Chống vi phạm tiếp sóng truyền hình trên internet bằng thủy vân số với quy mô lớn

Hứa Phước Trường, Hà Xuân Sơn, Lê Ngọc Hùng Dũng, Vũ Huy Hoàng, Nguyễn Song Thiên Long, Quản Thành Thơ
Khoa Khoa Học và Kỹ Thuật Máy Tính, Trường đại học Bách Khoa Thành phố Hồ Chí Minh
Đại học Quốc Gia Thành phố Hồ Chí Minh

e-mail: truong.huaphuoc@hcmut.edu.vn, ha.son@rmit.edu.vn, dung.lengochung@hcmut.edu.vn, hoang.vuhanks@hcmut.edu.vn, long.nguyencse2023@hcmut.edu.vn, qttho@hcmut.edu.vn

Abstract - Bảo vệ bản quyền kỹ thuật số cho nội dung phát trực tiếp nhằm ngăn chặn nạn vi phạm bản quyền và đặc biệt chống tiếp sóng bất hợp pháp là một nhu cầu mới nổi đối với bất kỳ nhà sản xuất và phát sóng nội dung truyền hình hiện đại nào, đặc biệt là các công ty truyền hình internet. Nghiên cứu này tập trung vào việc đặt thủy vân số lên vùng không gian của nội dung, thủy vân số sẽ được thiết kế để toàn vẹn trước các thiết bị ghi hình bên ngoài, thủy vân số có khả năng tự nhận diện (không yêu cầu lưu trữ phiên bản gốc hoặc ánh xạ, so sánh với dữ liệu bên ngoài nội dung). Phạm vi của nghiên cứu này tập trung vào độ chính xác, khả năng phát hiện tỷ lệ lỗi thấp và thích ứng tốt với các kỹ thuật truyền dẫn HTTP streaming trên internet (như DASH, HLS), tối ưu về chi phí vận hành, đảm bảo khả năng mở rộng số người dùng lớn tới vô hạn, và có chấp nhận tương đối cho việc làm giảm chất lượng nội dung. Nghiên cứu cung cấp kết quả là một cài đặt chạy với HTTP streaming, có thể thích ứng với các hệ thống CDN (Content Delivery Network) và có tỷ lệ nhận diện thủy vân số thành công cao thông qua phương pháp tiếp sóng bằng cách quay lại bằng thiết bị ngoại vi.

Keywords: thủy vân số; HTTP streaming; bảo vệ bản quyền; chống tiếp sóng; truyền hình internet; OTT; vùng không gian; MPEG-DASH; HLS; CDN

1 GIỚI THIỆU

Ngày nay, các nội dung quan trọng thường được bảo vệ bản quyền bằng các cơ chế mã hóa đối xứng hoặc bất đối xứng, vốn đã đặt ra rào cản cao hơn cho bất kỳ hành vi tấn công nào. Các cơ chế này thường dựa vào sự bảo vệ từ các thiết bị giải mã phần cứng như Widevine của Google, Fairplay của Apple và Playready của Microsoft. Mặc dù vậy, trong thực tế việc ngăn chặn hoàn toàn việc kẻ xấu làm rò rỉ nội dung phát trực tiếp là cực kỳ khó khăn, nhất là thông qua thiết bị thu phát lại ngoại vi. Do đó, một cơ chế có thể bảo vệ từ phía nguồn của luồng phát trực tiếp vẫn là một nhu cầu như một lớp cuối cùng cho bất kỳ phương pháp bảo vệ nào khác. Việc xử lý và chèn thủy vân số cho nội dung phát trực tiếp (Live Stream) trên phương thức HTTP khá khác so với loại nội dung video phát theo yêu cầu (Video On Demand), vì nội dung phát trực tiếp không có thời gian bắt đầu hoặc kết thúc cố định, hay nói cách khác, đó là một video không xác định thời gian. Vì vậy, một số kỹ thuật nhúng thủy vân số đòi hỏi phải so sánh giữa phiên bản đã gắn thủy vân số và phiên bản gốc có thể hoạt động bình thường trên nội dung video thông thường nhưng lại không hoạt động đối với nội dung phát trực tiếp. Một điểm đáng lưu ý nữa là việc lưu trữ video 24 giờ ở chất lượng 4Mbps để làm bản so sánh nhận diện thủy vân số sẽ làm tốn 43GB dung lượng lưu trữ, khá tốn kém khi nhu cầu bảo vệ thường kéo dài nhiều ngày, thậm chí nhiều tháng hay năm. Một số người có thể lập luận rằng chúng ta có thể sử dụng thuộc tính PTS/DTS của luồng trực tiếp để ánh xạ giữa phiên bản nhúng dấu bản quyền và phiên bản gốc, nhưng trên thực tế, có rất nhiều bộ truyền dẫn trung gian hoặc bộ chuyển tiếp có thể điều chỉnh lại hay ghi đè các thuộc tính này, hoặc việc ghép lại chủ động đơn giản từ phía người rỏ rỉ có thể làm ảnh hưởng đến việc nhận diện thủy vân số, và chi phí để làm việc này là không lớn.

