2007－FI－88 (7)
2007／11／8

プレプリントへのタイムスタンプ付与システムの構築

山地一禎
国立情報学研究所 学術ネットワーク研究開発センター
yamaji@nii.ac.jp
片岡俊幸
国立情報学研究所 学術ネットワーク研究開発センター
kataoka@nii.ac.jp
曽根原登
国立情報学研究所 情報社会相関研究系
sonehara@nii.ac.jp
行木孝夫
北海道大学 大学院理学研究院数学部門
nami@math.sci.hokudai.ac.jp

概要

数学や物理学などの分野では，プレプリントの公開により研究成果の先取権が確保される．現在では，印刷物の出版と共に，電子ファイルをインターネット上で公開する方式が主流となっている．今後，電子ファイルのみの公開となった場合，プレプリントの存在証明と非改ざん証明が担保されることが重要になる．本研究では，ボーンディジタル時代におけるこうした問題に対して，電子署名とタイムスタンプ技術を適用し，学術情報をよりセキュアに流通させるシステムの構築を行った．

キーワード

長期署名，電子署名，タイムスタンプ，プレプリント，機関リポジトリ，アーカイブ

Time Stamping Preprint Server Environment

Kazutsuna YAMAJI
R & D Center for Academic Networks, National Institute of Informatics
yamaji@nii.ac.jp
Toshiyuki KATAOKA
R & D Center for Academic Networks, National Institute of Informatics
kataoka@nii.ac.jp
Noboru SONEHARA
Information and Society Research Division, National Institute of Informatics
sonehara@nii.ac.jp
Takao NAMIKI
Department of Mathematics, Hokkaido University
nami@math.sci.hokudai.ac.jp

Abstract

In many academic fields, exchange of preprints plays an important role to establish the priority of new research ideas and results. Distribution of preprints by electronic file via the Internet has become a primary method in addition to paper publication. Electronic preprints should be certified in terms of existence proof and tamper resistant, especially in the paperless era. We developed a secure preprint service environment by that uses an electronic signature and timestamp technique.

Keywords

Long-term Electronic Signature, Electronic Signature, Timestamp, Preprint, Institutional Repository, Archives.

1. はじめに

最近では，学術情報流通の多様化が進み，ボーンディジタルとして作成された電子文書などの成果物がインターネット上で公開されることが多くなってきた [1]．研究者が成果を公開する場合には，知的財産保護の観点からも「誰が」「何を」「いつ」公開したものであるかを共通のフォーマットで明示できることが望ましい．特に，電子ファイルは複製，改変が容易であることから，内容が改ざんされていないことを簡便に確認できることが望まれる．ただし，特に大学などからローカルに公開される電子的な成果物については，そうしたセキュリティ保護は殆ど利用されていない．今後，インターネット上での学術情報流通をより活発にし，かつ，機関としての知的財産の管理を明確にしていくためには，セキュリティにも十分留意する必要がある．

数学，物理学などの論文の投稿から査読，受理までに比較的長い期間（数年）を要する研究分野では，研究成果の速報性を高めるために投稿以前あるいは投稿中の論文をプレプリントとして機関から出版する [2, 3]．プレプリントは，主要研究機関および研究者間で交換され，学術分野における先使用権の主張する文化とも捉えることができる．これを電子的に実現したものがプレプリントサーバである．世界的には，コーネル大学で運営されている arXiv.org が有名であり，40 万レコード以上のプレプリントを保持している [4]．しかしながら，公開される電子ファイルの信頼性保証について十分な技術導入がなされていないのが現状である．プレプリントサーバが普及してきているとはいえ，紙媒体による先使用権の保証モデルを前提に運用されているとみなすこともできる．今後，紙媒体の発行を行わず，完全に電子化されたプレプリントの流通を実現するためには，電子ファイルの信頼性を保証できる基盤技術の導入が重要な鍵となる．

一方，e-文書法の制定に伴い，企業内では，財務・税務関係の法定保存文章を電子的にセキュアに保存するための環境整備が進んでいる．このとき，文章の真実性を担保するために，電子署名とタイムスタンプの付与が求められている [5]．こうした事務的な文章保存以外にも，先使用権を立証する目的で，タイムスタンプの有効利用も実施されている [6]．産業分野では，知的財産を確保するために特許出願，公開による公知化という手段がとられるが，いずれの場合にもタイムスタンプが効果を発揮する [7, 8]．こうした企業内での応用に習い，ボーンディジタル時代における学術情報発信基地は，よりセキュリティを強化させる必要がある．そこで，本研究では，北海道大学が運営する EPrintsHUMATH[2] と連携して，電子署名およびタイムスタンプ技術を導入した新しいプレプリントサーバの枠組みの構築した．

