TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

VIỆN TOÁN ỨNG DỤNG VÀ TIN HỌC

HỆ MẬT SINH TRẮC DỰA TRÊN DẤU VÂN TAY VÀ MẬT MÃ

ĐỒ ÁN II

Chuyên ngành: HỆ THỐNG THÔNG TIN QUẢN LÝ

Giảng viên hướng dẫn: PGS. TS. NGUYỄN ĐÌNH HÂN

Sinh viên thực hiện: HOÀNG VĂN THÀNH

Mã số sinh viên: 20173586

Lớp: TƯD.03 – K62

HÀ NỘI – 2021

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

1. **Mục đích và nội dung của đồ án:**

2. **Kết quả đạt được:**

3. **Ý thức làm việc của sinh viên:**

Hà Nội, ngày 7 tháng 1 năm 2021

Giảng viên hướng dẫn

(Ký và ghi rõ họ tên)

MỤC LỤC

DANH MỤC HÌNH ẢNH, ĐỒ THỊ

PHẦN MỞ ĐẦU

1. Lý do chọn đề tài

Hiện nay, Việt Nam chúng ta là một trong các nước có sự phát triển mạng internet rất mạnh. Với dân số 97.72 triệu người, thì có 68.17 triệu dân sử dụng mạng internet(bằng 69.76% dân số). Trong đó có 148.5 triệu thuê bao di động(bằng 152% dân số) nằm trong top 30 của thế giới, có 65 triệu tài khoản mạng xã hội(bằng 67% dân số). Chúng ta cũng là một trong các nước bắt kịp các xu thế công nghệ mới về viễn thông như tỷ lệ phủ sóng 4G đạt trên 95%, thử nghiệm phát sóng 5G và mục tiêu triển khai thương mại đầu năm 2021, tắt sóng các dịch vụ lỗi thời như 2G và tiến tới là cả 3G. Ngoài ra, tỷ lệ sử dụng smartphone là 45% dân số và tiến tới mục tiêu 100% với các mẫu điện thoại 4G được trợ giá chỉ 600 nghìn đồng. Và nhà mạng Viettel lớn nhất nước ta là nhà mạng duy nhất ở Đông Nam Á nằm trong top 30 thương hiệu giá trị nhất thế giới và vào top 20 nhà mạng viễn thông lớn nhất thế giới theo quy mô thuê bao. Đó là một trong số ít những con số minh chứng cho sự phát triển hệ thống internet tại Việt Nam chỉ sau hơn 20 năm bắt đầu từ ngày đầu tiên có mạng internet vào tháng 11 năm 1977.

Cùng với sự phát triển nhanh chóng của hệ thống mạng internet còn kèm theo vô số các vấn đề liên quan, đặc biệt là bảo mật thông tin truyền tải trên không gian mạng và xác thực người dùng các giao dịch. Thực tế hiện nay phương pháp để xác thực người dùng với một hệ thống phổ biến vẫn là sử dụng mật khẩu hoặc mã PIN. Điều này dẫn đến rất nhiều nguy cơ người dùng có thể bị quên hoặc lấy cắp bằng các phương pháp như social engineering. Một số phương pháp bảo mật khác như sử dụng xác thực 2 lớp bằng OTP cũng tồn tại một số rủi ro cho do nhận thức về bảo mật của người dùng Việt Nam còn chưa cao.

Để giải quyết các vấn đề trên, một phương pháp xác thực được sử dụng ở nhiều hệ thống hiện nay đó là sử dụng dữ liệu sinh trắc học của người dùng. Tuy

nhiên để đảm bảo tính bảo mật, toàn vẹn dữ liệu, xác thực đúng và đảm bảo tốc độ không phải hệ thống nào cũng có thể đảm bảo được. Trong đồ án này, tôi đã xây dựng một hệ mật xác thực người dùng sử dụng dữ liệu sinh trắc học và mã PIN để có thể xác thực đúng với tốc độ cao. Cụ thể ở đây tôi sử dụng dấu vân tay của người dùng do phổ biến và tiện lợi hơn các dữ liệu sinh trắc học khác. Các dữ liệu sinh trắc học khác như mống mắt, gương mặt, v.v. có thể được thay thế dễ dàng.

2. Đối tượng và phạm vi nghiên cứu của đề tài

Đề tài nghiên cứu dữ liệu sinh trắc học của người dùng là dấu vân tay. Các phương pháp, kỹ thuật xử lí, nhận dạng, trích xuất và so khớp dấu vân tay để xác thực người dùng. Đề tài cũng nghiên cứu hệ mật MAS và ứng dụng vào việc xác thực người dùng.