ISBN: 978-604-45-2224-3

2 TỔNG QUAN

2.1 Các phương pháp trong miền không gian

Thủy vân số trong video đã được nghiên cứu rộng rãi nhằm bảo vệ nội dung đa phương tiện khỏi việc phân phối trái phép, với các kỹ thuật trải rộng trên miền không gian, miền tần số và miền nén. Trong miền không gian, Jindal et al. [1] mô tả phép sửa đổi bit ít quan trọng (least significant bit — LSB), một phương pháp kinh điển nhúng dữ liệu thủy vân vào các bit ít quan trọng của thành phần điểm ảnh. Mặc dù hiệu quả tính toán cao, các phương pháp dựa trên LSB có độ bền kém trước nén và chuyển mã, khiến chúng rất nhạy cảm với các tấn công vi phạm bản quyền vì việc tái chuyển mã để chuẩn hóa nội dung là cực kỳ phổ biến trong ngành. Cox et al. [2] giới thiệu thủy vân trải phổ (spread spectrum), phân bố năng lượng thủy vân trên các điểm ảnh để tăng độ bền trước nhiễu và nén. Tuy nhiên, cách tiếp cận này thường đòi hỏi nội dung gốc cho giai đoạn dò tìm, điều không khả thi với bối cảnh phát trực tiếp có độ dài lớn tới vô hạn.

Nhằm nâng cao độ bền và tính vô hình, Bayoudh et al. [3] đề xuất thuật toán thủy vân động "multi-sprites" sử dụng đặc trưng SURF, nhúng thông tin thủy vân vào không gian màu YUV bằng cách chỉnh sửa các bit trung gian (MIDSB) và LSB. Phương pháp này đạt độ bền cao trước tấn công câu kết (collusion) và chuyển mã, có liên quan đến khả năng phát hiện qua thiết bị ghi ngoài — một yêu cầu then chốt để chống quay màn hình hoặc ghi HDMI. Tương tự, Preda et al. [4] sử dụng trải phổ kết hợp mã sửa sai chu kỳ, nhúng thủy vân vào điểm ảnh độ chói với dư thừa không gian và thời gian. Cách tiếp cận này hỗ trợ dò mù (blind detection), phù hợp với yêu cầu tự phát hiện của nghiên cứu, và tăng khả năng chống lỗi bit, dù tính khả dụng cho HTTP streaming thời gian thực vẫn còn hạn chế. Masoumi et al. [5] đề xuất phương pháp dò mù quét khung hình thành tín hiệu một chiều và điều chế thông tin thủy vân bằng kỹ thuật trải phổ; dù hiệu quả cho dò mù, phương pháp chưa bao quát trọn vẹn đặc thù phân đoạn động của HTTP streaming. Li et al. [6] phát triển thuật toán thủy vân miền không gian chống tái nén co giãn và chuyển mã, dùng chênh lệch biểu đồ (histogram differences) để chọn khung trước điểm chuyển cảnh nhằm chống xóa khung; cách này rất phù hợp với HTTP streaming — nơi chuyển bitrate và chuyển mã là phổ biến — dù chưa đề cập trực tiếp đến khả năng phát hiện khi bị ghi ngoài.

