BÀI TẬP VỀ NHÀ – MÔN: AN TOÀN VÀ BẢO MẬT THÔNG TIN
Chủ đề: Chữ ký số trong file
PDF Giảng viên: Đỗ Duy Cốp
Thời điểm giao: 2025-10-24 11:45
Đối tượng áp dụng: Toàn bộ sv lớp học phần 58KTPM
Hạn nộp: Sv upload tất cả lên github trước 2025-10-31 23:59:59

I. CÁC YÊU CẦU CỤ THỂ

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).
- Liệt kê object refs quan trọng và giải thích vai trò của từng object trong lưu/truy xuất chữ ký.
- Đầu ra: 1 trang tóm tắt + sơ đồ object (ví dụ: Catalog → Pages → Page → /Contents; Catalog → /AcroForm → SigField → SigDict).

Object	Mô tả	Vai trò liên quan chữ ký
Catalog (/Root)	Gốc của tài liệu PDF.	Chứa tham chiếu đến /Pages và có thể đến /AcroForm.
Pages Tree (/Pages)	Cấu trúc cây quản lý các trang.	Không trực tiếp chứa chữ ký, nhưng xác định thứ tự trang được hiển thị.
Page Object	Đại diện cho mỗi trang.	Tham chiếu đến /Resources, /Contents (nội dung trang).
Resources	Font, ảnh, XObject...	Cung cấp tài nguyên cho nội dung trang.
Content Stream	Dòng lệnh vẽ nội dung trang (text, hình…).	Không chứa chữ ký, nhưng thay đổi nội dung sẽ làm invalid chữ ký.
XObject	Object có thể tái sử dụng (ảnh, form).	Có thể chứa con dấu đồ họa (visible signature appearance).
AcroForm	Biểu mẫu tương tác PDF.	Gốc chứa các **Form Field**, bao gồm **Signature Field**.
Signature Field (Widget)	Một field trong AcroForm có kiểu /Sig.	Liên kết trực tiếp đến **Signature Dictionary**.
Signature Dictionary (/Sig)	Object chứa thông tin chữ ký.	Đây là phần **trọng tâm**, định nghĩa các khóa như: /Filter, /SubFilter, /Name, /M, /ByteRange, /Contents, ...
/ByteRange	Mảng chỉ định các đoạn byte của file được ký.	Xác định vùng dữ liệu được hash → phục vụ xác minh chữ ký.

/Contents	Chứa chữ ký (thường là CMS/PKCS#7 DER).	Là dữ liệu nhị phân kết quả ký.
Incremental Update	Cơ chế thêm nội dung mà không ghi đè file cũ.	Mỗi chữ ký mới là một incremental update mới, giúp ký nhiều lần.
DSS (Document Security Store)	Phần lưu trữ dữ liệu xác minh lâu dài (PAdES-LTV).	Lưu chứng thư, OCSP, CRL, timestamp, v.v. phục vụ xác minh lâu dài.

2) Thời gian ký được lưu ở đâu?

- Nêu tất cả vị trí có thể lưu thông tin thời gian:
  + /M trong Signature dictionary (dạng text, không có giá trị pháp lý).
  + Timestamp token (RFC 3161) trong PKCS#7 (attribute timeStampToken).
  + Document timestamp object (PAdES).
  + DSS (Document Security Store) nếu có lưu timestamp và dữ liệu xác minh. - Giải thích khác biệt giữa thông tin thời gian /M và timestamp RFC3161

Vị trí	Định dạng / Ý nghĩa	Giá trị pháp lý
/M trong Signature Dictionary	Dạng chuỗi text, ví dụ: D:20251031120000+07'00'	**Không có giá trị pháp lý**, chỉ do phần mềm ký ghi vào.
Timestamp Token (RFC 3161)	Attribute timeStampToken trong CMS/PKCS#7	**Có giá trị pháp lý**, vì được TSA (Time-Stamp Authority) ký xác nhận.
Document Timestamp (PAdES)	Một chữ ký đặc biệt dùng SubFilter=/ETSI.RFC3161	Là **chữ ký thời gian độc lập**, không gắn với người ký.
DSS (Document Security Store)	Có thể chứa timestamp và dữ liệu xác minh (OCSP/CRL)	Hỗ trợ xác minh lâu dài (LTV).

