Môn: An toàn và bảo mật thông tin

Bài tập 2

Sinh viên: Dương Quang Minh

Mssv: K225480106047

Nội dung:

1. Cấu trúc PDF liên quan chữ ký:

Thành phần	**Vai trò**
Catalog (Root)	Đối tượng gốc của tài liệu. Tham chiếu đến (/Pages) và (/AcroForm) chứa trường chữ ký.
Pages tree (/Pages)	Quản lý danh sách các trang (/Kids), chỉ đến từng **Page object**. Không chứa trực tiếp chữ ký.
Page object	Đại diện cho 1 trang PDF; chứa nội dung trang (/Contents) và tài nguyên hiển thị (/Resources).
Resources	Tập hợp font, hình ảnh, XObject… dùng trong trang. Không liên quan trực tiếp đến chữ ký, nhưng ảnh hưởng đến băm dữ liệu.
Content streams	Luồng lệnh vẽ và văn bản hiển thị. Nội dung này được tính vào hash trong quá trình ký.
XObject	Đối tượng con (thường là hình ảnh, form con) có thể được tái sử dụng; có thể nằm trong phần hash của dữ liệu ký.
AcroForm	Biểu mẫu chứa các trường tương tác, gồm **Signature field** (/SigField). Tham chiếu đến tất cả các field của biểu mẫu.
Signature field (widget annotation)	Ô chữ ký hiển thị trên trang PDF. Liên kết đến **Signature dictionary (/Sig)** để lưu dữ liệu chữ ký.
Signature dictionary (/Sig)	Chứa thông tin chữ ký: /Filter, /SubFilter, /ByteRange, /Contents, /Cert, /M, /Name,… Đây là lõi chứa dữ liệu ký và chứng chỉ.
/ByteRange	Mảng chỉ định các đoạn byte trong file được tính vào hash. Cho phép loại trừ vùng /Contents (vùng chứa chữ ký thực tế).
/Contents	Lưu **giá trị chữ ký số (CMS/PKCS#7)** đã mã hóa (Base16). Đây là phần duy nhất không được băm khi tính chữ ký.
Incremental updates	PDF cho phép cập nhật chữ ký mới mà không sửa file cũ. Mỗi chữ ký thêm vào tạo 1 “incremental update section” chứa cross-reference mới.
DSS (Document Security Store – PAdES)	Lưu trữ dữ liệu mở rộng theo **PAdES**: chứng chỉ, CRL, OCSP, timestamp (/VRI, /Certs, /OCSPs, /CRLs). Giúp xác thực lâu dài (LTV – Long Term Validation).

Các object refs quan trọng và vai trò:

Object Ref	**Khi nhúng chữ ký (signing)**	**Khi xác thực chữ ký (verification):**

