BÀI TẬP SỐ 2 — MÔN: AN TOÀN VÀ BẢO MẬT THÔNG TIN

Sinh viên: Nguyễn Thị Xuân Phương— MSSV: K225480106054
Lớp: K58KTPM
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 (tham khảo ISO 32000-1 và PAdES).

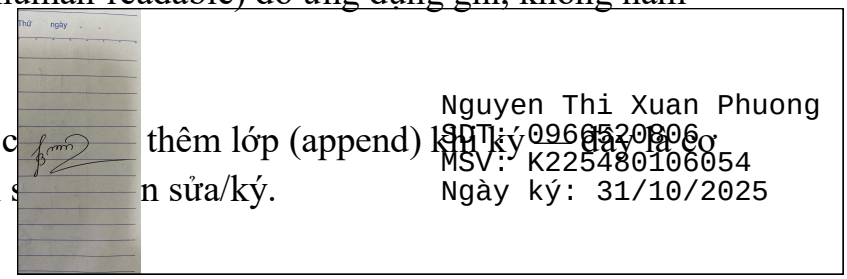
1. MỤC TIÊU BÀI TẬP

Trình bày, giải thích và minh hoạ bằng ví dụ cách thức chữ ký số được lưu và bảo vệ trong file PDF. Nhấn mạnh:
- Vị trí và cấu trúc lưu chữ ký trong PDF (AcroForm, Signature Field, Signature Dictionary).
- Cách lưu thời gian ký (khác nhau giữa /M và timestamp RFC-3161).
- Các rủi ro bảo mật phổ biến và biện pháp giảm thiểu.

2. TÓM TẮT CẤU TRÚC PDF LIÊN QUAN CHỮ KÝ

PDF lưu chữ ký như một tập hợp object liên kết: Catalog → AcroForm → Signature Field (widget) → Signature Dictionary (/Sig).

Các thành phần cần biết:
- Catalog: entry root trỏ tới AcroForm khi tài liệu có form.
- AcroForm: chứa danh sách fields (trong đó có field type = /Sig).
- Signature Field (Widget): annotation hiển thị vị trí chữ ký trên trang; trường này tham chiếu tới một Signature Dictionary.
- Signature Dictionary (/Sig): chứa /Filter, /SubFilter, /ByteRange, /Contents, /M, ...
  - /Contents chứa blob chữ ký (thường PKCS#7/CMS, hex or binary).
  - /ByteRange xác định hai đoạn byte của file được băm (phần trừ /Contents).
  - /M là chuỗi thời gian (human-readable) do ứng dụng ghi, không nằm trong vùng được ký.



Incremental update: PDF c… thêm lớp (append) khi ký … đây là cơ chế tiêu chuẩn để giữ lịch s… n sửa/ký.

3. LƯU THỜI GIAN KÝ — SO SÁNH NGẮN GỌN

/M (PDF Signature Dictionary)

- Định dạng: D:YYYYMMDDHHmmss±TZ.
- Đặc điểm: text hiển thị, không được bảo vệ bởi chính chữ ký (nằm ngoài vùng băm) → dễ bị chỉnh sửa.

Timestamp token (RFC-3161 trong PKCS#7)

- Là một thuộc tính trong gói CMS/PKCS#7 (timeStampToken), được ký bởi TSA.
- Được bảo vệ bởi chữ ký TSA → có giá trị pháp lý để chứng minh thời điểm tạo chữ ký.

Document Timestamp (PAdES)

- Là signature độc lập để bảo vệ toàn bộ trạng thái tài liệu tại một thời điểm; hữu ích cho LTV (long-term validation).

Kết luận: /M chỉ phục vụ hiển thị; để chứng minh thời điểm ký thực tế phải dùng timestamp RFC-3161 và/hoặc lưu timestamp vào DSS theo PAdES.

4. QUY TRÌNH KÝ PDF (TÓM TẮT 8 BƯỚC)

1. Chuẩn bị file gốc (original.pdf).
2. Tạo Signature Field (widget) trên trang, dành chỗ cho /Contents (reserve).
3. Xác định và ghi placeholder cho /ByteRange.
4. Tính toán hash (ví dụ SHA-256) trên hai đoạn byte được /ByteRange chỉ định.
5. Tạo PKCS#7/CMS detached chứa messageDigest và chuỗi chứng thư (và tùy chọn timeStampToken từ TSA).
6. Chèn blob PKCS#7 vào /Contents (ghi đè placeholder hoặc append incremental update).
7. Cập nhật /ByteRange chính xác (offsets) và ghi incremental update (thêm xref/trailer).
8. (Tùy chọn) Nhúng dữ liệu LTV (OCSP/CRL, certs, timestamp token) vào DSS để hỗ trợ xác minh sau này.

5. RỦI RO CHÍNH VÀ BIỆN PHÁP GIẢM THIỂU

Rủi ro 1 - Thay đổi nội dung (tampering)

- Mô tả: sửa content trước/ngoài vùng được ký hoặc sửa ByteRange.
- Phát hiện: verify sẽ so sánh hash trên ByteRange với messageDigest trong PKCS#7 và báo invalid nếu khác.
- Biện pháp: luôn dùng incremental update đúng chuẩn; trình verify phải kiểm tra ByteRange và modification level.

Rủi ro 2 — Replay / incremental abuse

- Mô tả: lợi dụng incremental updates để thêm các SigDict giả, hoặc che dấu hành vi sửa đổi.
- Giảm rủi ro: bắt buộc timestamp từ TSA và lưu các điểm timestamp/trailer vào DSS; trình kiểm tra cần phân tích toàn bộ lịch sử incremental để xác định modification_level.

Rủi ro 3 — Không kiểm tra revocation (OCSP/CRL)

- Mô tả: signer certificate đã bị thu hồi nhưng verifier không kiểm tra revocation.
- Giảm rủi ro: nhúng OCSP responses/CRLs vào DSS và kiểm tra trong quy trình verify để hỗ trợ LTV.

Rủi ro 4 — Lộ private key / quản trị yếu

- Giảm rủi ro: dùng HSM/smartcard, quản lý truy cập, khoá riêng tư không lưu trữ công khai.

6. KHUYẾN NGHỊ KỸ THUẬT (TÓM TẮT)

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

7. MINH HỌA FILE ĐÍNH KÈM (TỆP MẪU)

Trong bài nộp kèm các file mẫu (ví dụ):
- 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 NGẮN

Bài tập này yêu cầu nắm vững cơ chế lưu chữ ký trong PDF (ByteRange, /Contents, incremental update) và phân biệt rõ trạng thái thời gian ký (/M vs timestamp RFC-3161). Để đạt tính pháp lý và khả năng xác minh lâu dài cần kết hợp PKCS#7 + timestamp từ TSA và nhúng dữ liệu revocation vào DSS theo PAdES.