2. 長期署名

電子署名とは，電子ファイルの発信者（署名者）が特定でき，その内容の非改ざん性を保証するための行為あるいはシステムのことである．図 1 における ES（Electronic Signature）の部分がこれに該当し，電子文書，署名属性，署名値から構成される．署名値は，電子文章のハッシュ値を署名者鍵で暗号化した値であり，署名文書の受取者はこのハッシュ値を利用して非改ざん性を検証する [9]．ただし，電子署名のみでは電子文書に対し「いつ」署名が行われたのかが不明確であり，また，署名を行った本人は内容を改変し署名を再作成することが可能であることから，電子署名による「誰が」「何を」の保証に加えて，「いつ」電子ファイルが作成されたかを担保する規格が必要となる．これがタイムスタンプであり [10]，図 1 の場合では ES-T（ES with Timestamp）までのフォーマットが該当する [11]．

電子署名やタイムスタンプは，いずれもディジタル署名技術を利用しており，アルゴリズムの脆弱化や鍵漏洩への対応として，有効期限や失効機能を有している．すなわち，長期的に証明力を保持できない仕組みとなっており，研究成果のように長期にわたって存在証明と非改ざん証明の保証が必要とされる対象には適していない．これに対処するために提案されたのが長期署名である．長期署名では，署名の再検証に必要な情報として，署名時点において証明書が有効であったことを示す情報，すなわち，認証パス上の

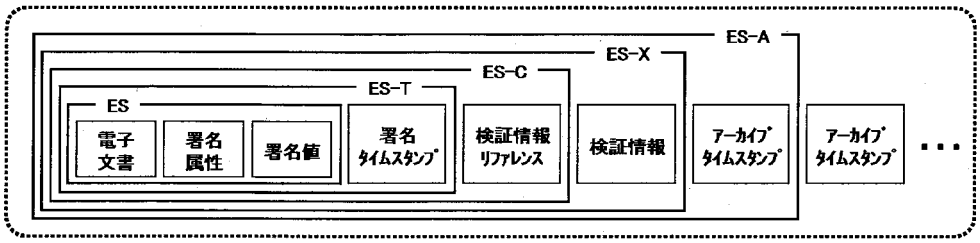

図 1：長期署名フォーマット（RFC3126）[11]

全ての証明書と失効情報を付与する（図 1 ES-X：ES with eXtended validation data）．さらに，それらの情報が改ざんされないように，長期署名フォーマット全体をタイムスタンプ（アーカイブタイムスタンプ（図 1 ES-A：ES with Archive validation data））で保護するものである．本研究では，プレプリントファイルの保護フォーマットとして，国際標準 RFC3126 に準拠した長期署名フォーマットを採用し，システムの構築を行った．

3. プレプリントサーバーへの応用

3.1 システム構成

構築したシステムの構成を図 2 に示す．電子ファイルに長期署名を付加するためには，タイムスタンプトークンを発行するための時刻認証局，電子証明書を発行・証明するための認証局，長期署名を付加するための長期署名サーバを準備する必要がある．本研究では，プレプリントの電子ファイルフォーマットとして広く利用されている PDF ファイルに長期署名を付与することとして，国際標準 RFC3126 に準拠した長期署名の署名部分とそれ以降の処理を独立して行えること，プレプリントとして 1 つのファイルで扱えるよう長期署名を PDF ファイル内に格納する方式を採用した．また，発行されるタイムスタンプの公証性を高めるため，第三者機関である商用の時刻認証局（タイムスタンプ局）を利用した．認証局には，オープンソースソフトウエアとして，認証局サーバや必要なコマンド群を提供する NAREGI-CA Version 2.2 を利用した [12, 13]．プレプリントサーバとしては，EPrints Version 3 を利用した [14]．

3.2 データおよび操作フロー

以下の手順に基づき，長期署名付きのプレプリント PDF を作成する．

1. 電子証明書の取得
2. プレプリント PDF の用意
3. プレプリント PDF に電子署名を付与
4. 3 の PDF をプレプリントサーバに登録
5. プレプリントサーバが 4 の PDF を長期署名サーバに自動送付
6. 長期署名サーバが時刻認証局からタイムスタンプトークンを取得し，長期署名付き PDF を作成

