォーマットを PDF に変換できる仕組みを持つことが要求される。また、PDF 内のフォントも埋め込まれている必要がある。

昨今のウェブブラウザは、アドオンとして Adobe Reader を組み込むことが標準的なので、ブラウザとしてウェブブラウザを用いて、ナビゲーションを HTML で記述することも考えられるが、条件 2 を満たすことができない。

### 2.2 会議録のファイル構成

会議録のファイル構成は以下のように設計する：

1. 起動用ページ (PDF; オプション)  
Windows 環境であれば、メディアを PC に挿入したとき、ブラウザが自動起動し、以下のナビゲーションページを表示するように設計してもよいが、そうでなければ INDEX.pdf のように目立つ名前の PDF を置くなどし、ユーザにブラウザの起動を促す。中身は次のナビゲーションページをロードするようにしておく。
2. ナビゲーションページ (PDF)  
一枚ものの PDF であり、各ページをリンクボタンを押すことにより呼び出せるように設計する。
3. 前付ページおよび講演プログラム目次ページ (PDF)  
序言、会議の概要、会議組織、施設の紹介、広告、地図、地域情報、スケジュールを載せる。続いて、講演プログラムや索引を納める。各講演論文名には論文 PDF へのリンクが張ってある。多くの会議参加者はスケジュールに合わせて論文を閲覧しようとするので、上記ナビゲーションページよりはこのファイルをハブとして閲覧するものと思われる。
4. 各論文 (フォルダに納められた PDF)  
各論文にはページ番号が割り付けられ、会議のロゴ等が刷り込まれる。また、ナビゲーションページや目次ページへのしおりを付加する。なお、前付・目次ページと全論文を 1 つの PDF にまとめると条件 2 を満たすことができる。

### 2.3 ツールの設計思想

信頼性の高い内容の会議録生成のためには、編集システム中の各ツールは以下の要件を満たす必要がある。

1. 投稿システムのデータベースを常に参照し、二次的な情報の生成・参照・変更を抑制すること。二次的情報はデータの一

貫性が失われ、更新漏れによるバグが混入する可能性がある。

2. できるだけ手順が機械的で再現性が高いこと。手作業が入ればそれに比例してミスが混入する機会が増えるので、各機能を果たすスクリプトがバッチ的に動かされるべきである。できれば make 一発で全コンテンツが生成されることが望ましい。
3. データベースに入らないローカルな情報はどうしても編集に手を入れざるをえないが、それらをできるだけ L<sup>A</sup>T<sub>E</sub>X ソースとして記述し、Makefile において依存ファイルとして登録し、更新漏れのないようにすること。

これらの設計思想をもとにシステム設計を行い、パッケージを開発・リリースした。本格的な国際会議の出版業務をこのシステムで手掛けることができる。満足のいく出来映えの会議録を作成するにはやや手動による作業を加えねばならないが、PDF 論文と講演プログラムだけが必要な簡単な集会であれば、採択論文の決定後はほぼ無手順で会議録ができあがる。

### 3. システム構成

本システムの構成を図 1 に示す。必要なハードウェアは UNIX が動作する PC 1 台である。本論文では PC では FreeBSD 6.2 RELEASE を、Dell PowerEdge 840 (Xeon 2.66 GHz dual, 2GB memory) で用いたが、本システムの動作にはごく一般的な PC の性能で十分である。

必要なソフトウェアは、Apache などのウェブサーバソフトウェア、PHP 4.x, MySQL 4.x, L<sup>A</sup>T<sub>E</sub>X, Commence Conference Manager (CCM)<sup>1)</sup>, PDFlib であり、PDFlib を除いて全てフリーで入手できるものである。PHP, MySQL は CCM の動作条件に含まれるため、ウェブサーバソフトウェアは PHP, MySQL のアクセスのためのモジュールを組み込んでおく必要がある。

PDFlib<sup>5)</sup> は本システム中で唯一有償のソフトウェアであり、PDI (PDF Import Library) と共に購入する必要がある。PDFlib は PDF を操作するインタフェースであり、多くのコンピュータ言語バインディングを提供している。本システムでは PHP で利用した。投稿論文に会議タイトル、ロゴ、ページ番号等を刷り込むために PDFlib が必要となる。PHP は CCM の枠組みとは独立に、CCM が内部的に生成する MySQL データベースへアクセス出来るので、PDFlib の言語バインディングと併せて、PHP で統一してシステム作りができる。