Trong đề tài nghiên cứu này, tôi quan tâm tới việc xác thực người dùng bằng dữ liệu sinh trắc học và so sánh về mặt tốc độ, hiệu năng, độ chính xác, rủi ro bảo mật, v.v. với các hệ mật khác có mục đích tương tự nhưng sử dụng các hệ mật khác như RSA, AES, DES, v.v.

3. Phương pháp nghiên cứu

Quá trình nghiên cứu đề tài có sử dụng kết hợp nhiều phương pháp khác nhau.

Bước 1: Nghiên cứu thực trạng của bài toán trong thực tế

Bước 2: Khảo sát, phân tích dữ liệu sinh trắc học và tổng quan hệ mật

Bước 3: Nghiên cứu lý thuyết liên quan

Bước 4: Xây dựng chương trình xử lí mẫu ảnh dấu vân tay và trích xuất đặc trưng.

Bước 5: Xây dựng thuật toán xác thực hai mẫu vân tay dựa trên các đặc trưng trích xuất được.

Bước 6: Kết hợp với thuật toán mã hoá để xây dựng thành hệ mật xác thực.

4. Ý nghĩa khoa học và thực tiễn của đề tài

Kết quả của đề tài nghiên cứu đưa ra một giải pháp mới cho việc xác thực người dùng. Có thể ứng dụng vào việc xác thực người dùng cho các giao dịch hiện nay, bảo mật dữ liệu trong thời đại internet phát triển nhanh chóng.

5. Lời cảm ơn

Để hoàn thành được đề tài nghiên cứu này, em xin chân thành cảm ơn sự giúp đỡ, góp ý nhiệt tình của thày giảng viên hướng dẫn **PGS. TS. Nguyễn Đình Hân**. Nhờ sự giúp đỡ, chỉ bảo, theo dõi sát sao tiến độ nghiên cứu để em có thể hoàn thành được đề tài nghiên cứu này.

Em xin chân thành cảm ơn tới các thày cô Viện Toán ứng dụng và Tin học đã chỉ dạy, giúp đỡ em trong quá trình học tập kỳ này. Em cũng xin cảm ơn bạn bè cùng lớp đã động viên, thảo luận, trao đổi kiến thức trong quá trình thực hiện đề tài.

Em xin chân thành cảm ơn!

PHẦN NỘI DUNG

CHƯƠNG 1: TỔNG QUAN VỀ BÀI TOÁN XÁC THỰC DẤU VÂN TAY

1.1. XÁC THỰC NGƯỜI DÙNG

Ngày nay, con người dành một phần của cuộc sống của họ trên các trang web mạng xã hội. Ở trên đó là một môi trường ảo nơi mà họ có thể tìm bạn bè, gặp gỡ và trò truyện thông qua video call. Họ chia sẻ thông tin hoạt động xã hội mỗi ngày. Họ không chỉ tham gia một mạng xã hội mà có thể nhiều mạng xã hội cùng một lúc như Facebook, Zalo, WeChat, v.v. Ngoài tài khoản mạng xã hội ra thì người dùng còn có nhiều tài khoản khác ở các hệ thống khác nhau như tài khoản ngân hàng, tài khoản email, tài khoản sinh viên, tài khoản nhân viên, v.v. Bất cứ hệ thống nào phục vụ nhiều người dùng thì mỗi người dùng trong hệ thống đều được quản lý dưới dạng các tài khoản. Mỗi khi người dùng muốn sử dụng các chức năng thì cần phải được xác thực. Giống như một toà nhà bị khoá, chỉ những người được phép mới được cấp chìa khoá và mở cửa vào được. Tức các người dùng cần phải được xác thực với hệ thống.

Vấn đề xác thực người dùng là một vấn đề rất quan trọng trong việc xây dựng các hệ thống phục vụ nhiều người dùng. Đặc biệt là hiện nay nhiều nghiệp vụ, quy trình được số hoá lên không gian mạng nên càng phải đặc biệt chú ý về bài toán định danh người dùng. Hiện nay có rất nhiều phương pháp đã áp dụng trong các hệ thống để cho phép xác thực người dùng. Mỗi phương pháp sẽ có các ưu nhược điểm riêng.

