BÀI TẬP VỀ NHÀ – MÔN: AN TOÀN VÀ BẢO MẬT THÔNG TIN

Chủ đề: Chữ ký số trong file PDF

1) Cấu trúc PDF liên quan chữ ký (Nghiên cứu)

- Mô tả ngắn gọn: Catalog, Pages tree, Page object, Resources, Content streams, XObject, AcroForm, Signature field (widget), Signature dictionary (/Sig), /ByteRange, /Contents, incremental updates, và DSS (theo PAdES)

Thành phần	Loại Object	Vai trò
Catalog	Dictionary (Gốc)	Điểm khởi đầu của tài liệu, trỏ đến cây Trang và Form tương tác (/AcroForm).
Pages tree	Dictionary	Cấu trúc phân cấp chứa các Page Object.
Page object	Dictionary	Đại diện cho một trang. Chứa tham chiếu đến nội dung (/Contents) và các chú thích/trường (/Annots).
Resources	Dictionary	Chứa các tài nguyên dùng trên trang (Font, XObject...).
Content streams	Stream	Dòng lệnh mô tả nội dung hiển thị (văn bản, hình ảnh).
XObject	Stream/Dictionary	Đối tượng ngoại vi (ví dụ: hình ảnh) được nhúng và tham chiếu trong Content streams.
AcroForm	Dictionary	Chứa thông tin về Form tương tác, bao gồm danh sách các trường (/Fields), trong đó có trường chữ ký.
Signature field (widget)	Dictionary (Annotation)	Đại diện cho vị trí và giao diện trực quan của chữ ký trên trang. Nằm trong /Annots của Page Object. Trỏ đến dữ liệu chữ ký qua khóa /V.
Signature dictionary (/Sig)	Dictionary	Chứa dữ liệu kỹ thuật của chữ ký số. Các khóa quan trọng:
$/ByteRange	Mảng số nguyên	Xác định chính xác các vùng dữ liệu trong file đã được sử dụng để tính hàm băm (hash). Quan trọng nhất là loại trừ chính vùng chữ ký ra khỏi phép tính hash.
$/Contents	Chuỗi hex	Dữ liệu chữ ký số thực tế (thường là gói PKCS#7/CMS), bao gồm hàm băm đã được mã hóa bằng khóa riêng.
Incremental Updates	Cơ chế	Cơ chế thêm chữ ký hoặc thay đổi mà không làm hỏng tính toàn vẹn của dữ liệu đã ký. Tạo ra một phiên bản tài liệu mới (revision) bằng cách nối thêm các đối tượng mới vào cuối file.
DSS (Document	Dictionary (PAdES)	Chứa thông tin xác thực lâu dài (LTV): Các chứng chỉ liên quan (/Certs) và dữ liệu xác

Security Store)		minh trạng thái chứng chỉ (/VRI chứa OCSP/CRL).

Liệt kê object refs quan trọng và giải thích vai trò của từng object trong lưu/truy xuất chữ ký.

Thành phần	**Loại Object**	**Vai trò**
Catalog	Dictionary (Gốc)	Điểm khởi đầu của tài liệu, trỏ đến cây Trang và Form tương tác (/AcroForm).
Pages tree	Dictionary	Cấu trúc phân cấp chứa các Page Object.
Page object	Dictionary	Đại diện cho một trang. Chứa tham chiếu đến nội dung (/Contents) và các chú thích/trường (/Annots).
Resources	Dictionary	Chứa các tài nguyên dùng trên trang (Font, XObject...).
Content streams	Stream	Dòng lệnh mô tả nội dung hiển thị (văn bản, hình ảnh).
XObject	Stream/Dictionary	Đối tượng ngoại vi (ví dụ: hình ảnh) được nhúng và tham chiếu trong Content streams.
AcroForm	Dictionary	Chứa thông tin về Form tương tác, bao gồm danh sách các trường (/Fields), trong đó có trường chữ ký.
Signature field (widget)	Dictionary (Annotation)	Đại diện cho vị trí và giao diện trực quan của chữ ký trên trang. Nằm trong /Annots của Page Object. Trỏ đến dữ liệu chữ ký qua khóa /V.
Signature dictionary (/Sig)	Dictionary	Chứa dữ liệu kỹ thuật của chữ ký số. Các khóa quan trọng:
$/ByteRange	Mảng số nguyên	Xác định chính xác các vùng dữ liệu trong file đã được sử dụng để tính hàm băm (hash). Quan trọng nhất là loại trừ chính vùng chữ ký ra khỏi phép tính hash.
$/Contents	Chuỗi hex	Dữ liệu chữ ký số thực tế (thường là gói PKCS#7/CMS), bao gồm hàm băm đã được mã hóa bằng khóa riêng.

