**BÁO CÁO BÀI TẬP VỀ NHÀ**

**MÔN: AN TOÀN VÀ BẢO MẬT THÔNG TIN**

**Chủ đề: Chữ ký số trong file PDF**

**Sinh viên thực hiện:** Nông Hồ Nhật

**MSSV:**K225480106094

**Lớp:**K58KTP

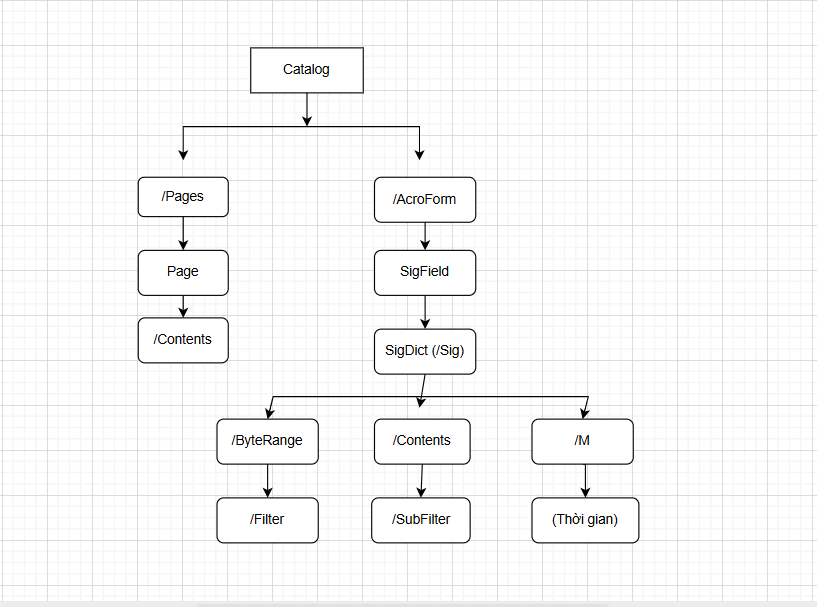
**Nội dung**: Tệp PDF dùng để minh họa quy trình ký số theo 8 bước; báo cáo tóm tắt cấu trúc PDF liên quan chữ ký, nơi lưu thời gian ký, phân tích rủi ro và biện pháp giảm thiểu.

**Nghiên cứu: Cấu trúc PDF liên quan chữ ký và thời gian ký**

**1. Cấu trúc PDF liên quan chữ ký**

Trong một tệp PDF, chữ ký số được nhúng vào thông qua các đối tượng (objects) có cấu trúc phân cấp.  
Các thành phần chính bao gồm:  
- Catalog (Root): Gốc của toàn bộ cấu trúc tài liệu PDF.  
- Pages Tree: Liên kết tới các trang của tài liệu.  
- Page Object: Đại diện cho từng trang, chứa thông tin tài nguyên và nội dung.  
- Resources: Danh sách các tài nguyên được dùng trong trang (phông chữ, hình ảnh, biểu mẫu...).  
- Content Streams: Dòng nội dung của từng trang (chứa các lệnh vẽ, text,...).  
- XObject: Đối tượng nhúng, có thể là hình ảnh hoặc biểu mẫu.  
- AcroForm: Đối tượng chứa các trường biểu mẫu (form fields), bao gồm cả trường chữ ký.  
- Signature Field (Widget): Là phần tử giao diện hiển thị vùng chữ ký.  
- Signature Dictionary (/Sig): Chứa thông tin chữ ký số (người ký, thời gian, thuật toán, chứng chỉ...).  
- /ByteRange: Xác định các byte của file được bảo vệ bởi chữ ký.  
- /Contents: Chứa giá trị chữ ký mã hóa (PKCS#7/CMS).  
- Incremental Updates: Cơ chế thêm nội dung hoặc chữ ký mới mà không ghi đè lên phần trước.  
- DSS (Document Security Store): Lưu thông tin xác minh chữ ký (chứng chỉ, CRL, OCSP, timestamp) theo chuẩn PAdES.

**Sơ đồ quan hệ giữa các đối tượng:**



**2. Thời gian ký được lưu ở đâu?**  
Trong tài liệu PDF, thông tin thời gian ký có thể được lưu ở nhiều vị trí khác nhau:  
1. Trường /M trong Signature Dictionary:  
 - Lưu thời gian ký dưới dạng chuỗi văn bản (text string).  
 - Không có giá trị pháp lý vì không được xác thực bởi bên thứ ba.  
2. Timestamp Token (RFC 3161) trong PKCS#7:  
 - Là một thuộc tính (attribute) của cấu trúc chữ ký CMS (timeStampToken).  
 - Được cung cấp bởi Time Stamping Authority (TSA), có giá trị pháp lý.  
3. Document Timestamp Object (PAdES):  
 - Là một dạng chữ ký đặc biệt áp dụng cho toàn bộ tài liệu (không gắn người ký).  
 - Được dùng để đóng dấu thời gian cho tài liệu điện tử.  
4. DSS (Document Security Store):  
 - Có thể chứa thông tin timestamp, OCSP, CRL và dữ liệu xác minh khác.

**Khác biệt giữa /M và timestamp RFC3161:**

- /M chỉ là thông tin thời gian do phần mềm ký chèn vào, không được bảo vệ bởi cơ chế xác thực.  
- Timestamp RFC3161 được tạo bởi bên thứ ba (TSA) và được ký số, đảm bảo giá trị pháp lý và xác thực thời gian.