Phương pháp xác thực và định danh người dùng phổ biến nhất chính là sử dụng tên người dùng(username) và mật khẩu(password). Đây là một phương pháp đơn giản, dễ thực hiện và tốc độ nhanh nhưng tồn tại rất nhều lỗ hổng bảo mật. Đầu tiên là có thể bị tấn công vét cạn(brute-force attack) tìm mật khẩu. Vấn

đề này có thể giải quyết nếu tăng độ khó cho mật khẩu. Ví dụ mật khẩu dài trên 8 ký tự và có chứa chữ in hoa, in thường, chữ số và ký tự đặc biệt. Nếu người dùng sử dụng mật khẩu quá yếu thì dễ đoán và tấn công. Nhưng nếu mật khẩu phức tạp thì người dùng lại dễ quên. Đặc biệt mỗi người thường có rất nhiều tài khoản ở nhiều hệ thống khác nhau. Nên họ thường đặt mật khẩu giống nhau cho các tài khoản đó. Vậy nên nếu bị lộ một tài khoản thì có thể các tài khoản khác cũng bị tấn cônng. Và khi bị lộ mà người dùng muốn thay đổi mật khẩu thì phải đổi lại hết các tài khoản khác có mật khẩu giống. Tuy rằng tốc độ của phương pháp này rất tốt do hệ thống chỉ phải so khớp mật khẩu lưu trong hệ thống và mật khẩu người dùng nhập vào nhưng có khá nhiều rủi ro bảo mật liên quan đến phương pháp này. Một số biện pháp có thể nâng cao tính bảo mật cho phương pháp này như xác minh hai lớp 2FA, mật khẩu dùng một lần OTP, v.v.

Một phương pháp đang là xu hướng hiện nay dần phổ biến đó chính là sử dụng dữ liệu sinh trắc học của người dùng. Sinh trắc học hay Công nghệ sinh trắc học là công nghệ sử dụng những thuộc tính, đặc điểm sinh học riêng của mỗi cá nhân nhưu vân tay, khuôn mặt, mống mắt, giọng nói tĩnh mạch ngón tay, tĩnh mạch lòng bàn tay, v.v. để nhận diện và định danh giữa các người khác nhau. Do các thuộc tính sinh trắc học cá nhân của mỗi người là duy nhất và hầu như không thể tái tạo hoặc giả mạo, ên các yếu tố sinh trắc học được ứng dụng và phát triển thành các công nghệ bảo mật sinh trắc học giúp cho việc bảo mật và xác thực trở nên an toàn và tin cậy hơn. Dữ liệu sinh trắc học gắn liền với cơ thể mỗi chúng ta nên người dùng không bao giờ quên mà cũng chẳng cần phải nhớ. Nó đem lại sự tiện lợi cho người dùng và độ chính xác rất cao hiện tại.

1.2. DỮ LIỆU SINH TRẮC HỌC – DẤU VÂN TAY

Dấu vân tay là một mô hình của các đường vân biểu bì trên ngón tay của chúng ta. Nó được hình thành vào tuần thứ 13 của thai kỳ và hoàn thiện vào tuần 19 đến 24 của thai kỳ. Nó rất độc đáo, khó thay đổi được và tồn tại bền bỉ trong suốt cuộc đời của mỗi chúng ta. Về mặt sinh học, chúng làm tăng ma sát khi ta cầm nắm đồ vật, giúp cho việc cầm nắm được tốt hơn trong các điều kiện ẩm ướt. Ngoài ra, chúng cũng khuếch đại các kích thích từ môi trường tới ngón tay và làm cho việc truyền tín hiệu tới các dây thần kinh xúc giác được tốt hơn.

Ảnh 1: Dấu vân tay của con người

Nhà khoa học tài ba Sir Francis Galton đã tính toán được rằng xác suất để có hai dấu vân tay giống hệt nhau là 1/64 tỷ người. Nó có nghĩa là có 7.2 tỷ dấu vân tay khác nhau tồn tại trên hành tinh của chúng ta. Đặc điểm cá nhân hoá đó của dấu vân tay là điều được nghiên cứu và ứng dụng rất nhiều hiện nay. Do tính đặc thù như vậy nên dấu vân tay được úng dụng rất nhiều trong các lĩnh vực khoa học khác nhau.

Trong lĩnh vực giám định pháp y, thu thập dấu vân tay từ hiện trường rồi so khớp với database về dữ liệu sinh trắc học dấu vân tay của công dân để tìm ra kẻ tình nghi. Ngoài các dữ liệu về đường vân trên ngón tay, các mô ngón tay cũng thường xuyên tiết ra các chất dịch dầu và tế bào chết. Mỗi người khác nhau có dư lượng chất và thành phần chất bài tiết khác nhau do quá trình trao đổi chất và thói

quen sinh hoạt của mỗi người khác nhau. Ví dụ với những người hay hút thuốc lá thì dấu vân tay của họ sẽ chứa nhiều dấu vết của chất cotinine do chất nicotine trong thuốc lá chuyển hoá.

