AN TOÀN VÀ BẢO MẬT THÔNG TIN – BÀI TẬP SỐ 2

SINH VIÊN: LƯƠNG HOÀNG VIỆT – K225480106073

1) CẤU TRÚC PDF LIÊN QUAN ĐẾN CHỮ KÝ

a) Cấu trúc cơ bản của PDF:

- **Catalog:** Đối tượng gốc, điểm khởi đầu của tệp PDF.
- **Pages tree:** Cấu trúc cây tổ chức tất cả các trang.
- **Page object:** Định nghĩa thuộc tính của một trang (kích thước, nội dung).
- **Resources:** Các tài nguyên (font, ảnh) mà trang cần để hiển thị.
- **Content streams:** Luồng lệnh chỉ thị cách vẽ nội dung lên trang.
- **XObject:** Đối tượng có thể tái sử dụng, thường là hình ảnh (Image XObject) hoặc một nhóm đồ họa (Form XObject).
- **AcroForm:** Cấu trúc chứa và quản lý tất cả các trường biểu mẫu (form fields).
- **Signature field (widget):** Hình ảnh trực quan của chữ ký bạn thấy trên trang.
- **Signature dictionary (/Sig):** Đối tượng chứa dữ liệu *thực tế* của chữ ký số (ai ký, khi nào, giá trị băm...).
- **/ByteRange:** Xác định chính xác phần nào của tệp PDF đã được ký.
- **/Contents:** Chứa dữ liệu chữ ký số (thường là chuỗi PKCS#7).
- **Incremental updates:** Kỹ thuật lưu chữ ký bằng cách *ghi thêm* vào cuối tệp mà không thay đổi nội dung gốc.
- **DSS (theo PAdES):** Kho lưu trữ (Document Security Store) chứa dữ liệu (chứng thư, OCSP) để xác thực chữ ký lâu dài (LTV).

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

Sơ đồ truy xuất cơ bản là: /Catalog → /AcroForm → Mảng /Fields → Signature Field (Widget) → /V (Value) → Signature Dictionary (/Sig)
Vai trò của từng đối tượng trong chuỗi tham chiếu đó:

1. /Catalog (Đối tượng Gốc)

- Vai trò: Là điểm khởi đầu của tài liệu.
- Liên kết (Key): /AcroForm

- Giải thích:
  - Truy xuất: Trình xem PDF (viewer) bắt đầu từ Catalog, tìm key /AcroForm để lấy tham chiếu đến đối tượng quản lý biểu mẫu của tệp.
  - Lưu: Khi lưu chữ ký (nếu đây là chữ ký đầu tiên), trình ký có thể phải thêm key /AcroForm vào Catalog nếu nó chưa tồn tại.

2. /AcroForm (Đối tượng Quản lý Biểu mẫu)

- Vai trò: Chứa và quản lý *tất cả* các trường biểu mẫu (fields) trong tài liệu.
- Liên kết (Key): /Fields
- Giải thích:
  - Truy xuất: Trình xem đọc mảng (array) /Fields từ AcroForm. Mảng này chứa các tham chiếu gián tiếp (indirect references) đến *từng* trường biểu mẫu trong PDF. Trình xem sẽ lặp qua mảng này để tìm các trường chữ ký.
  - Lưu: Khi tạo một chữ ký mới, một tham chiếu đến Signature Field (mục 3) mới được tạo sẽ được *thêm* vào cuối mảng /Fields này.

3. Signature Field (Trường Chữ ký / Widget Annotation)

- Vai trò: Đây là đối tượng định nghĩa *vị trí trực quan* của chữ ký trên trang (cái hộp bạn nhìn thấy). Về mặt kỹ thuật, nó là một loại Chú thích (Annotation) con /Widget.
- Liên kết (Key): /V (viết tắt của Value)
- Giải thích:
  - Truy xuất: Khi người dùng nhấp vào hình ảnh chữ ký trên trang, trình xem sẽ tìm đối tượng Widget này. Sau đó, nó sẽ đọc key /V để lấy tham chiếu đến đối tượng chứa *giá trị* của trường đó, chính là Signature Dictionary.
  - Lưu: Khi ký, đối tượng này được tạo ra, liên kết với một trang cụ thể, và key /V của nó được đặt để trỏ đến Signature Dictionary (mục 4) mới được tạo.

4. Signature Dictionary (Từ điển Chữ ký /Sig)

- Vai trò: Đây là đối tượng *cốt lõi* chứa dữ liệu chữ ký số thực tế (dữ liệu mật mã).
- Liên kết (Key): Không có (nó là đích đến). Nó chứa dữ liệu /Contents và /ByteRange.
- Giải thích:
  - Truy xuất: Đây là đích cuối cùng của chuỗi tham chiếu. Trình xem đọc đối tượng này (được trỏ đến bởi key /V của mục 3) để lấy dữ liệu chữ ký trong key /Contents và phạm vi byte đã ký trong key /ByteRange để tiến hành xác thực.

- Lưu: Đây là đối tượng quan trọng nhất được tạo ra trong quá trình ký. Nó được điền bằng dữ liệu chữ ký (PKCS#7) và các thông tin liên quan, sau đó được lưu vào tệp (thông qua *incremental update*).

2) VỊ TRÍ LƯU THỜI GIAN KÝ

a) Tất cả các vị trí có thể nêu thông tin thời gian

1. Thời gian do ứng dụng ký tự khai báo (Không đáng tin cậy)

Đây là thời gian được lấy từ đồng hồ hệ thống của máy tính ký, không được một bên thứ ba nào xác thực.