PDFlib と同様な機能は、Adobe Acrobat の JavaScript や FOP, iText, JasperReports など で実現されるが、いずれも自動化を困難にするため、上記設計思想 2 を満たせなくなる。

L<sup>A</sup>T<sub>E</sub>X は本システムのキーワードであり、PHP の処理結果が、目次ページに対応する L<sup>A</sup>T<sub>E</sub>X ソースを自動で生成する。hyperref<sup>7)</sup> を用いて論文 PDF への相対的なパスを、目次や索引にリンクとして埋め込むことができる。対応する dviware や Adobe Distiller を用いることにより、印刷にも十分堪えうる高品質な PDF 版下を生成できる。

システム全体は、投稿システムおよび編集システムに分かれる。

#### 3.1 投稿システム

CCM は IAPR によって開発されたフリーウェア<sup>1)</sup> であり、PHP によって書かれ、内部で MySQL を用いている。著者(投稿者)は、ウェブとメールの組合せで、アカウント・パスワード取得、著者情報入力、論文投稿が行える。1 つのアカウントでの複数の論文投稿ができ、締切まで何度でもそれらの再投稿が可能である。CCM に組み込まれた認証基盤の管理者権限を獲得すれば、ウェブ上で管理業務の全て、すなわち、論文の投稿、査読(割り当て、結果収集)、著者への連絡・コメント送信、再投稿などのプロセスが実行・管理できる。