Trong xác thực người dùng của các thiết bị thông minh. Do không thể tồn tại hai dấu vân tay giống nhau nên các thiết bị thông minh đã ứng dụng mạnh mẽ vào sản xuất và ứng dụng. Ứng dụng đầu tiên là khoá điện tử nhận dạng bằng dấu vân tay. Khi dấu vân tay nhập vào trùng khớp với dấu vân tay được đăng ký ban đầu trong hệ thống thì khoá sẽ được mở. Và những ứng dụng này đã được tiếp tục đưa vào hàng loạt các thiết bị khác như: máy chấm công, khoá vân tay trên thiết bị di động, ổ khoá cửa thông minh, v.v.

Trong lĩnh vực quốc phòng. Dấu vân tay thu thập tại hiện trường vụ án hoặc các bằng chứng tội phạm đã được sử dụng trong khoa học pháp y để xác định nghi phạm hay nạn nhân. Nhận dạng vân tay được coi như hệ thống quan trọng trong các cơ quan cảnh sát ở thế kỷ 19, khi nó thay thế đo nhân trắc học là một phương pháp đáng tin cậy hơn để xác định tội phạm.

Trong việc theo dõi hồ sơ công dân. Dấu vân tay phục vụ các chính phủ trên toàn thế giới trong vòng 100 năm qua để cung cấp xác định chính xác hồ sơ tội phạm. Vì vậy, dấu vân tay là công cụ cơ bản cho việc xác định người có tiền sử tội phạm trong cơ quan cảnh sát.

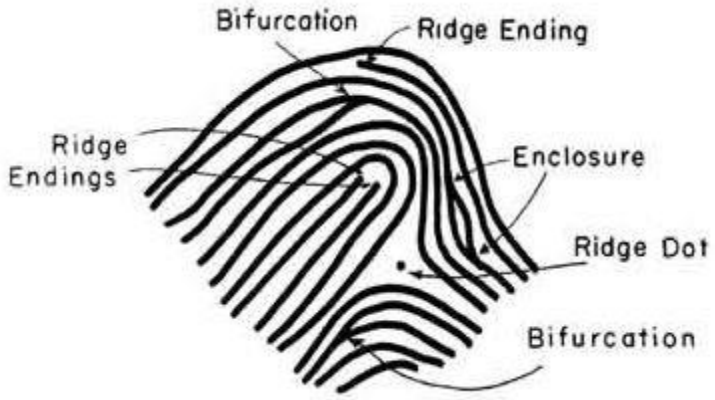
Dấu vân tay tồn tại suốt cuộc đời của một con người và nó sẽ không bị thay đổi về cấu trúc, đường nét theo thời gian. Nếu có thể thay đổi theo thời gian thì chắc chắn chỉ là sự thay đổi lớn lên của các mô mà thôi. Chính vì vậy, dấu vân tay là cơ sở khoa học quan trọng khám phá tiềm năng của con người. Việc đánh giá dấu vân tay sẽ giúp chúng ta dễ dàng đánh giá thiên khiếu bẩm sinh, hiểu được tính cách... của mỗi người.

CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

2.1. DẤU VÂN TAY

2.1.1. Cấu trúc sinh học của dấu vân tay

Vân tay là các mô biểu bì trên bề mặt da của bàn tay. Các mô biểu bì nhô lên



Ảnh 2: Các thành phần của dấu vân tay

thành các đường nổi trên bề mặt được gọi là đường vân(ridge hoặc streak). Trong mẫu ảnh dấu vân tay, các đường vân chạy dài và có hai đầu kết thúc(ridge ending). Phần lõm xuống nằm giữa hai đường vân cạnh nhau gọi là thung lũng. Một số đường vân giao nhau hoặc tách thành hai nhánh con(gọi là bifurcation). Một số vân rất nhỏ như một điểm chấm nổi lên trên vân tay.

Người ta cũng phân loại dấu vân tay theo hình dạnh của đường vân. Gồm ba loại chính: vân vòm, vân gấp khúc, vân xoáy ốc. Vân vòm là đường vân chỉ hơi cong lên, không quá cuộn vào nhau. Vân gấp khúc là các đường vân ở vùng trung tâm vân tay gấp lại thành một khúc lớn. Vân xoáy ốc là các đường vân ở trung tâm vân tay cuộn tròn lại với nhau. Hay thường được gọi là “hoa tay”. Có thể phân loại nhỏ hơn nữa dựa vào hướng gấp của các dấu vân tay gấp khúc từ bên trái hay phải, số lượng gấp khúc(một hoặc hai), v.v.

