

## **BÀI TẬP SỐ 2**

### **MÔN: AN TOÀN VÀ BẢO MẬT THÔNG TIN**

**Sinh viên: Lê Ngọc Tú – MSSV: K225480106069**

**Lớp: K58KTPM**

#### **I Giới thiệu**

Chữ ký số trong PDF (Digital Signature) là một dạng chữ ký điện tử có mã hóa, được gắn trực tiếp vào file PDF nhằm:

- Xác nhận người ký (dựa trên chứng chỉ số – Certificate).
  - Bảo đảm toàn vẹn nội dung (nếu PDF bị chỉnh sửa sau khi ký → chữ ký bị vô hiệu).
  - Ghi lại thời gian, phần mềm, và thuộc tính ký.
- 
- Cấu trúc PDF liên quan chữ ký (Nghiên cứu)

Chữ ký số trong PDF được lưu dưới dạng các object trong cấu trúc PDF và liên kết chặt chẽ thông qua Catalog → AcroForm → Signature Field → Signature Dictionary.

Các thành phần chính:

Thành phần	Vai trò
Catalog	Gốc của tài liệu, tham chiếu đến tất cả cấu trúc như Pages và AcroForm.
Pages Tree	Cây phân cấp quản lý tất cả các trang của PDF.
Page Object	Đại diện cho từng trang, liên kết tới nội dung hiển thị.
Resources	Danh sách tài nguyên (phông, hình, XObject...) dùng trong trang.
Content Streams	Chuỗi lệnh vẽ nội dung trang (văn bản, hình ảnh...).
XObject	Các đối tượng đồ họa hoặc hình ảnh tái sử dụng.
AcroForm	Biểu mẫu tương tác, chứa danh sách các trường (fields) — bao gồm trường chữ ký.
Signature Field (Widget Annotation)	Trường biểu mẫu chứa chữ ký số (vị trí ký hiển thị).
Signature Dictionary (/Sig)	Nơi lưu dữ liệu chữ ký (hash, người ký, thời gian...).
/ByteRange	Chỉ định các đoạn byte trong file được bao phủ bởi chữ ký (phần đã ký và chưa ký).

/Contents	Chứa dữ liệu chữ ký số (thường là PKCS#7/CMS dạng hex).
Incremental Updates	Cơ chế thêm phần ký mới mà không làm mất tính toàn vẹn các phần trước đó.

II Thời gian ký lưu ở đâu & khác biệt /M vs RFC3161\* \*Các vị trí có thể lưu thông tin thời gian ký:

1. **\*\*/M\*\*** trong **\*Signature dictionary\*** — kiểu chữ (date string, PDF date). Là metadata, không có tính chứng thực cryptographic (như plain text do signer đặt). \* Ví dụ: /M (D:20251030...).
2. **\*timeStampToken (RFC 3161)\*** nằm trong **\*PKCS#7 SignedAttributes\*** (cụ thể timeStampToken attribute): đây là timestamp được chữ ký bởi TSA (Time Stamping Authority). Có tính pháp lý và chống giả mạo (TSA ký trên digest của signature).
3. **\*Document Timestamp object\*** (theo PAdES): một kiểu signature/timestamp đặc biệt có mục đích timestamp toàn bộ document; khác với signature của signer.
4. **\*DSS (Document Security Store)\***: có thể lưu các timestamp token (RFC3161) kèm chứng cứ xác minh (OCSP/CRL, cert chain) để hỗ trợ LTV. **\*\*Khác biệt /M vs RFC3161 timestamp:\*\*** \* /M: chỉ là **\*metadata text\*** của PDF; signer có thể tự sửa; không có chữ ký độc lập xác nhận thời điểm. \* **\*RFC3161 timeStampToken\***: là một token **\*được TSA ký\*** (bên thứ ba đáng tin cậy) xác nhận thời điểm; token này nằm bên trong

PKCS#7 (hoặc DSS) và \*không thể bị thay đổi\* mà không làm mất tính hợp lệ của token — do đó có tính pháp lý/cryptographic.

### III Rủi ro bảo mật trong chữ ký số PDF (PAdES)

Các rủi ro chính gồm:

#### 1. Rò rỉ hoặc lộ Private Key

- Nguy cơ

Private Key là chìa khóa bí mật duy nhất của người ký. Nếu bị rò rỉ, kẻ tấn công có thể giả mạo chữ ký y hệt bản gốc.

- Nguyên nhân phổ biến

- Lưu private key trực tiếp trên máy tính thay vì thiết bị phần cứng (HSM / USB Token).
- File .pem hoặc .pfx không có mật khẩu bảo vệ.
- Lập trình sai trong quá trình sinh hoặc nạp khóa (ví dụ hardcoded key trong mã nguồn).
- Tải file ký / chia sẻ nhầm trên mạng (Google Drive, GitHub,...).