Catalog (Root) `/Root`	Trỏ đến `/Pages` và `/AcroForm`. Dùng để thêm hoặc liên kết trường chữ ký (`SigField`) vào tài liệu.	Là điểm truy cập đầu tiên để tìm `/AcroForm` → `/Fields` → `/SigField` → `/SigDict`.
AcroForm `/Root /AcroForm`	Chứa tất cả các field của form (bao gồm chữ ký). Khi tạo chữ ký, phần mềm thêm `/SigFlags` và cập nhật `/Fields`.	Xác định vị trí chữ ký, số lượng chữ ký, và kiểm tra `/SigFlags` để biết có chữ ký.
SigFlags `/AcroForm /SigFlags`	Đặt giá trị (1 hoặc 3) để báo rằng có trường chữ ký hoặc đã có chữ ký áp dụng.	Giúp trình đọc PDF xác định rằng đây là tài liệu có ký số.
Signature Field (Widget Annotation) `/AcroForm /Fields[n]`	Tạo một field kiểu `/FT /Sig`, xác định vị trí hiển thị (rect), tên (`/T`), và trỏ đến `/V` → Signature Dictionary.	Kiểm tra xem có chữ ký hợp lệ, vị trí hiển thị, và trích xuất tham chiếu `/V` để xác thực chữ ký.
Signature Dictionary `/V` (Value of SigField)	Là **đối tượng trung tâm của quá trình ký**. Phần mềm ký chèn các khóa: • `/Filter` – tên bộ xử lý (Adobe.PPKLite, etc.)• `/SubFilter` – kiểu chữ ký (adbe.pkcs7.detached, etc.)• `/ByteRange` – các vùng được băm• `/Contents` – giá trị chữ ký CMS/PKCS#7.	Trình xác thực dùng `/ByteRange` để đọc vùng dữ liệu được ký, trích xuất `/Contents`, giải mã PKCS#7, xác minh bằng chứng chỉ trong `/Cert` hoặc chuỗi tin cậy.
/ByteRange (Array)	Khi ký, xác định các vùng dữ liệu sẽ được tính hash, **bỏ qua vùng `/Contents`**.	Khi xác thực, dùng để đọc lại đúng các vùng byte đã được ký → tính hash và so sánh với giá trị trong `/Contents`.
/Contents (Stream/HexString)	Phần mềm ký chèn chữ ký CMS (base16) vào đây — chứa hash, chứng chỉ, timestamp, thuộc tính ký.	Trình xác thực giải mã CMS, kiểm tra digest, chứng chỉ, timestamp, CRL/OCSP (nếu có).
/Cert	Có thể chứa chứng chỉ X.509 của người ký.	Dùng để xây dựng chuỗi chứng chỉ (trust chain) và xác thực khóa công khai.
/Reference	Chỉ đến các đối tượng hoặc phần dữ liệu khác (ví dụ DSS).	Nếu có, giúp kiểm tra các thông tin mở rộng, như xác minh LTV.
/M	Ghi thời điểm ký (creation date).	So sánh với timestamp hoặc kiểm tra tính hợp lệ thời gian chứng chỉ.
/Name	Tên người ký hoặc định danh chủ thể.	Hiển thị thông tin chủ thể đã ký.
/Pages Tree `/Root /Pages`	Không liên quan trực tiếp đến ký, nhưng mọi nội dung trong `/Contents` đều được tính hash trong `/ByteRange`.	Dữ liệu hiển thị bị thay đổi → sẽ làm hash khác → chữ ký không hợp lệ.

/Page Object(s)	Chứa `/Contents` (nội dung trang), `/Annots` (các widget, gồm SigField).	Xác minh rằng field chữ ký thuộc trang nào, đảm bảo không bị di chuyển/sửa.
/Contents (Page content stream)	Nội dung vẽ trang; được băm trong vùng `/ByteRange`.	So sánh lại byte data; mọi thay đổi sau khi ký làm hỏng hash → chữ ký invalid.
DSS (Document Security Store) `/Root /DSS`	Thêm sau khi ký (nếu theo PAdES). Lưu chứng chỉ, OCSP, CRL phục vụ xác minh dài hạn.	Dùng để xác thực chữ ký **offline / lâu dài (LTV)**: xác minh ngay cả khi chứng chỉ gốc đã hết hạn hoặc máy không có mạng.
Incremental Update Section	Khi ký, không sửa file cũ mà thêm 1 "update" chứa SigDict và xref mới.	Khi xác thực, parser đọc toàn bộ incremental chain để kiểm tra chữ ký trước/sau.

2. Vị trí lưu thời gian ký

Vị trí	**Thuộc object / cấu trúc**	**Dạng dữ liệu**	**Vai trò và ý nghĩa**
/M trong Signature Dictionary	/Type /Sig (Signature dictionary)	Chuỗi text ISO-8601: (D:YYYYMMDDHHmmSS+TZ)	Thời gian hiển thị trong PDF viewer (Adobe, Foxit…). Không được ký bảo vệ, chỉ để hiển thị hoặc metadata.
Timestamp Token (RFC 3161) trong PKCS#7 / CMS	Bên trong /Contents của Signature dictionary	Cấu trúc ASN.1 (timeStampToken) trong SignedAttributes	Thời gian được cấp bởi TSA (Time-Stamp Authority). Gắn kèm trong attribute id-aa-signingTime hoặc id-aa-signatureTimeStampToken. Đảm bảo thời điểm ký được chứng thực.
Document Timestamp (PAdES Document Timestamp Signature)	Là một Signature dictionary riêng biệt với /SubFilter /ETSI.RFC3161	Object /Type /Sig, /SubFilter /ETSI.RFC3161, /Contents chứa TST (RFC 3161)	Chữ ký đặc biệt chỉ chứa timestamp (không thay đổi nội dung), áp dụng cho toàn tài liệu. Dùng trong PAdES-LTV để niêm phong thời gian tài liệu.