Ảnh 3: Ba loại vân tay theo hình dáng đường vân. (Trái) vân vòm. (Giữa) vân xoáy ốc. (Phải) vân gấp khúc.

Thông thường các hệ thống sử dụng dấu vân tay để xác thực dựa vào các điểm như hai đầu các đường vân, điểm phân tách hoặc gộp đường vân làm các đặc trưng sinh học.

2.1.2. Lấy mẫu ảnh dấu vân tay

Khi trình độ khoa học kĩ thuật chưa phát triển, con người thường lấy mẫu dấu vân tay bằng cách lăn ngón tay trên mực đen rồi lăn lại trên một mặt phẳng khác

Ảnh 4: Lấy mẫu vân tay thủ công

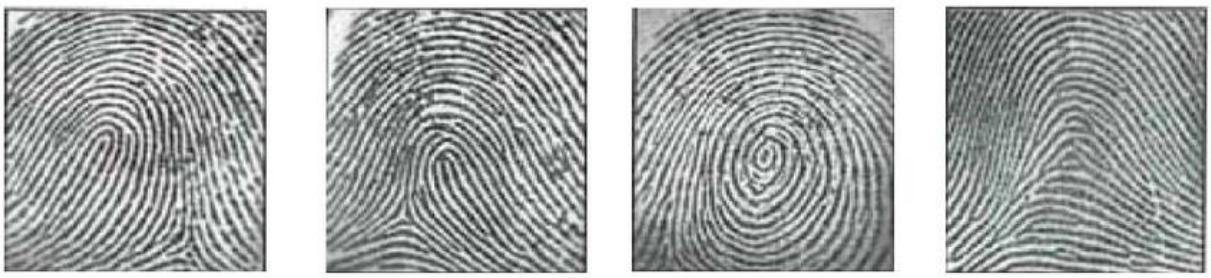
có màu đối lập lại, thường là màu trắng. Ngoài ra, các mẫu dấu vân tay có thể lấy ở một bề mặt nào đó. Khi tay ta chạm vào hầu hết các mặt phẳng đều có lưu lại dấu vân tay ẩn ở đó mà không thể nhìn thấy bằng mắt thường, gọi là Latent Print. Đây là cơ hội để pháp y có thể lấy được dấu vết kẻ tình nghi ở hiện trường. Các mẫu này tuy không thể nhìn thấy bằng mắt thường nhưng có thể sử dụng một số phương pháp vật lý, hoá học như sử dụng ánh sáng từ đèn pin, bột có thể thu được mẫu từ bề mặt. Để so khớp hai dấu vân tay bằng mẫu này thì thường sẽ dùng phương pháp thủ công là bằng mắt thường. Phương pháp này mất nhiều thời gian, hiệu quả không cao. Đồng thời việc lưu trữ, bảo quản mẫu dấu vân tay để quản lý cũng gặp nhiều khó khăn. Trong quá trình bảo quản mẫu có thể gây mất mát, hư hỏng hoặc giảm chất lượng của mẫu dấu vân tay.

Hiện tay, với trình độ khoa học kỹ thuật tiên tiến, dấu vân tay được lấy bằng các máy quét dấu vân tay, gọi là fingerprint sensor hoặc fingerprint scanner. Các công nghệ hiện nay được sử dụng để đọc dấu vân tay kỹ thuật số là: Opptical scanner, Capacitive hay CMOS scanner, Ultrasound fingerprint scanner, Thermal scanner. Opptical scanner là cảm biến vân tay quang học, sử dụng một camera kỹ thuật số để chụp lấy hình ảnh của dấu vân tay. Capacitive or CMOS scanner sử dụng các tụ điện để lấy hình ảnh dấu vân tay bằng cách ghi nhận sự thay đổi điện tích tại các vị trí tiếp xúc với các đường vân. Hai loại trên có độ chính xác khá

Ảnh 5: Cảm biến vân tay sử dụng công nghệ Capacitive scanner

cao nhưng không hiệu quả nếu dấu vân tay có dính nước. Loại cảm biến thứ ba là Ultrasound fingerprint scanner. Bằng cách sử dụng sóng siêu âm và phân tích sóng phản hồi về có thể nhận được sơ đồ về các đường vân của dấu vân tay. Loại cảm biến Thermal scanner thì lại dùng cảm biến nhiệt để cảm nhận sự chênh lệch nhiệt độ giữa các điểm tiếp xúc với vân tay và các điểm còn lại. Hai loại này có thể hoạt động tốt với các dấu vân tay có dính mồ hôi, nước.

