BÁO CÁO BÀI TẬP 2

AN TOÀN VÀ BẢO MẬT THÔNG TIN

Sinh viên thực hiện: Nguyễn Trung Hiếu

Mã số sinh viên: K225480106019

Lớp: K58KTP

1. Cấu trúc PDF liên quan chữ ký số

PDF là định dạng dựa trên Object với chữ ký số nhúng qua Signature Dictionary và AcroForm. Chữ ký sử dụng PKCS#7/CMS để mã hóa dữ liệu.

Theo chuẩn ISO 32000-2 / PDF 2.0, một file PDF được tạo thành từ nhiều object liên kết với nhau theo cấu trúc cây:

Object	Vai trò
Catalog (/Root)	Gốc của toàn bộ tài liệu, chứa tham chiếu đến cây trang (/Pages) và biểu mẫu (/AcroForm).
Pages tree (/Pages)	Quản lý toàn bộ các trang trong tài liệu. Mỗi trang là một object con (/Page).
Page object	Đại diện cho từng trang cụ thể, chứa tham chiếu đến /Resources, /Contents, và các chú thích (annotations) như trường chữ ký.
Resources	Tài nguyên của trang (phông chữ, hình ảnh, XObject...).
Content streams (/Contents)	Dòng lệnh vẽ nội dung trang (text, hình ảnh...).
XObject	Đối tượng đồ họa có thể tái sử dụng (ví dụ ảnh, form).

Object	Vai trò
AcroForm	Mẫu biểu mẫu điện tử chứa danh sách các trường (fields), bao gồm **trường chữ ký (Signature field)**.
Signature field (widget)	Là một loại annotation hiển thị vùng chữ ký trên trang PDF. Tham chiếu đến **Signature dictionary (/Sig)**.
Signature dictionary (/Sig)	Chứa thông tin chữ ký số: người ký, thời gian, lý do, chứng chỉ, và nội dung chữ ký PKCS#7 trong /Contents.
/ByteRange	Mảng số xác định vùng byte trong file được hash (ngoại trừ vùng /Contents chứa chữ ký).
/Contents	Lưu chữ ký số thực tế (dạng nhị phân hoặc hex PKCS#7/CMS).
Incremental update	Khi ký, PDF không ghi đè mà **thêm phần cập nhật** ở cuối file để bảo toàn tính toàn vẹn.
DSS (Document Security Store)	(Theo chuẩn **PAdES**) – chứa dữ liệu hỗ trợ xác minh lâu dài (chứng chỉ, CRL, OCSP, timestamp...).

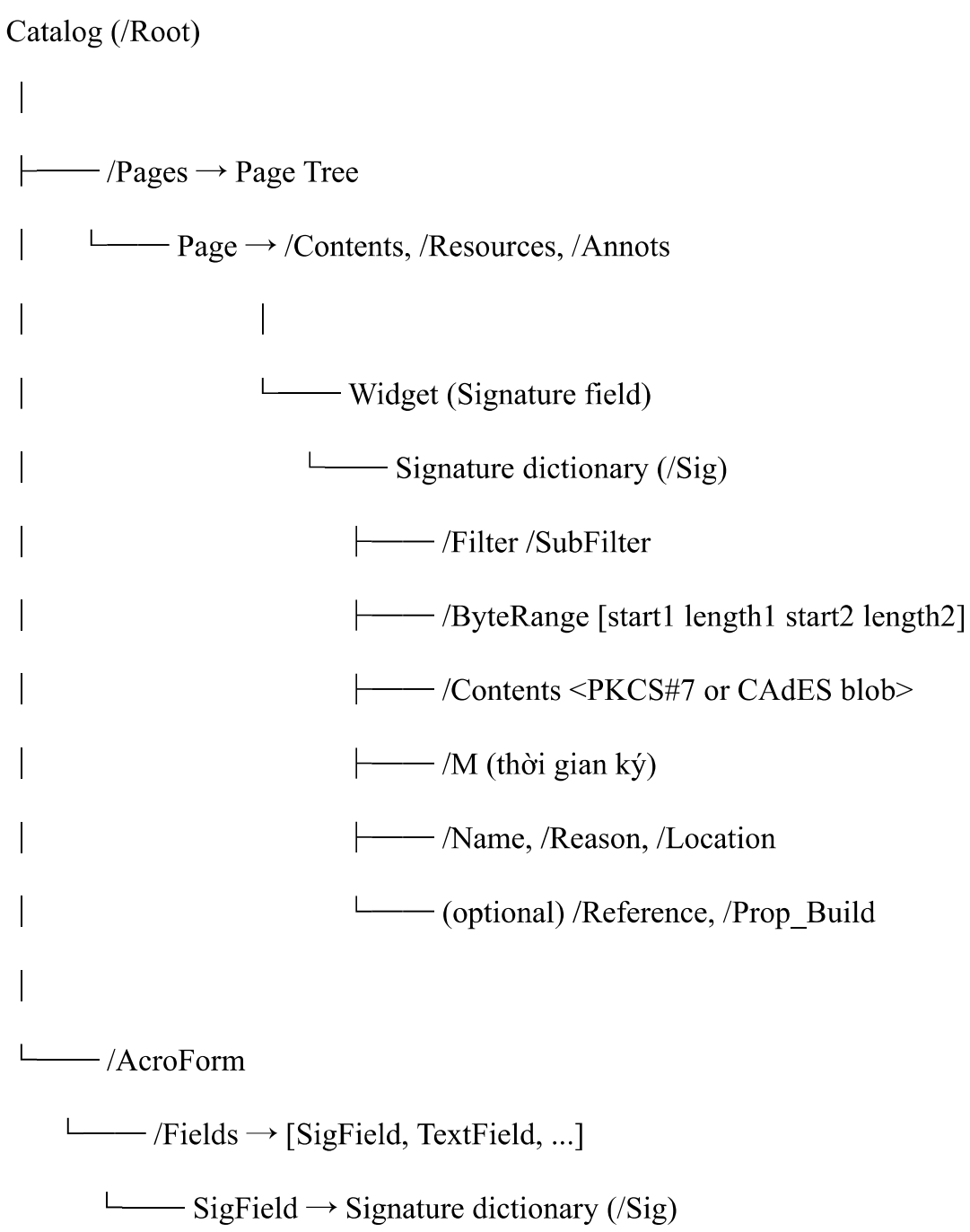
Mối quan hệ giữa ***/ByteRange*** và ***/Contents***