Vị trí	Thuộc object / cấu trúc	Dạng dữ liệu	Vai trò và ý nghĩa
DSS (Document Security Store)	/DSS dictionary trong Catalog	Có thể chứa các timestamp trong /VRI → /TS hoặc timestamp tokens trong /Certs hoặc /OCSPs	Lưu trữ dữ liệu xác minh lâu dài (chứng chỉ, OCSP, CRL, timestamp). Cho phép xác minh offline sau này.

Sự khác biệt giữa thông tin thời gian /M và timestamp RFC3161:

	/M	timestamp RFC3161
Cách tạo	Phần mềm ký (signing application) tự ghi	Time Stamping Authority (TSA) cung cấp
Định dạng	Chuỗi text ISO-8601: (D:YYYYMMDDHHmmSS+TZ)	Cấu trúc ASN.1 (timeStampToken) trong SignedAttributes
Giá trị pháp lý	Không	Có
Mục đích	Hiển thị	Chứng minh thời gian ký

3. Rủi ro bảo mật

Sửa /M:

- /M là nội dung text hiển thị, không được bảo vệ.
- Hậu quả: thời gian bị làm giả dễ gây hiểu nhầm

ByteRange:

- Thay đổi ByteRange hoặc bố trí object để khiến trình đọc xác minh sai vùng được ký.
- Hậu quả: Duy trì chữ ký "hợp lệ" trên nội dung đã bị thay đổi.

Forged:

- Timestamp token có thể bị giả nếu TSA bị giả mạo hoặc giao tiếp không an toàn.
- Hậu quả: Thời điểm ký bị giả, làm sai lệch giá trị pháp lý.

Làm giả/thay đổi DSS:

- Kẻ tấn công chèn DSS/OCSP/CRL/Timestamp giả trong incremental update.

- Hậu quả: Trình xem tin dữ liệu xác minh không chính xác, dẫn đến kết luận sai về hợp lệ chữ ký.

Padding oracle:

- Padding-oracle là tấn công cho phép kẻ tấn công từ từ thu hồi dữ liệu được mã hoá hoặc ký bằng cách lợi dụng phản hồi khác nhau (error/timeout) của hệ thống khi xử lý padding không hợp lệ. Thường gắn với chế độ mã hóa có padding (ví dụ CBC + PKCS#7 padding) hoặc padding RSA PKCS#1 v1.5 (Bleichenbacher).
- Hậu quả:
    - Lộ nội dung mã hoá trong EnvelopedData.
    - Tạo được dữ liệu CMS hợp lệ giả mạo, có thể làm cho trình đọc chấp nhận dữ liệu bị sửa.
    - Trong trường hợp timestamp exchange (TSA) không cẩn thận, có thể giả mạo phản hồi TSA.

Replay attack:

- Replay là việc dùng lại chữ ký, timestamp hoặc token hợp lệ ở tài liệu A để áp lên tài liệu B (hoặc tái dùng token cũ) nhằm tạo ấn tượng hợp lệ.
- Hậu quả:
    - Giả mạo thời điểm / chữ ký trên tài liệu không đúng.
    - Gây nhầm lẫn pháp lý, lừa đảo hợp đồng.
    -

Key leak:

- Lộ khóa riêng (private key compromise), nếu khóa ký của signer hoặc khóa của TSA/CA bị lộ, kẻ tấn công có thể ký/chấp nhận mọi tài liệu và timestamp.
- Hậu quả:
    - Mất toàn bộ tính tin cậy hệ thống chữ ký: tài liệu giả mạo được xem là hợp lệ.
    - Phải thu hồi certificate, điều tra, và (thường) thông báo pháp lý.
    - Tổn thất niềm tin, pháp lý, tài chính.