Trong khuôn khổ đề tài này, tôi chỉ trình bày các quá trình xử lí từ khi nhận được ảnh kỹ thuật số các dấu vân tay từ các thiết bị trên. Các vấn đề liên quan đến thiết bị đọc dấu vân tay sẽ không được nghiên cứu đến. Mẫu dấu vân tay được sử dụng theo tiêu chuẩn FBI là ảnh xám có mật độ điểm ảnh tối thiểu 500 ppi và có kích thước 500x500.



Ảnh 6: Một số mẫu dấu vân tay

Tuy nhiên, do chất lượng cảm biến hoặc quá trình lấy mẫu khiến cho chất lượng ảnh thu được có chất lượng khác nhau. Mặt khác để máy tính có thể hiểu được mẫu dấu vân tay và so khớp, định danh người dùng gắn với mẫu vân tay đó thì phải trích xuất ra được các đặc trưng của dấu vân tay đó. Vậy nên cần có một quá trình tiền xử lí ảnh dấu vân tay để có thể dễ dàng trích xuất ra các đặc trưng của nó.

2.2. HỆ MẬT MAS

Mật mã là một nghệ thuật và khoa học để giữ cho các thông tin được an toàn, không bị xem trộm bởi những kẻ không được phép. Các hệ mật có hai quá trình là mã hoá thông tin thành bản mã và giải mã ngược trở lại thành thông tin. Các hệ mật chia thành hai loại: đối xứng và bất đối xứng. Hệ mật mã đối xứng là hệ mật sử dụng khoá mã hoá và giải mã giống hệt nhau và được giữ bí mật giữa hai bên. Trước khi gửi tin nhắn đi, người gửi sẽ sử dụng khoá bí mật để mã hoá thành bản mã, sau đó gửi bản mã cho người nhận. Sau khi nhận được bản mã, người nhận sẽ sử dụng khoá bí mật để giải mã lại thành tin nhắn. Nếu khoá bị lộ ra ngoài thì hệ mật sẽ không còn tác dụng. Hệ mật mã bất đối xứng là hệ mật sử dụng cặp khoá bí mật – công khai. Người nhận sinh ra cặp khoá công khai và bí mật. Khoá công khai sẽ được gửi cho người gửi kể cả qua môi trường không an toàn. Người gửi sẽ sử dụng nó để mã hoá thành bản mã. Rồi gửi bản mã cho người nhận. Sau đó người nhận sử dụng khoá bị mật để giải mã lại thành tin nhắn.

Về lý thuyết, hệ mật mã là một hệ bao gồm 5 thành phần $(\boldsymbol{\mathcal{P}}, \boldsymbol{\mathcal{C}}, \boldsymbol{\mathcal{K}}, \boldsymbol{\mathcal{E}}, \boldsymbol{\mathcal{D}})$ với:

- $\boldsymbol{\mathcal{P}}$: là tập hữu hạn các bản tin nhắn
- $\boldsymbol{\mathcal{C}}$: là tập hữu hạn các bản mã hoá
- $\boldsymbol{\mathcal{K}}$: là tập hữu hạn các khoá
- Với mỗi $K \in \boldsymbol{\mathcal{K}}$, có một phép mã hoá $e_K \in \boldsymbol{\mathcal{E}}$ và một phép giải mã tương ứng là $d_K \in \boldsymbol{\mathcal{D}}$. Mỗi $e_K: \boldsymbol{\mathcal{P}} \rightarrow \boldsymbol{\mathcal{C}}$ và $d_K: \boldsymbol{\mathcal{C}} \rightarrow \boldsymbol{\mathcal{P}}$ là các ánh xạ sao cho: $d_k(e_k(x)) = x$ với mọi $x \epsilon \boldsymbol{\mathcal{P}}$.

Một hệ mật tốt phải đảm bảo an toàn, khó phá vỡ và tốc độ nhanh. Có rất nhiều hệ mật mã phổ biến hiện nay như AES, DES, RSA, v.v. Trong đề tài là, tôi sử dụng một hệ mật là MAS có nhiều ưu điểm hơn các hệ mật trên.