本システムでは CCM 自体にいくつかのバグフィックスと、著者の所属カラムを増やすため

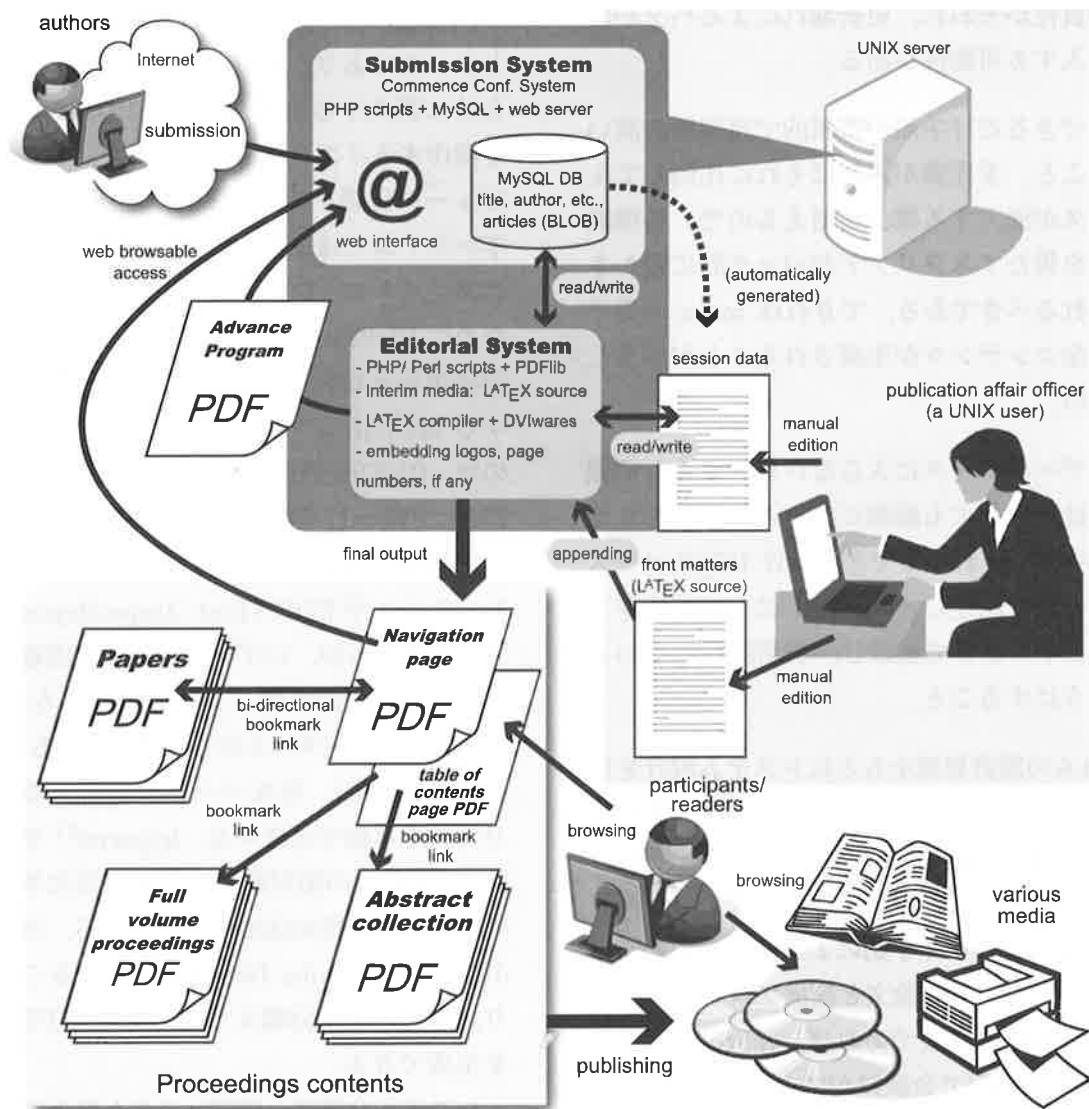


図1 システム構成図

のパッチを適用しているのみで、ほぼオリジナルのまま機能を利用している。そのため CCM の特徴の詳細は文献<sup>1)</sup> の情報に譲ることにする。また、ターゲットとしては英語の会議録を仮定する。日本語をデータとして扱うには CCM をさらに修正せねばならない。しかし、原稿自体は埋め込まれた日本語フォントを含む PDF であっても何ら問題はない。

投稿原稿は BLOB 形式として MySQL のフィールドデータに直接に格納される。これは、後にアプリケーションを開発し、CCM と連携させる場合、SQL クエリのみで原稿を取り出せるというメリットがある。しかし、他の多くの投稿システムでは処理パフォーマンスの点か

ら、原稿は MySQL 内部には置かずにはファイルシステムに書き出す実装が多い。経験的ではあるが、一つの論文が数 MByte 以内で 200 件程度の投稿規模の会議であれば BLOB 形式で問題とはならなかった。

フリーウェアである投稿システムは、START V1<sup>2)</sup>、CyberChair<sup>3)</sup> などがある。START V1 は Perl で書かれた軽いシステムであるが、高度な認証機能や、再投稿の機構がない。CyberChair は Python で書かれているが、データベースは用いられておらず、よってカスタマイズが難しい。また、商用の投稿システムとしては、Epapers<sup>4)</sup>、START II<sup>2)</sup> などがある。以上のシステムはどれも採録論文を出力する機能はあるもの

の、会議録としての体裁を整える編集機能は付属していない。

### 3.2 編集システム

編集システムは以下の各ツール (PHP スクリプト) によって構成される。

#### 3.2.1 論文抽出ツール

採録される論文について、その論文番号や氏名をファイル名として PDF ファイルに書き出す。

#### 3.2.2 刻印ツール

セッションデータに従って論文データをデータベースから抽出し、PDF の実体をファイルシステムに書き出す。そして PDFlib により各論文のノブルとして、通し番号によるページ番号および会議タイトル、ロゴ等を刷り込む。また、しおりにはナビゲーションページや各種会議情報へのリンクを張る。また、目次ページと双方向リンクをしおりに書き、ユーザに便宜を図る。最後に、このツールで生成されたページ番号情報をセッションデータに書き戻す。

セッションデータは、時間、場所、セッションタイトル、座長、それに続いて当該セッションで予定されている講演に対応する論文番号が記述される。現バージョンの CCM ではまだ講演プログラムを生成する機能は未実装であるため\*、上記の設計思想 1 に抵触してしまうが、セッションデータは出版担当者が直接記述し、また、ページ番号も刻印ツールによってセッションデータに書き戻す仕様になっている。