Cách khắc phục

- Luôn lưu private key trong HSM / smart card.
- Bảo vệ khóa bằng mật khẩu mạnh + mã hóa AES256.
- Không bao giờ để private key trong code hoặc gửi qua email.
- Giới hạn quyền truy cập file .pem (chmod 600, ACL).

---

## 2. Tấn công Side-Channel (kênh kề)

### Mục tiêu

Không tấn công trực tiếp vào thuật toán RSA/ECDSA, mà khai thác rò rỉ vật lý hoặc thời gian xử lý khi phần mềm ký thực hiện phép toán.

### Cách thức

Khi thiết bị thực hiện ký (ví dụ RSA modular exponentiation), thời gian hoặc năng lượng tiêu thụ phụ thuộc vào bit của private key. Kẻ tấn công có thể:

- Đo thời gian ký nhiều lần để suy ra key (Timing Attack).
- Theo dõi điện năng tiêu thụ của HSM (Power Analysis).
- Khai thác cache hoặc branch prediction (Spectre/Meltdown).

### Phòng tránh

- Sử dụng thư viện ký có cơ chế chống side-channel (constant-time operations).
- Hạn chế chạy ký trên môi trường ảo hóa không tin cậy.
- Ưu tiên dùng HSM có chứng chỉ FIPS 140-2.

---

## 3. Incremental Update Abuse – Lạm dụng cập nhật tuần tự

### Cơ chế PDF

PDF cho phép cập nhật nội dung mà không ghi đè toàn bộ file — tức là mỗi lần sửa, nó chỉ thêm (append) dữ liệu mới vào cuối file. Phần này được gọi là Incremental Update.

### Vấn đề

Khi ký, chữ ký chỉ bảo vệ ByteRange được chỉ định. Kẻ tấn công có thể:

- Chèn thêm nội dung mới sau phần ký (ví dụ thay đổi văn bản, thêm chữ ký giả).
- Ẩn nội dung gốc bằng cách overlay layer mới.
- Giữ nguyên chữ ký → Acrobat vẫn hiển thị là “hợp lệ” nếu không kiểm tra kỹ.

### Ví dụ tấn công nổi tiếng


“Incremental Save Attack” – được công bố 2018 bởi nhóm nghiên cứu Ruhr University Bochum (Đức).  
→ Adobe sau đó đã vá và cập nhật quy trình xác minh.

### Giải pháp

- Luôn khóa hoàn toàn (Lock document) sau khi ký.
- Tránh dùng “certifying signature” không giới hạn.
- Kiểm tra cờ /ByteRange và /Contents xem có trỏ đến cuối file hay không.
- Dùng PAdES-LTV / DocMDP để chống sửa sau ký.

## IV KẾT LUẬN VỀ CHỮ KÝ SỐ TRONG FILE PDF

- Chữ ký số trong PDF là sự kết hợp giữa **mật mã học hiện đại** và **chuẩn tài liệu điện tử**, cho phép xác thực danh tính người ký, đảm bảo tính toàn vẹn và chống chối bỏ cho tài liệu. Chuẩn **PAdES (PDF Advanced Electronic Signature)** đã mở rộng định dạng chữ ký PDF truyền thống thành một cơ chế ký an toàn, có giá trị pháp lý quốc tế, tương thích với các hệ thống xác thực điện tử và hạ tầng khóa công khai (PKI).
- Trong file PDF, chữ ký số được lưu dưới dạng **trường chữ ký (Signature Field)** trong vùng **AcroForm**, chứa **Signature Dictionary** gồm dữ liệu mã hóa PKCS#7, vùng dữ liệu được ký (ByteRange) và các thông tin về người ký, thời gian, lý do ký.  
Cấu trúc này cho phép phần mềm xác minh dễ dàng kiểm tra tính hợp lệ, đồng thời hỗ trợ nhiều chữ ký trên cùng một tài liệu.
- Tuy nhiên, **bảo mật chữ ký số trong PDF không tuyệt đối**. Nếu **private key bị lộ**, hoặc file PDF bị **lợi dụng tính năng Incremental Update** hay **Object Injection**, chữ ký có thể bị giả mạo hoặc nội dung bị thay đổi mà người dùng không nhận ra. Vì vậy, việc sử dụng **hạ tầng khóa an toàn (HSM, token)**, **phần mềm ký đạt chuẩn PAdES**, và quy trình xác minh nghiêm ngặt là vô cùng cần thiết.

	<p>Lê Ngọc Tú  SDT: 0582007343  MSV: K225480106069  Ngày ký: 31/10/2025</p>
-------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------

