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 tffợng 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. MÔ TẢ CHUNG
Sinh viên thực hiện báo cáo và thực hành: phân tích và hiện thực việc nhúng, xác thực chữ ký số trong file PDF.
Phải nêu rõ chuẩn tham chiếu (PDF 1.7 / PDF 2.0, PAdES/ETSI) và sử dụng công cụ thực thi (ví dụ iText7, OpenSSL, PyPDF, pdf-lib).

II. 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 lffu/truy xuất chữ ký.
- Đầu ra: 1 trang tóm tắt + sơ đồ object (ví dụ: Catalog → Pages → Page → /Contents ; Catalog → /AcroForm → SigField → SigDict).

2) Thời gian ký đfffợc lffu ở đâu?
- Nêu tất cả vị trí có thể lffu 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 lffu timestamp và dữ liệu xác minh.
- Giải thích khác biệt giữa thông tin thời gian /M và timestamp RFC3161.

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

4) Các bffớc xác thực chữ ký trên PDF đã ký
- Các bffớc kiểm tra:
  1. Đọc Signature dictionary: /Contents, /ByteRange.
  2. Tách PKCS#7, kiểm tra định dạng.
  3. Tính hash và so sánh messageDigest.
  4. Verify signature bằng public key trong cert.
  5. Kiểm tra chain → root trusted CA.
  6. Kiểm tra OCSP/CRL.
  7. Kiểm tra timestamp token.
  8. Kiểm tra incremental update (phát hiện sửa đổi).
- Nộp kèm script verify + log kiểm thử.

III. YÊU CẦU NỘP BÀI
1. Báo cáo PDF ≤ 6 trang: mô tả cấu trúc, thời gian ký, rủi ro bảo mật.
2. Code + README (Git repo hoặc zip).
3. Demo files: original.pdf, signed.pdf, tampered.pdf.
4. (Tuỳ chọn) Video 3–5 phút demo kết quả.

IV. TIÊU CHÍ CHẤM
- Lý thuyết & cấu trúc PDF/chữ ký: 25%
- Quy trình tạo chữ ký đúng kỹ thuật: 30%
- Xác thực đầy đủ (chain, OCSP, timestamp): 25%
- Code & demo rõ ràng: 15%
- Sáng tạo mở rộng (LTV, PAdES): 5%

V. GHI CHÚ AN TOÀN
- Vẫn lffu private key (sinh random) trong repo. Tránh dùng private key thffơng mại.
- Dùng RSA ≥ 2048-bit và SHA-256 hoặc mạnh hơn.
- Có thể dùng RSA-PSS thay cho PKCS#1 v1.5.
- Khuyến khích giải thích rủi ro: padding oracle, replay, key leak.

VI. GỢI Ý CÔNG CỤ
- OpenSSL, iText7/BouncyCastle, pypdf/PyPDF2.
- Tham khảo chuẩn PDF: ISO 32000-2 (PDF 2.0) và ETSI EN 319 142 (PAdES).