2.2 Các phương pháp trong miền tần số

Trong miền tần số, các phương pháp dựa trên DCT/DWT thường đạt độ bền cao hơn miền không gian nhưng đi kèm chi phí tính toán lớn, không phù hợp cho xử lý thời gian thực của HTTP streaming. Li et al. [7] cho thấy năng lượng khung hình sau biến đổi DCT tập trung ở hệ số DC và dải tần thấp; nhiều nghiên cứu đã khai thác điều này để nhúng thủy vân vào các hệ số được lựa chọn, kết hợp tương quan tần số cao, mã CDMA hoặc các đặc trưng hình học để tăng độ bền [8]–[10]. Các hướng khác tận dụng đặc trưng HVS và điểm đặc trưng để định vị vùng nhúng, đồng thời sử dụng QIM, DCT giả–3D hoặc mô hình JND để cân bằng giữa tính vô hình và độ bền [11]–[18]. Tuy nhiên, phần lớn các phương pháp này yêu cầu đồng bộ chặt chẽ và không đề cập rõ ràng tới khả năng phát hiện khi video bị ghi lại bằng thiết bị ngoại vi, cũng như khó mở rộng cho luồng trực tiếp có độ dài (gần như) vô hạn và số lượng người dùng rất lớn.

2.3 Các phương pháp trong miền nén

Trong miền nén, Wang et al. [19] trình bày phương pháp thủy vân MPEG-2 nhúng vào khung "shadow" bằng biến đổi DCT, chỉ cần giải mã một phần để chống tấn công co–giãn; cách tiếp cận này tối ưu cho xử lý thời gian thực nhưng bị ràng buộc bởi cấu trúc codec MPEG-2. Sun et al. [20] mô tả phương pháp thủy vân chống tái nén, duy trì chất lượng thị giác và tỷ lệ nén — đặc biệt quan trọng với HLS trải qua nhiều bước chuyển mã. Dù các phương pháp miền nén tiết kiệm chi phí do tránh chu trình giải–mã đầy đủ, tính phụ thuộc vào codec hạn chế khả năng áp dụng cho yêu cầu "agnostic" của HLS hiện đại — vốn hỗ trợ nhiều định dạng (H.264, H.265, AV1) ở các mức chất lượng khác nhau. Hơn nữa, phần lớn phương pháp miền nén chưa giải quyết bài toán phát hiện khi nội dung đã được giải mã, hiển thị trên màn hình và bị quay lại bằng camera hoặc thiết bị ghi HDMI, trong khi đây lại là kịch bản trung tâm của bài toán chống tiếp sóng.