2.2.1. Khái iệm k-không nhập nhằng

Xét một ngôn ngữ $X \subseteq A^+$ và số tự nhiên $k \geq 0$. Khi đó:

i. Tập X được gọi là k-không nhập nhằng nếu: với mọi $k \geq m \geq 1$ và với mọi $x_1, x_2, \dots, x_k, y_1, y_2, \dots, Y_m \in X$ nếu $x_1 x_2 \dots x_k = y_1 y_2 \dots y_m$ thì $k = m$ và $x_i = y_i$ với $i = 1,2,..,k$.

Ngược lại, X được gọi là k-nhập nhằng.

ii. Nếu tồn tại k lớn nhất làm cho X là k-không nhập nhằng thì ta nói X có độ không nhập nhằng k. Ngược lại, nếu không tồn tại k, ta nói X có độ không nhập nhằng là ∞.

2.2.2. Hệ mật đa trị và nhập nhằng

Cho $A = \{u_1, u_2, \dots, u_n\}$ là bảng hữu hạn các chữ cái. Xét ngôn ngữ X có độ nhập nhằng $k > 0$ sao cho X có thể phân hoạch được thành n tập con: $X_1, X_2, \dots, X_n, X_i \cap X_j = \emptyset, \forall i \neq j, X_1 \cup X_2 \cup \dots \cup X_n = X$.

Gọi X_P là tập các phân hoạch có thể có của X. TỪ đó ta định nghĩa hệ mật:

Cho $\boldsymbol{\mathcal{P}} = A^{\leq k}$, $\boldsymbol{\mathcal{C}} = X^*$. $\boldsymbol{\mathcal{K}}$ bao gồm tất cả các hàm đơn ánh đa trị $g: A \to X_P = \{X_1, X_2, \dots, X_n\}$. Với mỗi $g \in \boldsymbol{\mathcal{K}}$, ta định nghĩa:

$$e_g(x) = w \in g(x) \quad (1)$$

Và

$$d_g(w) = \{y | w \in g(y)\} \quad (2)$$

2.2.3. Lược đồ hệ mật MAS

Cho m là số nguyên dương cố định và S_k là một chuỗi bit bí mật độ dài m. Xét ngôn ngữ $X \subseteq \{0,1\}^*$ có độ không nhập nhằng k sao ccho thoả mãn điều kiện: $\forall x_1, x_2, \dots, x_k \in X$, ta có $|x_1| + |x_2| + \dots + |x_k| \leq m$.

Thủ tục mã hoá ENCODE của hệ mật được mô tả như sau:

```
procedure ENCODE(u):
    i = 1, j = 1;
    while i ≤ n do
```

count = 1;

while $(count \leq k)$ and $(|w_j| < m)$ do

if $|w_j e_g(u_i)| \leq m$ then

$w_j = w_j e_g(u_i)$;

count = count + 1;

$i = i + 1$;

else PAD(w_j);

if $|w_j| < m$ then PAD(w_j):

if $i == 1$ then $w_j' = w_j \oplus S$

else $w_j' = w_{j-1}' \oplus w_j$;

$j = j + 1$;

return $w = w_1' w_2' \ldots w_{j-1}'$

với hàm PAD(w_j) là hàm bù bit trong trường hợp không đủ khối m-bit.

Thủ tục giải mã DECODE của hệ mật được mô tả như sau:

procedure DECODE(w)

$i = 1, j = 1$;

while $j \leq q$ do

if $j == 1$ then $w_j = w_j' \oplus S$

else $w_j = w_{j-1} \oplus w_j'$;

EXTRACT(w_j, tmp);

$count = 1$;

while $(count \leq length(tmp))$ do

$u_i = d_g(tmp[count])$;

$count = count + 1$;