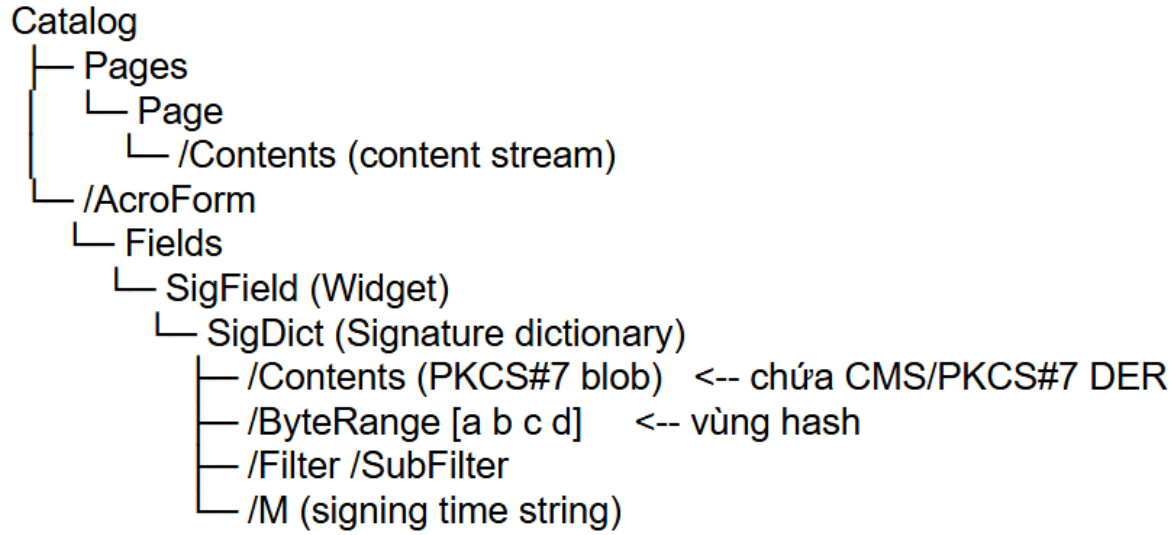
Incremental Updates	Cơ chế	Cơ chế thêm chữ ký hoặc thay đổi mà không làm hỏng tính toàn vẹn của dữ liệu đã ký. Tạo ra một phiên bản tài liệu mới (revision) bằng cách nối thêm các đối tượng mới vào cuối file.
DSS (Document Security Store)	Dictionary (PAdES)	Chứa thông tin xác thực lâu dài (LTV): Các chứng chỉ liên quan (/Certs) và dữ liệu xác minh trạng thái chứng chỉ (/VRI chứa OCSP/CRL).

Đầu ra: 1 trang tóm tắt + sơ đồ object (ví dụ: Catalog → Pages → Page → /Contents ; Catalog → /AcroForm → SigField → SigDict).

Thành phần	Mô tả	Đối tượng liên quan
Catalog	Gốc của tài liệu PDF, liên kết đến các thành phần khác như Pages và AcroForm	/Root
Pages	Danh sách các trang trong tài liệu	/Pages → /Page
Page	Một trang cụ thể chứa nội dung (text, hình, annotation, v.v.)	/Page → /Contents
Contents	Dữ liệu vẽ nội dung của trang (text, vector, hình ảnh)	/Contents
AcroForm	Form tương tác, chứa các trường (Field) trong tài liệu	/AcroForm
SigField	Trường chữ ký (signature field) trong AcroForm	/SigField

Object Ref	Vị trí/Khóa	Vai trò trong Lưu/Truy xuất Chữ ký
Catalog	/AcroForm	Điểm khởi đầu để truy xuất Form và tất cả các trường chữ ký
Signature Field	/V	Liên kết vị trí hiển thị với dữ liệu chữ ký kỹ thuật (Signature Dictionary).
Signature Dictionary	/Contents	Chứa gói dữ liệu chữ ký đã được mã hóa (PKCS#7).
Signature Dictionary	/ByteRange	Xác định dữ liệu nào của tài liệu đã được người dùng ký số

sơ đồ object



2.Thời gian ký được lưu ở đâu?

Thông tin thời gian ký có thể xuất hiện ở nhiều vị trí, nhưng chỉ những vị trí được chứng thực bởi bên thứ ba mới có giá trị pháp lý.