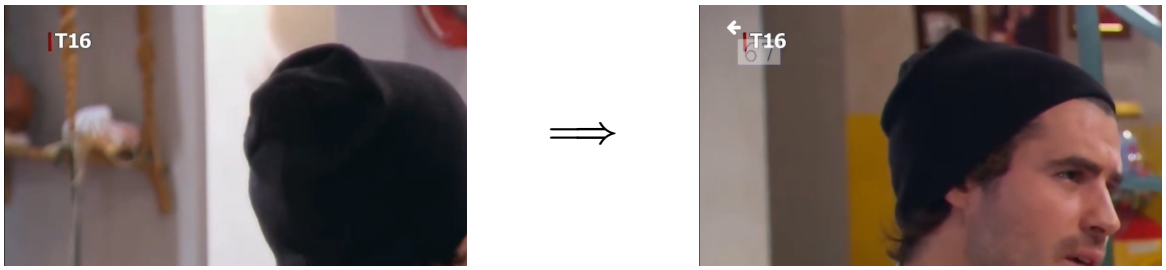
3 PHƯƠNG PHÁP ĐỀ XUẤT

3.1 Nguyên tắc Nhúng Thủy Vân Số

Kỹ thuật nhúng trong nghiên cứu này được thiết kế với mục tiêu đảm bảo tính gọn nhẹ và dễ triển khai, có thể thực hiện trên các bộ mã nguồn mở phổ biến như GPAC, ffmpeg hoặc OpenCV. Bên cạnh đó, yêu cầu về hiệu năng được đặt ra rất nghiêm ngặt: thời gian xử lý phải nhỏ hơn 16 ms cho mỗi khung hình nhằm đáp ứng nhu cầu truyền phát trực tiếp chất lượng cao với tốc độ lên tới 60 FPS. Một yêu cầu quan trọng khác là watermark phải duy trì được tính toàn vẹn ngay cả khi video bị ghi lại bởi các thiết bị quay màn hình trong điều kiện môi trường thông thường.

Nghiên cứu này sử dụng phương pháp chèn thủy vân số có thể nhận diện được bằng mắt người một cách dễ dàng, do tập trung vào giải quyết các vấn đề về khả năng mở rộng và khả năng chèn thủy vân số đáp ứng nhu cầu phát trực tiếp. Thủy vân số được chèn dưới dạng một hình vuông đè (overlay) lên trên nội dung tại các khung hình thể hiện bit 1, trong khi các khung hình thể hiện bit 0 giữ nguyên nội dung gốc. Trong cài đặt thực tế, vị trí và mức độ trong suốt (opacity) của hình vuông được thay đổi theo từng đoạn A/B và theo chuỗi bit, nhằm hạn chế các tấn công đơn giản kiểu che phủ một vùng cố định của màn hình. Để loại bỏ hoàn toàn thủy vân, kẻ tấn công buộc phải cắt hoặc che nhiều vùng khác nhau trong một khoảng thời gian đủ dài, dẫn đến suy giảm chất lượng nội dung hiển thị rõ rệt đối với người xem hợp pháp.

Thời lượng hiển thị thủy vân số có thể được điều chỉnh: ngắn để giảm khả năng nhận biết đối với người xem thông thường, hoặc dài để tăng cường tính bền vững. Trong phạm vi nghiên cứu, thời lượng đề xuất là 300 ms nhằm ưu tiên tính bền vững và đảm bảo mắt người có thể nhận thấy dễ dàng, đồng thời vẫn đảm bảo thủy vân số được duy trì khi video bị quay lại bằng camera phổ biến có tốc độ khung hình 24 FPS.

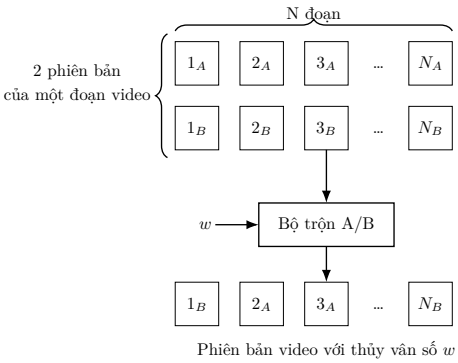


Frame hình gốc thể hiện bit 0 Frame có thủy vân số thể hiện bit 1

Hình 1: Ví dụ hai frames được chèn thủy vân số

4 Nguyên tắc đảm bảo khả năng mở rộng

Nghiên cứu của [21] đề xuất sử dụng thủy vân số hoán chuyển A/B như là một phương pháp phù hợp nhất (mặc dù Hannes và Glenn không cung cấp cài đặt và cũng không cài đặt phương pháp A/B trong nghiên cứu của mình) để thiết kế một hệ thống thủy vân số có thể tăng trưởng không giới hạn về số lượng người dùng và đặc biệt có thể thích ứng với các mô hình phát sóng trực tiếp hiện đại trên nền HTTP như MPEG-DASH hay HLS. Từ ý tưởng trên, để có thể biểu diễn một số nguyên dương (unsigned integer) có kích thước 4 bytes làm thủy vân số, một chuỗi thủy vân số A/B sẽ cần ít nhất 32 đoạn video. Với giả định mỗi đoạn dài 6 s (được đề xuất bởi Apple để cân bằng giữa độ trễ và hiệu suất), chuỗi thủy vân số sẽ tốn tối thiểu 192 giây ghi hình để có thể lấy đủ thông tin trước khi giải mã thủy vân số; khoảng thời gian này (khoảng hơn 3 phút) là có thể chấp nhận để phát hiện một nguồn



Hình 2: Phương pháp trộn thủy vân số A/B

phát sóng không mong muốn.