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

So sánh	/M	RFC 3161 Timestamp
Nguồn phát sinh	Phần mềm ký tự điền	Cấp bởi máy chủ TSA có chứng thư riêng
Dạng dữ liệu	Chuỗi văn bản (PDF text)	Token ASN.1 trong PKCS#7
Mức tin cậy	Không được bảo vệ, có thể sửa	Được ký số, có thể xác minh
Vai trò	Thông tin hiển thị	Dấu thời gian xác thực, có giá trị pháp lý

S

➢ **Tóm lại:**

- Chữ ký PDF nằm trong /Sig dictionary, tham chiếu từ một **signature field** trong **AcroForm**.
- Nội dung ký được xác định qua /ByteRange, dữ liệu ký là /Contents.
- Thời gian ký có thể xuất hiện ở nhiều nơi, nhưng chỉ **timestamp RFC3161 hoặc Document Timestamp (PAdES)** là hợp lệ để chứng minh thời điểm ký.
- DSS dùng để duy trì xác minh chữ ký lâu dài (LTV – Long Term Validation).

❖ Rủi ro bảo mật

Nhóm rủi ro	Mô tả chi tiết	Ảnh hưởng
1) **Chỉnh sửa nội dung sau khi ký (Incremental Update Abuse)**	PDF cho phép **incremental update**, nghĩa là thêm phần mới mà không xoá phần cũ. Kẻ tấn công có thể chèn nội dung mới (text/ảnh) sau vùng /ByteRange mà người dùng không để ý.	Làm sai lệch nội dung hiển thị mà vẫn giữ chữ ký "hợp lệ" theo phần mềm đọc PDF yếu.
2) **Invisible Signature Fields / Hidden Appearance**	Có thể tạo **form field vô hình** hoặc **xObject che phủ** nội dung thật → lừa người đọc xem nội dung giả.	Gây hiểu nhầm nội dung tài liệu đã ký, đặc biệt trong hợp đồng hoặc hóa đơn.
3) **ByteRange Manipulation**	/ByteRange xác định phần dữ liệu được hash. Nếu bị thay đổi (hoặc khai báo sai), trình xem PDF yếu có thể vẫn hiển thị "valid signature".	Có thể bị chèn mã độc hoặc nội dung giả ngoài vùng được ký.
4) **Lỗi xác thực của trình xem PDF**	Một số viewer (như bản cũ của Adobe Reader, Foxit, v.v.) chỉ kiểm tra một phần chữ ký hoặc bỏ qua lỗi nhỏ trong PKCS#7.	Có thể khiến chữ ký giả hoặc không hợp lệ vẫn hiển thị "hợp lệ".
5) **Lạm dụng XObject / Appearance Stream**	Chữ ký hiển thị đồ họa (con dấu, tên, ngày...) nằm trong **appearance stream**, không phải dữ liệu ký. Kẻ tấn công có thể thay đổi hình ảnh con dấu hoặc tên mà không ảnh hưởng chữ ký.	Người xem tin vào hình ảnh con dấu "giả" mà không kiểm tra chữ ký thực.
6) **Fake Timestamp / /M Field**	Trường /M có thể bị sửa bằng tay, vì không được ký. Nếu không có timestamp RFC3161 thật, người ký có thể "lùi ngày" hoặc "giả thời gian ký".	Mất tính pháp lý về thời điểm ký.
7) **DSS và LTV giả mạo**	Nếu phần DSS chứa dữ liệu chứng thư hoặc OCSP không được xác thực đúng, có thể dẫn đến xác minh sai hoặc lưu dữ liệu giả.	Rủi ro trong xác minh lâu dài (PAdES-LTV).

MA QUOC HIEU
SDT: 0355553996
MSV: K225480106089
Ngày ký: 31/10/2025