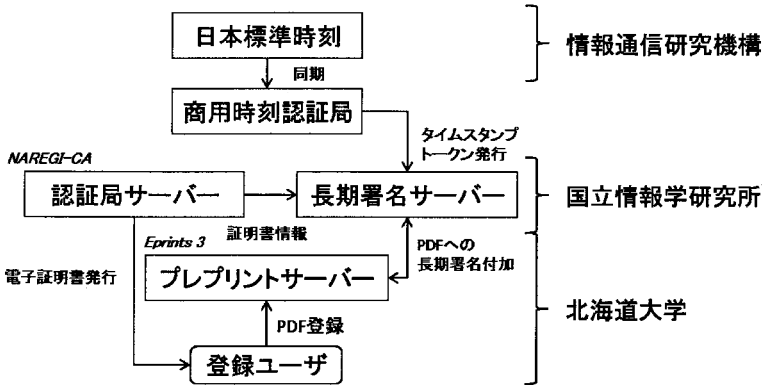

図 2：システム構成図

7. プレプリントサーバが 6 の PDF を長期署名サーバから自動取得

8. プレプリントサーバが 7 の PDF を公開

まず初めに，NAREGI-CA による認証局に対しブラウザ上から必要情報を入力し，電子証明書を取得する．この証明書を利用して，2 以降の作業を繰り返し行う．3 における電子署名の付与には，市販の PDF 生成ソフトウェアを利用した．EPrints の場合，アップロードされた PDF は既定のディレクトリに保存される．そのディレクトリ上のファイルを長期署名サーバーに送付することで，長期署名の付加を行った．PDF ファイルがプレプリントサーバに登録された後，通常は，ファイル公開可否の決定を促す電子メールがモデレータに送られる（設定により即公開も可能）．今回のシステムでは，この通知処理の前に 5〜7 の処理を行うプログラムを実行することにした．この処理は，上記通知処理を検出すると全て自動的に処理するため，PDF への署名付与の処理を除くと，長期署名の付与を行わない既存のプレプリント登録フローを全く変更しない．プレプリントサーバと長期署名サーバ間の通信には，SCP を用いた．

長期署名が付加された PDF のサンプルを図 3 左に示す．電子署名が付加されたことを明示的に示すために，サンプルでは，北海道大学大学院理学研究院のロゴを PDF 上に貼付し署名時刻と署名者を印字してある．こうした表示内容は，クライアントソフトにより，比較的自由にカスタマイズできる．また，図 3 右に示すように，長期署名の検証結果を確認することができる．

4. おわりに

現在，国立情報学研究所では，最先端学術情報基盤（Cyber Science Infrastructure：CSI）の構築を推進している [15]．本研究は，CSI における全国共同電子認証基盤構築事業（Inter University PKI：UPKI[16]）で開発された，認証局スタートアップパックを利用して NAREGI-CA を構築し [17, 18]，プレプリントサーバへの長期署名付与を実施した．NAREGI-CA は，本来，超高速コンピュータ網形成プロジェクト（National Research Grid Initiative：NAREGI）を利用するための認証局として構築されたものである．これらに加え，機関リポジトリなどの学術コンテンツ発信プロジェクトが，次世代学術情報ネットワーク（SINET3）の上で有機的に結合するのが CSI である．本研究で NAREGI-CA を用いて構築したような認証局は，将来的には各大学に用意され，UPKI を利用するための証明書発行局として利用される [19]．すなわち，タイムスタンプの利用に際しても，今後，より簡便に応用できる環境が全国レベルで整備されることが期待される．そうした発展に向けて，本研究では，