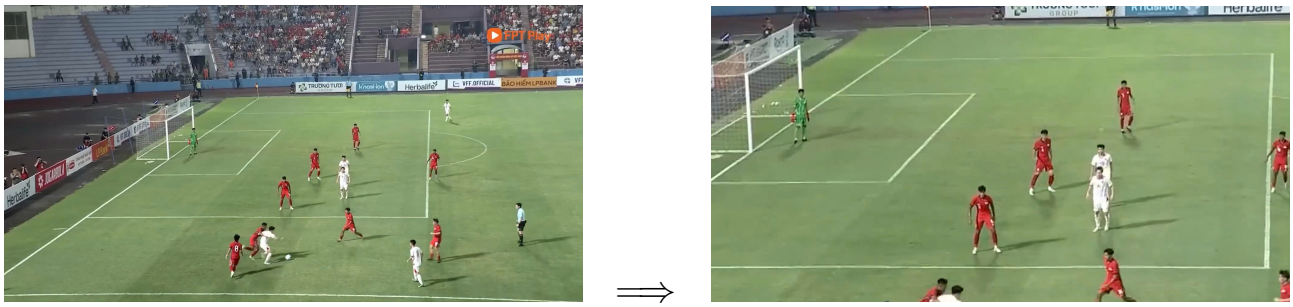
Về mặt triển khai, thủy vân A/B được trộn ở tầng biên hoặc tại các nút CDN bằng cách thao tác trên playlist/manifest (HLS, MPEG-DASH) và ánh xạ người dùng tới chuỗi lựa chọn A/B tương ứng, do đó không cần thay đổi codec hay pipeline mã hóa hiện tại. Cơ chế này hoạt động độc lập nhưng bổ sung cho các hệ thống DRM như Widevine, Fairplay hoặc Playready: DRM bảo vệ kênh truyền và nội dung trong miền mã hóa, trong khi thủy vân A/B cho phép truy vết nguồn rò rỉ ngay cả khi nội dung đã được giải mã và bị quay lại bằng thiết bị ngoại vi.

Thuật toán phát hiện được thiết kế theo hướng mù (blind), nghĩa là không cần một bản đối chiếu với video gốc như một số phương pháp khác, việc này giúp tối ưu rất nhiều chi phí khi sử dụng thủy vân số. Người giải mã thủy vân số sẽ theo dõi video được chèn mã và ghi lại thời gian xuất hiện các bit 1. Từ một chuỗi dữ liệu thời gian xuất hiện bit 1, các bit 0 sẽ được nội suy từ những khoảng thời gian trống và chu kỳ chèn thủy vân số.

5 ĐÁNH GIÁ KHẢ NĂNG CHỐNG TẤN CÔNG

Để đánh giá chặt chẽ độ bền của phương pháp thủy vân, bộ thử nghiệm bao quát các kịch bản xử lý, phân phối và thao tác ác ý có thể gặp trong thực tế. Ở đây, một tấn công (attack) là bất kỳ phép biến đổi tín hiệu nào — cố ý hay vô ý — làm suy giảm khả năng phát hiện thủy vân [22]. Các phép thử được chia thành ba nhóm chính: hình học (geometric), theo thời gian (temporal) và xử lý tín hiệu / nén (signal processing / compression) [23], [24].

