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

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).

    - Catalog : Là đối tượng gốc của tài liệu PDF. Tham chiếu đến cây trang (/Pages) và biểu mẫu (/AcroForm).
    - Pages tree : Tổ chức toàn bộ các trang dưới dạng cây. Gốc là /Pages, mỗi node có danh sách /Kids trỏ tới các Page object.
    - Page object: Đại diện cho một trang cụ thể. Chứa các thuộc tính /Resources (font, hình ảnh, form fields) và /Contents (nội dung trang).
    - Resources : Tập hợp tài nguyên được sử dụng trong trang (font, XObject, hình ảnh…).
    - Content streams : Chứa dữ liệu hiển thị (text, hình, vector, lệnh vẽ).
    - XObject : Đối tượng đồ họa tái sử dụng (thường là hình ảnh hoặc form XObject).

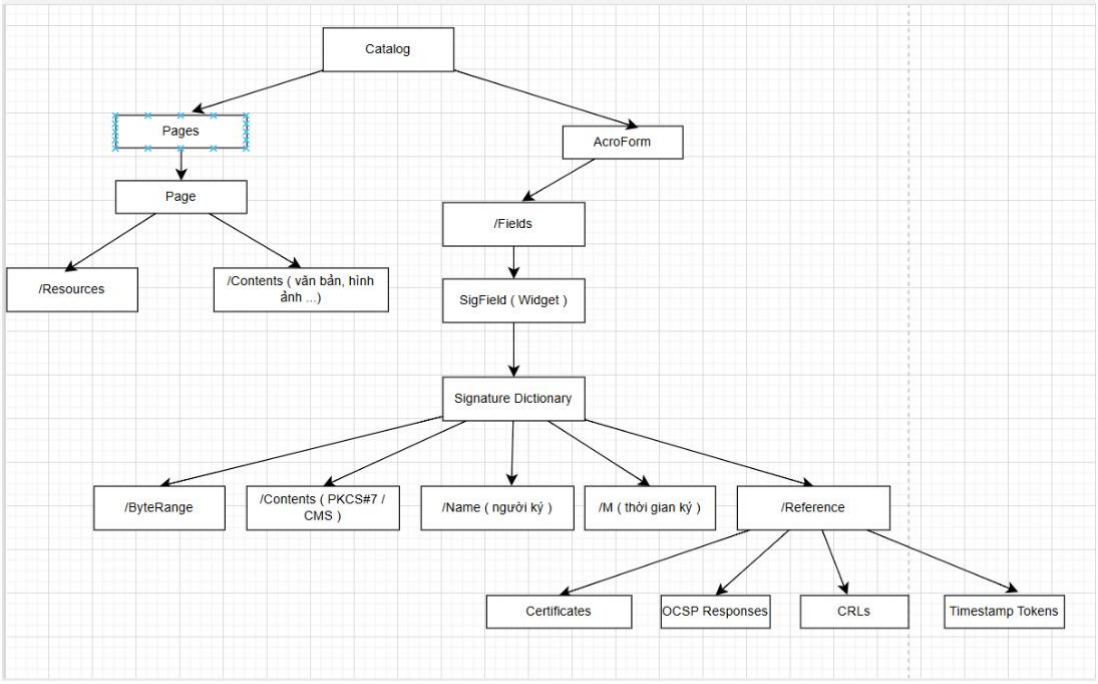
- AcroForm : Mô tả các biểu mẫu (form fields) trong PDF, gồm cả trường chữ ký.
- Signature field (widget) : Trường hiển thị vùng chữ ký trong tài liệu. Liên kết tới Signature Dictionary (/Sig) thông qua khóa /V.
- Signature dictionary (/Sig) : Chứa thông tin chi tiết về chữ ký /Name (người ký), /M (ngày ký), /Reason (lý do), /ByteRange, /Contents (dữ liệu chữ ký).
- /ByteRange : Mảng 4 số xác định các vùng byte của tệp được ký. Các vùng ngoài ByteRange bị bỏ qua (vì chứa chữ ký).
- /Contents : Dữ liệu chữ ký số (thường ở dạng CMS/PKCS#7, mã hóa Hex).
- Incremental updates : PDF được cập nhật bằng cách thêm phần mới ở cuối file (append-only), cho phép ký nhiều lần.
- DSS (theo PAdES) : Kho lưu trữ chứng thư, CRL, OCSP phục vụ xác minh chữ ký lâu dài (LTV).

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

- Root (Catalog) : Gốc của tài liệu, trỏ đến /Pages và /AcroForm.
- /Pages : Danh sách các trang.
- /Page : Trỏ đến /Contents và /Resources.
- /Contents : Luồng dữ liệu được băm và ký.
- /AcroForm : Chứa danh sách trường (Fields).
- /Fields : Tập các form field, trong đó có SigField.