#### 3.2.3 前付、講演プログラム目次ページ等生成ツール

セッションデータにしたがい、データベースからタイトル、著者リスト (氏名、所属) を抽

出し、論文のページ番号情報とともに講演プログラム目次 (リスト) を生成する。対応する論文ファイルへの相対パスを `\href` によってタイトルへのリンクとして埋め込む。前付に当たるデータは、自動化できる要素が低いので L<sup>A</sup>T<sub>E</sub>X ソースを直接編集し用意する枠組みとする。

著者索引にはハッシュを用いる。L<sup>A</sup>T<sub>E</sub>X の索引は、キーワードのあるページ番号しか参照できないので PDF で閲覧するには不便である。そこで著者氏名をキーとして、論文ファイルへの相対パス、ページ番号、セッション情報をハッシュに格納する。すると著者索引は箇条書き環境下において、次の簡単なコードで実現される：

```
ksort($list);
foreach ($list as $name => $ptrs){
    fwrite($F, "\\item $name $ptrs\n");
}
```

ここで `$list` はハッシュ、`$name` は著者姓であり、`$ptrs` は `\href`, `\pageref`, `\ref` を含む著者に関連した情報へのポインタである。

#### 3.2.4 アブストラクト集生成ツール

著者は、投稿システムにおいて論文とは別にアブストラクトを登録できる。このアブストラクトを前付、講演プログラム目次と合わせて組版するツールを提供する。アブストラクト集は、フルボリュームの会議録に較べコンパクトに作成できるので、印刷して参加者に配布し便宜を図ることができる。

### 3.3 その他のコンテンツの生成

ナビゲーションページは、自動生成も可能であるが、多くの場合、会議の特徴を表したデザインを背景に貼り、各コンテンツへのボタンを奇麗に配置すると見栄えがよい。よってナビゲーションページのみ Adobe Illustrator などのデザインアプリケーションなどを用いて作成するとよい。リンクボタンの処理は Acrobat Professional 等で付加することができる。

\*講演プログラムを生成するためのテーブル等は定義されている。

講演プログラム目次ページはそのままアドバンスプログラムとして公開可能である。また、必要であれば、ナビゲーションページ以下全てのファイルはオンライン公開・閲覧も可能となる。

CCM ではキーワード登録は実装されていない。別の投稿システムを運用して、キーワードが付加される場合は、著者索引と同様にキーワード索引を作ることができる。

現状では、論文本体に記載されている文献情報と、投稿システムがデータとして持つ文献情報とは独立しており、その一貫性を保障するのは投稿者自身となっている。しかし、そこにはタイポや入力ミスがつきまとう。これらのデータと論文との整合性のチェックは出版担当者の責任となるため、最終稿の手直しは試行錯誤で行えることが望ましい。本システムでは、エディタでの編集作業以外の行程は make を用いれば自動化できる。3. 節の PC 環境においては、約 10 秒で最終稿を生成できる。目次ページ PDF だけを生成する工程だけだと 1.54 秒で終了した。このように組版時間が極めて短いため、仕上がりを確認しながらの編集作業が容易であり、最終稿の品質を上げるための時間を最大限確保できる。

#### 4. 使用実績と評価

本システムを用いた国際会議を、表 1 に示す。多くの会議で本システムが運用され、継続使用されている。

本システム中で、投稿システムと編集システムは基本的には不可分であるが、編集システムの入力インタフェースをやや書き直せば分割することができる。表 1 中で (\*) が記されたケースでは、投稿システムと編集システムが別々のマシンで運営されている。CCM の場合は、投稿システム側 MySQL データベースをアーカイブし、それを編集システムに移植・復元して使用した。START V2, ePAPERS の場合は投稿システムプロバイダの方から投稿データと論文 PDF をまとめて送付してもらい、編集システムにおいて組版のみ行った。

表 1 システム使用実績

Conference	Submission system
RISP/NCSP'04	CyberChair
RISP/NCSP'05	CCM
RISP/NCSP'06	CCM
IEEE/NCN'06	CCM
RISP/NCSP'07	CCM (*)
IEEE/NDES'07	CCM
IEEE/NCN'07	CCM
IEICE/NOLTA'07	START V2 (*)
NOMA'07	CCM
RISP/NCSP'08	CCM (*)
IEICE/NOLTA'08	ePAPERS (*)
IEEE/SSJW'08	CCM
IWVCC'08	CCM
IEEE/NCN'08	CCM
RISP/NCSP'09	CCM (*) 予定
IEICE/NOLTA'09	ePAPERS (*) 予定

(\*): The submission server was managed by the other site.