Tấn công hình học thường nhắm vào việc phá đồng bộ hoá bộ dò hơn là loại bỏ hoàn toàn dấu vết, ví dụ: phóng/thu (scaling), cắt (cropping) hay biến dạng ngẫu nhiên. Khi thử nghiệm, ta kiểm tra cả scale đồng nhất và không đồng nhất, và điều chỉnh mức độ bằng hệ số tỉ lệ hoặc tỉ lệ phần trăm khung hình bị cắt [24].



(a) Trước khi tấn công hình học (b) Sau khi bị tấn công hình học

Hình 3: So sánh trực quan trước và sau tấn công hình học (Cropping)

Với video, các tấn công theo thời gian ảnh hưởng đến thứ tự khung và đồng bộ watermark — ví dụ loại bỏ khung (frame dropping) hay chuyển đổi tốc độ khung (FRC). Những phép này được mô phỏng ở nhiều mức: từ vài khung rời rạc đến loại bỏ một tỉ lệ lớn khung hình, hoặc chuyển đổi sang các frame rate khác nhau [25]. Nhóm cuối là các tấn công xử lý tín hiệu và nén. Lọc thấp (Gaussian/median/average blur) làm suy giảm thành phần tần số cao; nén mất mát (H.264, H.265) bằng thay đổi thông số CRF mô phỏng mức độ nén từ nhẹ đến nặng — đây thường là thử nghiệm quan trọng nhất vì nó phản ánh quá trình mã hoá thực tế [23], [25]. Tất cả thí nghiệm được chạy nhiều lần trong điều kiện có kiểm soát và lấy trung bình để so sánh công bằng [24]. Ngoài các tấn công trong miền số, chúng tôi còn thực hiện một tập thử nghiệm thực tế với kịch bản quay lại bằng thiết bị ngoại vi: luồng 1080p được phát trên màn hình máy tính để bàn trong điều kiện ánh sáng phòng thông thường và được ghi lại bằng camera tiêu dùng 24 fps, sau đó video quay lại được chuyển mã lại bằng H.264 trước khi tiến hành đọc thủy vân thủ công. Kịch bản này phản ánh sát hơn bài toán tiếp sóng trái phép trong thực tế.

Bảng 1 liệt kê các tham số thử nghiệm tham khảo.

ISBN: 978-604-45-2224-3

Table 1: Tham số tấn công dùng để đánh giá độ bền (tham khảo).

Loại tấn công	Mức	Tham số (ví dụ)
Nén (H.264/H.265)	Nhẹ / Trung bình / Nặng	CRF = 17 / 23 / 29
Làm mờ	Nhẹ / Trung bình / Nặng	Median kernel = 3×3 / 5×5 / 7×7
Phóng/thu (Scaling)	Nhẹ / Trung bình / Nặng	Tỷ lệ = 0.8× / 0.5× / 0.3×
Cắt (Cropping)	Nhẹ / Trung bình / Nặng	Cắt = 5% / 10% / 20%
Thao tác thời gian	Nhẹ / Nặng	Frame drop = 10%, 50%; FRC → 20/30 fps

6 KẾT QUẢ

Khả năng nhận diện thủy vân số sau khi tấn công được đánh giá bởi 3 người độc lập với khoảng 1 giờ video, kết quả được phân loại vào ba nhóm:

- Dễ dàng nhận diện: người đọc thủy vân số có thể dễ dàng nhận thấy thủy vân số và khôi phục lại được giá trị của thủy vân số chỉ trong một lần xem video.
- Khó nhận diện: người đọc thủy vân số có thể nhận diện được thủy vân số, tuy nhiên có thể phải xem đi xem lại video nhiều lần hoặc có thể phải ước đoán nhiều phiên bản thủy vân số, tuy nhiên sẽ bao gồm kết quả giải mã đúng.
- Không thể nhận diện: người đọc thủy vân số không thể nhận diện được hoặc không thể khôi phục được giá trị của thủy vân số.