Vị trí Lưu trữ	Kiểu Dữ liệu / Cấu trúc	Giá trị Pháp lý
Signature Dictionary	/M (Chuỗi text, Modification Time)	Không có giá trị pháp lý cao
. PKCS#7/CMS data	Timestamp token (RFC 3161) (Attribute timeStampToken)	Có giá trị pháp lý (Bằng chứng thời gian).
Tài liệu (PAdES)	Document timestamp object	Rất cao (Chữ ký độc lập, khóa thời gian cho toàn bộ tài liệu).
DSS	Dữ liệu xác minh kèm thời gian	Hỗ trợ Xác thực Lâu dài (LTV)

/M trong Signature dictionary (dạng text, không có giá trị pháp lý).

Khóa /M không được công nhận là bằng chứng thời gian độc lập và không thể chối bỏ (non-repudiation) vì lý do sau:

- Nguồn Gốc Không Tin Cậy: Thời gian này được cung cấp trực tiếp bởi đồng hồ hệ thống (Client Clock) của máy tính người ký.
- Dễ Dàng Giả Mạo: Bất kỳ người dùng nào cũng có thể thay đổi (tua nhanh/chậm) đồng hồ hệ thống của mình trước khi ký. Do đó, thời gian ghi trong /M không có

tính xác thực và không thể chứng minh một cách độc lập rằng hành động ký đã xảy ra vào thời điểm đó.

Thông tin thời gian có giá trị pháp lý (được chứng thực) trong chữ ký số PDF luôn được lấy từ một bên thứ ba đáng tin cậy và được nhúng vào tài liệu theo các cơ chế sau:

1. Timestamp Token (RFC 3161) trong PKCS#7
   - Vị trí: Được nhúng dưới dạng một thuộc tính (Attribute) có tên timeStampToken bên trong gói dữ liệu chữ ký PKCS#7/CMS (là nội dung của khóa /Contents trong Signature Dictionary).
   - Cơ chế: Khi người dùng ký, hàm băm (hash) của dữ liệu đã ký được gửi đến một Máy chủ Dấu thời gian (TSA - Time Stamping Authority). TSA ký số lên hash đó cùng với thời gian hiện tại của máy chủ, tạo ra Timestamp Token.
   - Giá trị Pháp lý: Rất cao. Đây là bằng chứng không thể chối bỏ (non-repudiation) về việc tài liệu đã được ký trước thời điểm ghi trên dấu thời gian. Nó chứng minh rằng chứng chỉ của người ký còn hiệu lực tại thời điểm ký (dù sau đó chứng chỉ có thể hết hạn hoặc bị thu hồi).
2. Document Timestamp Object (PAdES)
   - Vị trí: Là một chữ ký số riêng biệt trong tài liệu PDF (thường không hiển thị trực quan như chữ ký người dùng). Nó được thêm vào thông qua một Incremental Update mới.
   - Cơ chế: Document Timestamp đóng vai trò là một "khóa thời gian" cho toàn bộ tài liệu (bao gồm tất cả các chữ ký và nội dung trước đó). Nó được sử dụng để bảo vệ chuỗi chữ ký khỏi nguy cơ chứng chỉ bị hết hạn hoặc thuật toán mã hóa bị phá vỡ trong tương lai.
   - Giá trị Pháp lý: Rất cao. Đây là một phần của tiêu chuẩn PAdES (PDF Advanced Electronic Signatures), đặc biệt là cho Xác thực Lâu dài (LTV - Long-Term Validation).
3. DSS (Document Security Store)
   - Vị trí: Là một Dictionary được nhúng trong tài liệu (thường là một phần của Incremental Update mới nhất).
   - Cơ chế: DSS lưu trữ các dữ liệu cần thiết để xác minh tính hợp lệ của chữ ký trong tương lai (LTV), bao gồm:
   - OCSP/CRL: Dữ liệu trạng thái thu hồi chứng chỉ (Certificate Revocation List / Online Certificate Status Protocol).
   - Các chứng chỉ liên quan: Chuỗi chứng chỉ cần thiết.

- Vai trò Thời gian: Mặc dù bản thân DSS không lưu trữ "thời gian ký" trực tiếp, nó lưu trữ dữ liệu xác minh (OCSP/CRL) có kèm theo thời gian (thời điểm mà OCSP/CRL đó được cấp), cho phép hệ thống xác minh rằng chứng chỉ của người ký là hợp lệ tại thời điểm được ghi bởi Timestamp Token.

Giải thích khác biệt giữa thông tin thời gian /M và timestamp RFC3161.