ところで、CCM では査読委員確保・割り当て等を分散的に行うことが困難であるため、100 件程度以上の投稿が予想される会議の場合は、スムーズな査読プロセスを行うために商用投稿システムを利用する方が無難である。

本システムが継続的に利用されているのは、時間的・金銭的成本を抑えることができているためである。出版業務を業者に委託する場合、相当な余裕を持って締め切りやセッションデータの入稿期限が設定され、校正等にも手間がかかる。しかし、本システムでは、出版担当委員が直接編集システムを操作できる環境においては、編集をローカルに行えるため、業者にはデータを単に CD-ROM に焼き込むことだけを委託すればよく、出版に要する期間を極限まで短縮出来る<sup>†</sup>。

実際に軽減されたコストを計算することは難

<sup>†</sup>CD-ROM duplicator をローカルに持っている場合は、それを用いて複写を始める直前まで、また、認証を付けてオンライン公開・閲覧を行う場合は、公開直前まで編集が行える。

しいが、ある会議ではプログラム決定・入稿から納品まで20日ほどかかっており、その間のプログラム変更や追加は困難である。また、投稿システム、CD-ROMとアブストラクト集を合わせた全出版費用に関して、本システムを使用した場合は、全部委託した場合に比して約1/5～1/3となっている。

本システムを教育現場に用いることも興味深い応用である。実例として徳島大学大学院先端技術科学教育部システム創生工学専攻知能情報システム工学コースでは、学生の中間発表用の予稿集の収集システムとして本システムを3度稼働させた。従来は複写による予稿集を参加者全員分に配布していたため、本システムの導入により紙資源の節約に資することができた。また、実施には至らなかったが、査読プロセスを学生が体験することは教育効果は高いと考えられる。

なお、本システムの操作には、現状では若干のUNIXオペレーションが必要であるが、 $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ に関する知識をもって臨まねばならないのは、前付のソース執筆のみである。前付のコンテンツを自動で埋められるようになれば、全行程をウェブベースインタフェースでできるように設計することも将来的には可能である。

## 5. むすび

国際会議等の集會會議録を作成するための出版システムを開発した。費用面で大きなメリットを生み、出版担当者の負担を軽減している。また、會議録の読者に対してはPDFブラウザによる軽快な閲覧を可能にしている。

現在の所、技術系会議でのみ運用実績があるため、他の分野での運用に展開できるかを検討する必要がある。

## 謝辞

編集システムのうち、刻印ツールの初期バージョンは、辻 繁樹氏(JST/ERATO 合原複雑数理モデリングプロジェクト)の貢献による。ここに謝意を表す。

## 参考文献

- (1) Commence Conference System, International Association for Pattern Recognition, <http://iaprcommence.sourceforge.net/>
- (2) START: Submission Tracking And Review Toolset, <http://www.softconf.com/>の左メニュー中の“Legacy Code”から迎れるSTART V1がフリーウェアであり、START V2は商用である。
- (3) CyberChair, <http://borbala.com/cyberchair/>
- (4) ePAPERS: Web Based Submission and Review, <http://www.epapers.org/>
- (5) PDFlib, <http://pdflib.jp/>
- (6) M. Goossens et. al, The  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$  コンパニオン, アスキー出版局, 1998.
- (7) M. Goossens, S. Rahtz, 鷺谷好輝訳,  $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$ Web コンパニオン, アスキー, 2001.

(2008年11月14日原稿受付)

(2009年3月30日採録決定)

## 著者略歴



**上田 哲史** 1990年徳島大学工学部卒業, 1992年同大学院工学研究科博士前期課程修了, 同年同大学工学部助手, 1997年4月同講師, 2002年10月高度情報化基盤センター助教, 2007年4月准教授。

博士(工学)。電子情報通信学会, IEEE, 情報処理学会, 計測自動制御学会, 信号処理学会各会員。