Ngoài đánh giá chủ quan, chúng tôi đo thêm PSNR (Peak Signal-to-Noise Ratio) và SSIM (Structural Similarity Index Measure) để định lượng mức suy giảm chất lượng sau khi nhúng; trên nhiều video 1080p mã hóa H.264 với tổng cộng hơn 57 000 khung hình, PSNR trung bình đạt khoảng 44–46 dB và SSIM nằm trong khoảng 0.986–0.998, cho thấy chất lượng thị giác gần như không thay đổi và đáp ứng yêu cầu truyền hình internet.

Table 2: Kết quả nhận diện thủy vân số (gọn).

Tấn công	Nhẹ	Trung bình	Cao
Nén (H.264/H.265)	✓	✓	✓
Làm mờ (Median / Gaussian)	✓	✓	∼
Phóng / thu (Scaling)	✓	✓	∼
Cắt (Cropping)	✓	∼	✗
Frame drop	✓	✓	–
FPS Convert	✓	✓	–

Ghi chú: ✓ = dễ nhận diện; ∼ = khó nhận diện; ✗ = không thể nhận diện; "–" = không áp dụng / không thử.

Bài nghiên cứu chủ yếu tập trung vào khả năng xây dựng một thủy vân số có thể áp dụng trong môi trường truyền hình trực tuyến với quy mô lớn và không tập trung vào hình thức của thủy vân số. Chính vì vậy, nghiên cứu sử dụng một thủy vân đơn giản có thể nhận diện bằng con người chứ không sử dụng thủy vân có khả năng ẩn với mắt người và phù hợp cho nhận diện bằng máy nên sẽ không có phần so sánh mức độ nhận diện thành công của máy tính so với các nghiên cứu khác tập trung vào xây dựng hình thức của thủy vân số.

7 TỔNG KẾT

Nghiên cứu đã khảo sát nhiều phương pháp thủy vân số được đề xuất trong bảo vệ bản quyền nội dung số, đồng thời chỉ ra các hạn chế khiến các phương pháp đó khó có thể sử dụng trong môi trường phát sóng truyền hình, nhất là hoàn toàn không thích ứng với môi trường có số lượng người dùng lớn và mỗi người dùng phải có phiên bản thủy vân số khác nhau. Trên cơ sở đó, chúng tôi đề xuất một phương pháp chèn thủy vân số dựa trên hoán chuyển A/B, có thể triển khai trên hạ tầng HTTP streaming hiện đại, tương thích với CDN và các cơ chế DRM hiện có, đồng thời đã được cài đặt và kiểm thử trong môi trường thử nghiệm với kịch bản quay lại bằng thiết bị ngoại vi.

Kết quả thực nghiệm cho thấy thủy vân vẫn được nhận diện ổn định dưới nhiều dạng tấn công phổ biến (nén H.264/H.265, làm mờ, phóng/thu, frame drop, chuyển đổi FPS), trong khi chất lượng video được bảo toàn

ở mức gần như không suy giảm với PSNR khoảng 44–46 dB và SSIM trên 0.986. Hạn chế chính của phương pháp là thủy vân vẫn ở dạng nhìn thấy được và bộ dò hiện tại chủ yếu dựa trên quan sát thủ công; hướng phát triển tiếp theo là nghiên cứu các biến thể thủy vân bán vô hình hoặc vô hình và tích hợp các kỹ thuật nhận diện tự động để giảm phụ thuộc vào đánh giá chủ quan, từ đó tăng mức độ khả dụng trong các hệ thống sản xuất thực tế.

THAM KHẢO

[1] S. Jindal and others., "Performance analysis of lsb based watermarking for optimization of psnr and mse," International Journal of Security and Its Applications, vol. 10, pp. 345–350, 2016.

[2] I. Cox, J. Kilian, et al., "Secure spread spectrum watermarking for multimedia," IEEE Transactions on Image Processing, vol. 6, pp. 1673–1687, 1997.

[3] I. Bayoudh et al., "Online multi-sprites based video watermarking robust to collusion and transcoding attacks for emerging applications," Multimedia Tools and Applications, vol. 77, pp. 14 361–14 379, 2017.