Đặc điểm	/M (Modification Time)	Timestamp Token (RFC 3161)
Nguồn Gốc	Đồng hồ hệ thống máy tính người ký (Client Clock).	Máy chủ Dấu thời gian (TSA) độc lập và đáng tin cậy.
Cơ chế	Chỉ là siêu dữ liệu (Metadata) dưới dạng chuỗi văn bản đơn giản.	Dữ liệu được ký số bởi TSA và nhúng vào gói chữ ký.
Tính Toàn vẹn	Không tin cậy. Người dùng có thể dễ dàng thay đổi đồng hồ hệ thống trước khi ký.	Tin cậy cao. Được bảo vệ bằng chữ ký số của TSA, không thể thay đổi sau khi tạo.
Giá trị Pháp lý	Không có giá trị pháp lý cao. Chỉ dùng cho mục đích thông tin.	Có giá trị pháp lý cao. Là bằng chứng thời gian không thể chối bỏ, thiết yếu cho LTV.

Đặc điểm	/M (Modification Time)	Timestamp Token (RFC 3161)

/M chỉ là thông tin tự khai báo, trong khi Timestamp RFC 3161 là bằng chứng thời gian được chứng thực bởi một Bên thứ ba Đáng tin cậy và được bảo vệ bằng chữ ký số.

Rủi ro bảo mật

Rủi ro thay đổi nội dung sau ký

Nếu ByteRange không chính xác, nội dung có thể bị chỉnh sửa mà không bị phát hiện.

Rủi ro key/certificate

Private key bị lộ sẽ phá vỡ tính xác thực.

Certificate hết hạn hoặc bị thu hồi.

Rủi ro về timestamp

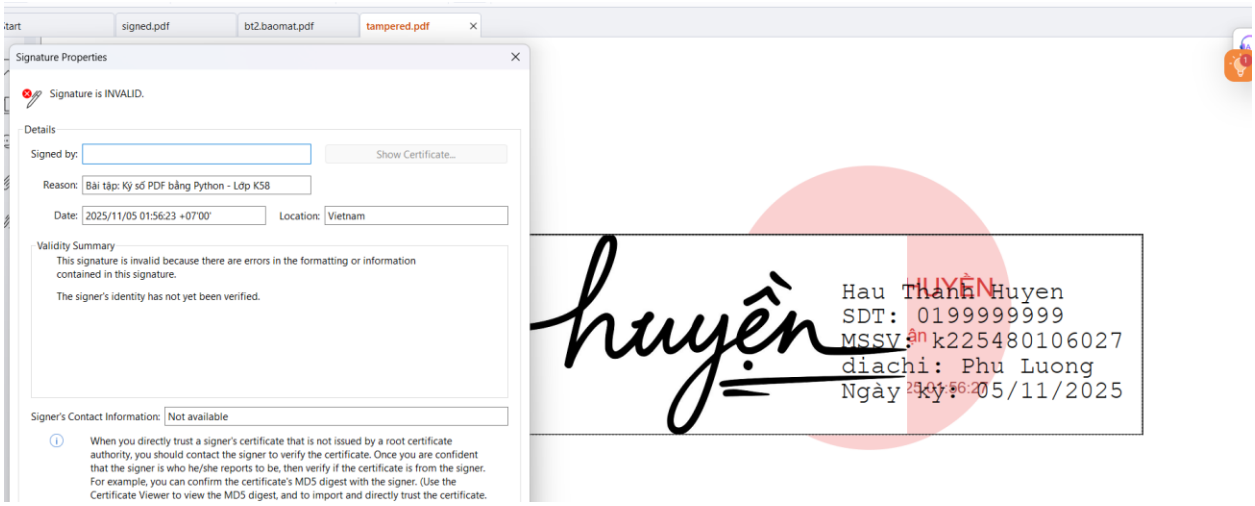
Không có TSA, thời gian ký có thể bị giả mạo.

Rủi ro về LTV

Nếu DSS không đầy đủ OCSP/CRL, chữ ký không chứng thực lâu dài.

Padding và thuật toán

Sử dụng sai RSA padding (PKCS#1 v1.5 vs PSS) có thể dẫn tới tấn công giả mạo


signed.pdf
bt2.baomat.pdf
tampered.pdf
Signature Properties
Signature is INVALID.
Details
Signed by:
Show Certificate...
Reason: Bài tập: Ký số PDF bằng Python - Lớp K58
Date: 2025/11/05 01:56:23 +07'00'
Location: Vietnam
Validity Summary
This signature is invalid because there are errors in the formatting or information contained in this signature.
The signer's identity has not yet been verified.
Signer's Contact Information: Not available
When you directly trust a signer's certificate that is not issued by a root certificate authority, you should contact the signer to verify the certificate. Once you are confident that the signer is who he/she reports to be, then verify if the certificate is from the signer. For example, you can confirm the certificate's MD5 digest with the signer. (Use the Certificate Viewer to view the MD5 digest, and to import and directly trust the certificate.
huyền
Hau Thanh Huyen
SDT: 0199999999
MSSV: k225480106027
diachi: Phu Luong
Ngày ký: 05/11/2025