- /SigField : Trường hiển thị chữ ký; trỏ đến Signature Dictionary /V.
- /V (Signature Dictionary) : Chứa dữ liệu chữ ký số thực tế.
- /ByteRange : Chỉ vùng byte được bao phủ bởi chữ ký.
- /Contents : Chứa giá trị chữ ký PKCS#7 (CMS).
- /M : Thời gian ký (dạng text, không có giá trị pháp lý).
- /DSS : Lưu thông tin xác minh lâu dài (timestamp, OCSP, CRL, cert).

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



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).
    - Dạng: chuỗi text ISO 8601, ví dụ "D:20251026 093000+07'00'".
    - Mục đích: hiển thị "ngày ký" do phần mềm ký thêm vào.
    - Không có giá trị pháp lý vì phụ thuộc vào đồng hồ hệ thống máy người ký.
- Timestamp token (RFC 3161) trong PKCS#7 (attribute timeStampToken).
    - Là chữ ký thời gian điện tử phát hành bởi Time Stamping Authority (TSA).
    - Lưu trong trường timeStampToken (thuộc nhóm "authenticatedAttributes").
    - Có giá trị pháp lý vì được ký bởi TSA, xác nhận tài liệu tồn tại tại thời điểm đó.
    - Thường nằm sâu trong nội dung Base64 hoặc DER của /Contents.
- Document timestamp object (PAdES).
    - Là một chữ ký đặc biệt (chỉ chứa timestamp, không có người ký).
    - Dạng object /Type /DocTimeStamp.
    - Được thêm vào PDF như một lần ký mới (incremental update).
    - Dùng để "đóng dấu thời gian" cho toàn tài liệu nhằm phục vụ xác minh lâu dài (LTV).

- DSS (Document Security Store) nếu có lưu timestamp và dữ liệu xác minh.
    - Thành phần tùy chọn của PAdES-LTV.
    - Có thể chứa:
    - Timestamps (bản sao token RFC3161).
    - Certs, OCSPs, CRLs dùng để kiểm tra lại tính hợp lệ chữ ký.
    - Không trực tiếp thể hiện "thời gian ký", mà lưu bằng chứng thời gian phục vụ xác minh sau này.

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

- /M :
    - Nằm trong Signature Dictionary.
    - Chỉ là chuỗi text ghi lại thời điểm ký.
    - Lấy từ đồng hồ máy tính người ký.
    - Không có giá trị pháp lý, dễ bị thay đổi.
    - Không được bảo vệ bởi chữ ký số.
    - Chỉ dùng để hiển thị trong phần mềm PDF.
- Timestamp (RFC 3161):
    - Nằm trong gói chữ ký PKCS#7 (trong /Contents).
    - Là token do TSA (Timestamp Authority) phát hành.
    - Chứa hash tài liệu + thời gian chính xác.
    - Được ký bởi TSA → có giá trị pháp lý.
    - Bảo đảm tài liệu đã tồn tại tại thời điểm đó.

3. Các bước tạo và lưu chữ ký trong PDF (đã có private RSA)

- Viết script/code thực hiện tuần tự:

  - Chuẩn bị file PDF gốc.
  - Tạo Signature field (AcroForm), reserve vùng /Contents (8192 bytes).

  - Xác định /ByteRange (loại trừ vùng /Contents khỏi hash).

  - Tính hash (SHA-256/512) trên vùng ByteRange.

  - Tạo PKCS#7/CMS detached hoặc CAdES:

    - Include messageDigest, signingTime, contentType.
    - Include certificate chain.
    - (Tùy chọn) thêm RFC3161 timestamp token.

  - Chèn blob DER PKCS#7 vào /Contents (hex/binary) đúng offset.

  - Ghi incremental update.

- (LTV) Cập nhật DSS với Certs, OCSPs, CRLs, VRI.
    - Phải nêu rõ: hash alg, RSA padding, key size, vị trí lưu trong PKCS#7.
    - Đầu ra: mã nguồn, file PDF gốc, file PDF đã ký.

    - Mã nguồn : ky_pdf_daydu.p

    - File PDF gốc : bt2.pdf

    - File PDF đã ký : final_signed.pdf

4. Các bước xác thực chữ ký trên PDF đã ký

- Các bước kiểm tra:
    - Đọc Signature dictionary: /Contents, /ByteRange.
    - Tách PKCS#7, kiểm tra định dạng.
    - Tính hash và so sánh messageDigest.
    - Verify signature bằng public key trong cert.
    - Kiểm tra chain → root trusted CA.
    - Kiểm tra OCSP/CRL.
    - Kiểm tra timestamp token.
    - Kiểm tra incremental update (phát hiện sửa đổi).
- Nộp kèm script verify + log kiểm thử.

5.