[4] R. Preda et al., "New robust watermarking scheme for video copyright protection in the spatial domain," UPB Scientific Bulletin, vol. 73, pp. 93–104, 2011.

[5] M. Masoumi et al., "A blind spatio-temporal data hiding for video ownership verification in frequency domain," AEU-International Journal of Electronics and Communications, vol. 69, pp. 1868–1879, 2015.

[6] X. Li et al., "A robust video watermarking scheme to scalable recompression and transcoding," in International Conference on Electronics Information and Emergency Communication, 2016.

[7] J. Li et al., "A digital video watermarking algorithm based on dct domain," in 5th International Joint Conference on Computational Sciences and Optimization, 2012.

[8] G. Liu et al., "A robust digital video watermark algorithm based on dct domain," in International Conference on Computer Application and System Modeling, 2010.

[9] M. Cheng et al., "Recoverable video watermark in dct domain based on code division multiple access (cdma) modulation," Journal of Computers, vol. 8, pp. 533–538, 2013.

[10] T. Nguyen et al., "A robust blind video watermarking in dct domain using even-odd quantization technique," in International Conference on Advanced Technologies for Communications, 2015.

[11] I. Bayoudh et al., "A robust video watermarking for real-time application," in 18th International Conference on Advanced Concepts for Intelligent Vision Systems, 2017.

[12] T. Thanh et al., "Robust semi-blind video watermarking based on frame-patch matching," in AEU-International Journal of Electronics and Communications, 2014.

[13] B. Chen et al., "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," IEEE Transactions on Information Theory, vol. 47, pp. 1423–1443, 2001.

[14] C. Yang et al., "An adaptive video watermarking technique based on dct domain," in 8th International Conference on Computer and Information Technology, 2008.

[15] H. Huang et al., "A video watermarking technique based on pseudo-3d dct and quantization index modulation," IEEE Transactions on Information Forensics and Security, vol. 5, pp. 625–637, 2010.

[16] P. Campisi et al., "3d-dct video watermarking using quantization-based methods," in 15th European Signal Processing Conference, 2007.

[17] A. Cedillo-Hernandez et al., "A spatiotemporal saliency-modulated jnd profile applied to video watermarking," Journal of Visual Communication and Image Representation, vol. 52, pp. 106–117, 2018.

[18] A. Abdulfetah et al., "Robust adaptive video watermarking scheme using visual models in dwt domain," Information Technology Journal, vol. 9, pp. 1409–1414, 2010.

[19] Y. Wang et al., "Blind mpeg-2 video watermarking in dct domain robust against scaling," IEE Proceedings - Vision, Image and Signal Processing, vol. 153, pp. 581–588, 2006.

[20] J. Sun et al., "An anti-recompression video watermarking algorithm in bitstream domain," Tsinghua Science and Technology, vol. 26, no. 2, pp. 154–162, 2021. DOI: 10.26599/TST.2019.9010050.

[21] H. Mareen, G. Van Wallendael, and P. Lambert, "Implementation-free forensic watermarking for adaptive streaming with a/b watermarking," in Proceedings of Sixth International Congress on Information and Communication Technology, X.-S. Yang, S. Sherratt, N. Dey, and A. Joshi, Eds., Singapore: Springer Singapore, 2022, pp. 325–339, ISBN: 978-981-16-2377-6.

[22] S. Voloshynovskiy et al., "Attacks on digital watermarks: Classification, estimation-based attacks, and benchmarks," IEEE Communications Magazine, vol. 39, no. 8, pp. 118–126, 2001. DOI: 10.1109/35.940053.

[23] X. Yu et al., "A survey on robust video watermarking algorithms for copyright protection," Applied Sciences, vol. 8, no. 10, p. 1891, 2018. DOI: 10.3390/app8101891.

[24] Z. Qin et al., "A dynamic watermark based proactive deepfake defense," arXiv preprint arXiv:2401.XXXXX, 2024.

[25] Y. Zhou et al., "Robust watermarking for video forgery detection with improved imperceptibility and robustness," arXiv preprint arXiv:2207.03409, 2022.