- **/M (trong Signature Dictionary):**
  - **Mô tả:** Là một mục (key) trong từ điển chữ ký, lưu thời gian (dạng text) mà ứng dụng *tạo ra* đối tượng chữ ký.
  - **Giá trị pháp lý:** Không có. Nó chỉ mang tính thông tin, rất dễ bị giả mạo vì nó không được ký và chỉ dựa vào đồng hồ máy khách.

2. Thời gian do người ký khai báo (Được ký, nhưng không có TSA)

Thời gian này được đưa vào *trước khi* tính toán chữ ký, vì vậy nó được bảo vệ khỏi việc thay đổi (nếu thay đổi, chữ ký sẽ mất hiệu lực). Tuy nhiên, nó vẫn dựa vào đồng hồ của máy khách.

- **Thuộc tính signingTime (trong PKCS#7):**
  - **Mô tả:** Một "thuộc tính đã xác thực" (authenticated attribute) bên trong cấu trúc PKCS#7 (CAdES). Nó được ký cùng với hash của tài liệu.
  - **Giá trị pháp lý:** Cao hơn /M vì nó được ký, nhưng vẫn chỉ chứng minh rằng *người ký* khai báo thời gian đó, chứ không chứng minh thời gian đó là *chính xác* theo một nguồn tin cậy.

3. Thời gian được xác thực bởi bên thứ ba (TSA)

Đây là các loại dấu thời gian (timestamp) đáng tin cậy nhất, được cấp bởi một Cơ quan Chứng thực Dấu thời gian (TSA - Time Stamping Authority) độc lập.

- **timeStampToken (trong PKCS#7):**
  - **Mô tả:** Đây là loại phổ biến nhất. Nó là một "thuộc tính chưa xác thực" (unauthenticated attribute) được *thêm vào* cấu trúc PKCS#7 *sau khi* chữ ký đã được tạo. Token này là một chữ ký riêng của TSA, ký lên giá trị băm (hash) của chữ ký số của người dùng.

- **Giá trị pháp lý:** Rất cao. Nó chứng minh một cách độc lập rằng chữ ký số của người dùng đã tồn tại *trước* thời điểm mà TSA cấp dấu.

❖ **Document Timestamp (DTS) (theo PAdES):**

- **Mô tả:** Đây là một loại chữ ký *riêng biệt* (một đối tượng Signature Dictionary hoàn chỉnh) trong PDF. Thay vì ký lên tài liệu, nó chỉ chứa một timeStampToken (giống mục trên) ký lên hash của toàn bộ nội dung tài liệu.
- **Giá trị pháp lý:** Rất cao. Nó được dùng để "niêm phong" toàn bộ tài liệu tại một thời điểm cụ thể, thường là trước khi có các chữ ký khác.

---

4. Dữ liệu thời gian hỗ trợ xác thực lâu dài (LTV)

Đây là các mốc thời gian nằm trong các tài liệu được thu thập để chứng minh chữ ký hợp lệ, thường được lưu trong DSS.

❖ **DSS (Document Security Store):**

- **Mô tả:** Bản thân DSS không phải là một mốc thời gian, mà là một *kho lưu trữ*. Nó chứa các bằng chứng xác thực (chứng thư, CRLs, OCSP responses) và có thể chứa các *dấu thời gian* (Validation Timestamps - VTS) riêng để "đóng dấu" lên chính các bằng chứng đó.
- **Ví dụ:** Một dấu thời gian (VTS) ký lên phản hồi OCSP để chứng minh rằng phản hồi đó đã tồn tại vào thời điểm T.

❖ **Dữ liệu xác thực (Validation Data):**

- **Mô tả:** Các dữ liệu được lưu trong DSS (hoặc đôi khi nhúng trong chính PKCS#7) đều chứa thông tin thời gian riêng của chúng.
- **Ví dụ:**
  - **Certificate:** Thời gian hiệu lực (notBefore, notAfter).
  - **OCSP Response:** Thời gian phản hồi (producedAt), thời gian cập nhật (thisUpdate, nextUpdate).
  - **CRL:** Thời gian phát hành (thisUpdate), thời gian hết hạn (nextUpdate)

b) Giải thích khác biệt giữa thông tin thời gian /M và timestamp RFC3161

Tiêu chí	/M (trong Signature Dictionary)	Timestamp (RFC 3161)
Nguồn gốc	Lấy từ đồng hồ hệ thống của máy tính người ký.	Lấy từ đồng hồ bảo mật của **Bên thứ ba tin cậy (TSA)**.
Độ tin cậy	**Rất thấp.** Dễ dàng bị giả mạo (chỉ cần đổi giờ máy tính).	**Rất cao.** Được bảo vệ bằng mật mã và cấp bởi tổ chức độc lập.
Giá trị pháp lý	**Không có.** Chỉ mang tính tham khảo, thông tin.	**Cao.** Cung cấp bằng chứng không thể chối bỏ (non-repudiation) về *thời điểm tồn tại* của chữ ký.
Cơ chế	Ứng dụng ký tự động *ghi* chuỗi thời gian vào.	Ứng dụng ký gửi hash của chữ ký đến TSA, TSA *ký* và trả về một "token" thời gian (Time-Stamp Token).
Vị trí lưu	Là một mục (key) riêng lẻ (`/M`) trực tiếp trong `Signature Dictionary` (/Sig).	Nằm *bên trong* chuỗi dữ liệu PKCS#7 (ở mục `/Contents`), dưới dạng một thuộc tính (`timeStampToken`).
Mục đích	Chỉ để tham khảo "người ký khai báo đã ký lúc nào".	Để chứng minh rằng chữ ký đã tồn tại *trước* một thời điểm cụ thể, phục vụ cho xác thực lâu dài (LTV).