- /***ByteRange*** chỉ định các vùng dữ liệu được băm (hash) để tạo chữ ký
- Vùng /***Contents*** chứa chính chữ ký số PKCS#7, nằm ngoài vùng được hash.

Khi xác minh, trình đọc PDF lấy các vùng /ByteRange để tính lại hash và so sánh với giá trị trong PKCS#7.

Các thành phần /***ByteRange*** và /***Contents*** là lõi của cơ chế đảm bảo tính toàn vẹn và xác thực chữ ký.

Sơ đồ Object:



2. Thời gian ký được lưu ở đâu?

Trong file PDF có chữ ký số, thời gian ký (signing time) có thể được ghi lại ở nhiều vị trí khác nhau. Dưới đây là toàn bộ các vị trí có thể chứa thông tin thời gian cùng ý nghĩa, định dạng, và giá trị pháp lý của chúng:

- **/M trong /Sig:** đây là thông tin đơn giản nhất về thời điểm ký, không có giá trị pháp lý – người ký có thể chỉnh sửa thời gian này trước khi ký.
- **Timestamp Token (RFC 3161) trong PKCS#7/CMS:**

Cấu trúc timeStampToken theo chuẩn RFC 3161 chứa:

  - Hash của chữ ký gốc.
  - Thời gian ký chính xác (genTime).
  - Chữ ký của TSA chứng nhận thời điểm đó

Đây là bằng chứng có giá trị pháp lý chứng minh rằng chữ ký đã tồn tại tại thời điểm được TSA xác nhận. Dữ liệu này không thể bị giả mạo nếu TSA hợp lệ.

- **Document Timestamp Signature (DTS) – (PAdES Document Timestamp):**

Thay vì gắn timestamp vào chữ ký, PDF chứa một chữ ký riêng chỉ để đánh dấu thời điểm toàn bộ tài liệu (bao gồm chữ ký khác) được "niêm phong".

Loại này dùng /SubFilter /ETSI.RFC3161, lưu trong một Signature dictionary riêng. Giá trị pháp lý rất cao, vì timestamp này xác nhận toàn bộ file PDF tại thời điểm chèn dấu thời gian.

- **DSS (Document Security Store):**

Nếu tài liệu có timestamp (Document timestamp hoặc RFC3161 token), DSS có thể lưu bản sao của timestamp đó để xác minh lâu dài ngay cả khi TSA hoặc CA hết hạn. Là nền tảng của PAdES-LTV (Long-Term Validation) – đảm bảo chữ ký có thể xác thực được nhiều năm sau.

So sánh giữa /M và RFC3161 Timestamp:

Tiêu chí	**/M**	**Timestamp (RFC 3161)**
Vị trí lưu	/Sig dictionary	PKCS#7/CMS (unauthenticated attribute)
Dạng dữ liệu	Văn bản (text string)	Token nhị phân (DER, chứa chữ ký TSA)
Ai cấp	Phần mềm ký (local)	Time Stamp Authority (TSA)

Tiêu chí	/M	Timestamp (RFC 3161)
Có chứng thực không?	❌ Không	✅ Có
Có giá trị pháp lý?	❌ Không	✅ Có
Dùng khi nào?	Ghi chú thời gian ký	Chứng minh thời điểm ký thực sự

3. Rủi ro bảo mật

Chữ ký PDF sẽ an toàn nếu tuân thủ PAdES, nhưng sẽ dễ gặp nguy cơ tấn công nếu cert hoặc tamper yếu.

Một số rủi ro chính:

- **Tamper nội dung (/Contents hoặc /ByteRange):** thay đổi text/ hình phá hash, dẫn đến invalid. Rủi ro giả mạo tài liệu, phát hiện qua verify (intergrity check).
- **Replay attack:** ký lại SigDict với Timestamp cũ. Rủi ro chối bỏ thời gian, giảm bằng RFC3161/DSS.
- **Cert Revocation (CRL/OCSP):** cert hết hạn hoặc thu hồi không kiểm tra. Rủi ro ký giả mạo
- **Incremental updates lạm dụng**: ký nhiều lớp, lớp sau che lớp trước. Rủi ro ẩn thay đổi.