**3) Các bước tạo và lưu chữ ký trong PDF (đã có private RSA)**

* Viết script/code thực hiện tuần tự:
* Chuẩn bị file PDF gốc : btvn2\_baomat.pdf
* Tạo Signature field (AcroForm), reserve vùng /Contents (8192 bytes).
* Xác định /ByteRange (loại trừ vùng /Contents khỏi hash).
* Tính hash (SHA-256/512) trên vùng ByteRange.
* Tạo PKCS#7/CMS detached hoặc CAdES:
* Include messageDigest, signingTime, contentType.
* Include certificate chain.
* (Tùy chọn) thêm RFC3161 timestamp token. \_ Chèn blob DER PKCS#7 vào /Contents (hex/binary) đúng offset. \_ Ghi incremental update. \_ (LTV) Cập nhật DSS với Certs, OCSPs, CRLs, VRI.
* Phải nêu rõ: hash alg, RSA padding, key size, vị trí lưu trong PKCS#7.
* Đầu ra: mã nguồn, file PDF gốc, file PDF đã ký.4) Các bước xác thực chữ ký trên PDF đã ký

**4) Các bước xác thực chữ ký trên PDF đã ký**

**Các bước kiểm tra:**

* Đọc Signature dictionary: /Contents, /ByteRange.
* Tách PKCS#7, kiểm tra định dạng.
* Tính hash và so sánh messageDigest.
* Verify signature bằng public key trong cert.
* Kiểm tra chain → root trusted CA.
* Kiểm tra OCSP/CRL.
* Kiểm tra timestamp token.
* Kiểm tra incremental update (phát hiện sửa đổi).
* Nộp kèm script verify + log kiểm thử.

**5. Rủi ro chính và biện pháp giảm thiểu**

|  |  |  |
| --- | --- | --- |
| **Rủi ro** | **Mô tả** | **Phát hiện và Biện Pháp** |
| **Thay đổi nội dung (Tampering)** | - Kẻ tấn công có thể sửa nội dung ngoài hoặc thay đổi hash, khiến trước vùng ByteRange, hoặc chỉnh sửa trực tiếp giá trị ByteRange. - Làm chữ ký không còn khớp với dữ liệu ban đầu. | - Trong quá trình xác minh (verify), so sánh hash trên ByteRange với messageDigest trong PKCS#7. - Nếu khác nhau → báo chữ ký không hợp lệ. *Biện pháp:* chỉ sử dụng incremental update đúng chuẩn PDF, trình verify phải kiểm tra ByteRange và modification level. |
| **Replay / Incremental Update Abuse** | - Kẻ xấu lợi dụng cơ chế incremental update của PDF để chèn thêm các Signature Dictionary giả hoặc che dấu sửa đổi trước đó. - Dễ khiến người dùng nhầm là tài liệu vẫn “được ký hợp lệ”. | - Yêu cầu timestamp bắt buộc từ TSA trong mỗi lần ký. - Ghi nhận và lưu toàn bộ trailer/timestamp vào DSS (Document Security Store). - Trình xác minh phải phân tích lịch sử incremental để phát hiện hành vi bất thường (ví dụ: signature xuất hiện sau cùng không liên quan đến bản gốc). |
| **Không kiểm tra thu hồi chứng chỉ (Revocation: OCSP / CRL)** | - Trường hợp chứng chỉ của người ký đã bị thu hồi nhưng hệ thống verify không kiểm tra OCSP/CRL, dẫn đến chấp nhận chữ ký không hợp lệ. | - Trong quá trình ký, nhúng OCSP responses / CRL vào DSS. - Trình xác minh cần thực hiện kiểm tra trạng thái chứng chỉ (revocation check). *Hỗ trợ LTV (Long-Term Validation)* để đảm bảo tài liệu vẫn xác minh được sau nhiều năm. |
| Lộ khóa riêng (Private Key Exposure) / Quản trị yếu | - Khóa riêng bị đánh cắp hoặc lưu trữ kém bảo mật, khiến kẻ xấu có thể ký thay cho người hợp pháp. - Thường do người dùng lưu khóa trong máy tính hoặc chia sẻ file .pfx không an toàn. | - Khóa riêng bị đánh cắp hoặc lưu trữ kém bảo mật, khiến kẻ xấu có thể ký thay cho người hợp pháp. - Thường do người dùng lưu khóa trong máy tính hoặc chia sẻ file .pfx không an toàn. |

**6. Khuyến nghị kỹ thuật**

**-** Dùng SHA-256 hoặc mạnh hơn cho message digest.

- Dùng RSA 2048+ hoặc RSA-PSS(khuyến nghị) cho chữ ký và server TSA đáng tin cậy cho timestamp RFC -3161.

- Thực hiện LTV (PAdES-LTV) bằng cách nhúng chứng thư, OCSP/CRL và timestamp token và DSS.

- Kiểm tra modification level và đảm bảo trình verify báo rõ ràng khi có incremental updates.

**7. Minh họa File đính kèm**

Trong bài nộp kèm các file mẫu :

- original.pdf -file gốc.

- signed.pdf- file sau khi đã ký ( chứa/Contents PKCS#7 và ByteRange hợp lệ).

- tampered.pdf - phiên bản đã bị chỉnh sửa ngoài vùng được ký ( dùng để minh chứng verify thất bại).

**8.Kết luận**

Bài tập này giúp hiểu rõ cơ chế lưu và xác minh chữ ký trong PDF thông qua các thành phần /ByteRange ,/Contents và incremental update. Trường /M chỉ lưu thời gian hiện thị không, không có giá trị pháp lý, trong khi timestamp RFC-3161 trong PKCS#7 mới chứng minh được thời điểm ký thực tế. Để đảm bảo tính pháp lý và xác minh lâu dài (LTV), cần kết hợp PKCS#7 + timestamp từ TSA và nhúng dữ liệu OCSP/CRL vào DSS theo chuẩn PAdES.