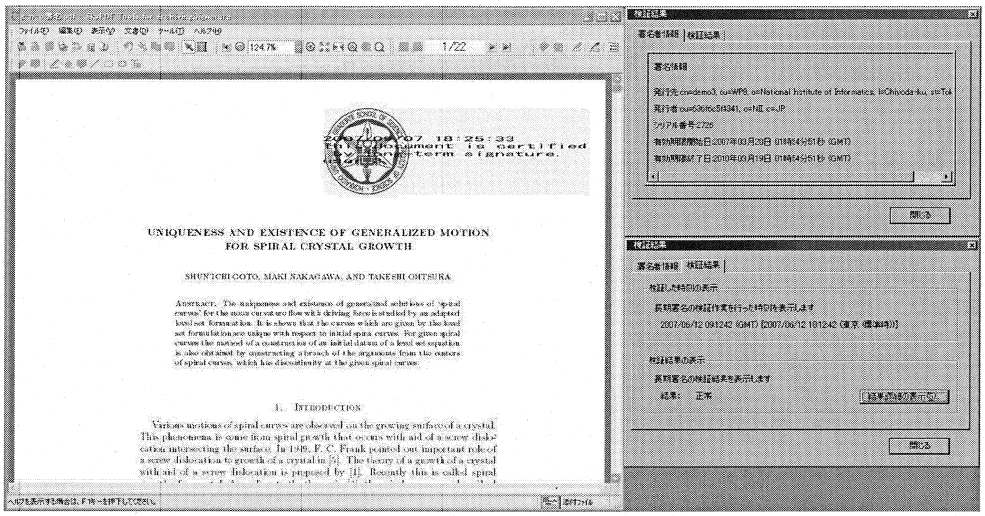

図 3：長期署名付き PDF とその検証結果

- 署名内にライセンスを明示するための仕様検討・システム改良
- 複数署名（共著者など）を可能にするための仕様検討・システム改良
- タイムスタンプサーバと連動したディジタルコンテンツ ID 取得システムの構築
- PDF 以外の電子ファイルフォーマットへの対応

などの改良に取り組み，長期署名の利便性の向上を計る予定である．こうした技術開発により，安心・安全な学術ネットワークコミュニティおよび学術コンテンツ流通が形成され，このことにより，産学連携の仕組みがより促進されるようなインフラの提供を目指したいと考えている．

参考文献

[1] 長塚　隆：インターネット上の情報資源の恒久的な保存と公開, 情報管理, 45(7), pp466-476, 2002.

[2] 行木孝夫, 畠山元彦：プレプリントサーバの構築と運営, 情報の科学と技術, 55(10), pp.434-438, 2005.

[3] 針谷喜久雄：インターネット時代のプレプリント交換, 固体物理, 31(1), pp.79-81, 1996.

[4] McKiernan, G. : arXiv.org: The Los Alamos National Laboratory E-Print Server, The International Journal on Grey Literature, 1(3), 127-138, 2000.

[5] 民間事業者等が行う書面の保存等における情報通信の技術の利用に関する法律, 内閣府 2004. (http://law.e-gov.go.jp/htmldata/H16/H16HO149.html)

[6] 池鳥裕介：先使用権制度の円滑な利用に関する調査研究, 知財研紀要, 16(11), pp.1-8, 2007.

[7] 林紘一郎：著作権の法と経済学, 勁草書房, 2004

[8] 安田　浩, 安原隆一：コンテンツ流通教科書, アスキー, 2003

[9] Housley, R. : Cryptographic Message Syntax, IETF RFC3852, 2004.

[10] Adams, C., Cain, P., Pinkas, D. and Zuccherato, R. : Internet X.509 Public Key Infrastructure Time-Stamp Protocol (TSP), IETF RFC3161, 2001.

[11] Pinkas, D., Ross, J. and Pope, N. : Electronic Signature Formats for Long Term Electronic Signatures, IETF RFC3126, 2001.

[12] Miura, K. : Overview of Japanese Science Grid Project NAREGI, Progress in Informatics, 3, pp.67-75, 2006.

[13] NAREGI-CA : http://www.naregi.org/download/index.html#capkg

[14] Millington, P. and Nixon, W. J. : EPrints 3 Pre-Launch Briefing, Ariadne, 50, 2007.

[15] Sakauchi, M., Yamada, S., Sonehara, N., Urushidani, S., Adachi, J., Konishi, K. and Matuoka, S. : Cyber Science Infrastructure Initiative for Boosting Japan's Scientific Research, CTWatch Quarterly Journal, 2(1), pp.20-26, 2006.

[16] 全国共同電子認証基盤構築事業 : https://upki-portal.nii.ac.jp/

[17] 片岡俊幸 : CSI向け認証局試用版の作成と配布, UPKIイニシアティブ発足式, WP5, 2006. (https://upki-portal.nii.ac.jp/item/idata/odatao/upkiinitiative20060830/UPKIinitiative5.pdf/view)

[18] 片岡俊幸 : これからのＰＫＩ：～ さまざまな大学内サービスの認証・認可の統合～, NIIオープンハウスCSIワークショップ, 2007. (https://upki-portal.nii.ac.jp/item/idata/odatao/CSIworkshop2007/s3appli-pki-kataoka-070608.pdf/view)

[19] 島岡政基, 谷本茂明, 片岡俊幸, 峯尾真一, 曽根原登, 寺西裕一, 飯田勝吉, 岡部寿男 : 大学間連携のための全国共同電子認証基盤UPKIにおける認証連携方式の検討, 電子情報通信学会技術研究報告, IA, 106(62), pp.13-18, 2006.