$$i = i + 1;$$

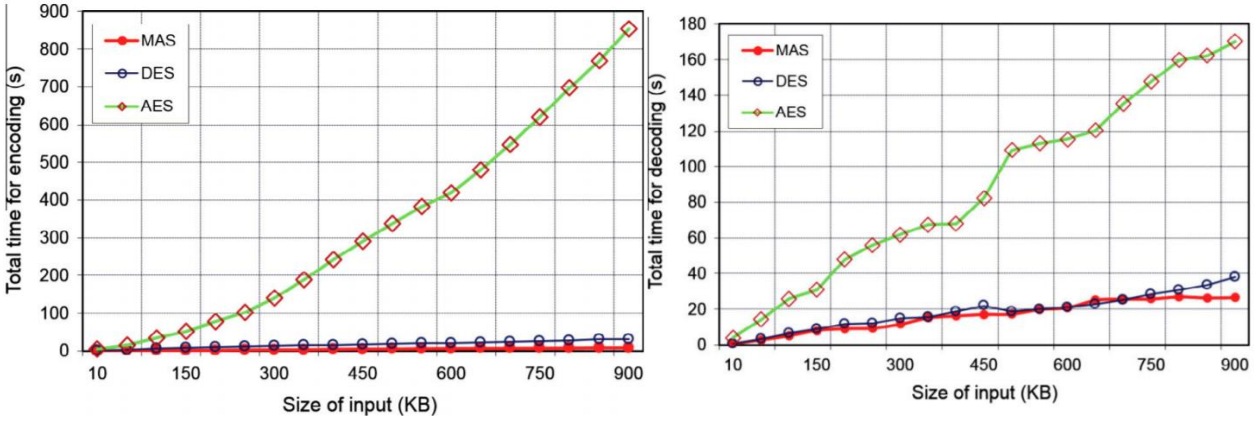
$$j = j + 1;$$

return $u = u_1 u_2 \dots u_{i-1}$

với hàm EXTRACT để tách phần bit bù đi ở phần mã hoá, trích xuất các từ trong X từ w_j rồi lưu vào mảng tmp.

2.2.4. Ưu nhược điểm so với các hệ mật còn lại

So với các hệ mật DES, AES thông dụng hiện nay thì hệ mật MAS có thời gian mã hoá và giải mã thấp hơn đáng kể.



Ảnh 7: So sánh tốc độ của hệ mật MAS với DES và AES. (Trái) Thời gian mã hoá. (Phải) Thời gian giải mã.

Tuy vậy, do là hệ mật đa trị nên việc giải mã sẽ tốn nhiều thời gian hơn so với việc mã hoá. Và việc kết hợp với các mô hình, kỹ thuật khác cũng cần phải tính toán đến số các bản mã có thể tạo ra từ một bản rõ.

CHƯƠNG 3: THUẬT TOÁN SO KHỚP DẤU VÂN TAY

3.1. XỬ LÍ ẢNH DẤU VÂN TAY

Sau khi đọc qua cảm biến thì ta nhận được hình ảnh về dấu vân tay. Tuy nhiên đây mới chỉ là dữ liệu sinh trắc học về dấu vân tay của người dùng. Muốn cho máy tính hiểu được và xác thực, so sánh được thì phải tiền xử lí để trích xuất được các đặc trưng từ ảnh. Tức là biến đổi dữ liệu dấu vân tay để được thông tin dấu vân tay. Ảnh trước khi xử lí được đưa về dạng ảnh xám. Ảnh xám

Ảnh 8: Ảnh dấu vân tay thu được từ cảm biến

hay ảnh đơn sắc là ảnh mà các điểm ảnh có giá trị nằm trong khoảng 0 đến 255, tức mỗi điểm ảnh dùng 8 bits để biểu diễn. Điểm ảnh nào có mức xám gần 0 thì càng đen và càng gần 255 thì càng sáng.

Ta thấy được do chất lượng cảm biến hoặc quá trình đo khiến cho việc các đường vân quá gần nhau như nhập lại thành mảng lớn không rõ ràng. Ảnh có quá nhiều nhiễu khiến cho việc trích xuất thông tin gặp khó khăn.

3.1.1. Chuẩn hoá ảnh

Chuẩn hoá ảnh được sử dụng để chuẩn hoá các giá trị cường độ trong hình ảnh bằng cách điều chỉnh phạm vi giá trị mức xám trong một khoảng mong muốn. Chuẩn hoá ảnh không làm thay đổi cấu trúc của các đường vân trong ảnh dấu vân tay. Mặt khác nó giúp cho thông tin về các đường vân được rõ ràng hơn. Mục đích chính của việc chuẩn hoá ảnh dấu vân tay là làm giảm sự thay đổi trong các giá trị mức xám của các điểm ảnh dọc theo đường vân và rãnh vân tay. Điều này giúp cho việc xử lí các bước tiếp theo được thuận lợi hơn.

Một phương pháp thường được sử dụng là dựa trên ngưỡng phương sai. Ban đầu hình ảnh được chia thành các khối nhỏ. Mỗi khối nhỏ sẽ được tính toán phương sai. Nếu phương sai nhỏ hơn hơn một ngưỡng nào đó được đặt trước thì coi như khối đó là nền và không có đường vân nào. Các điểm ảnh được cho là nền thì sẽ được gán hết thành màu trắng, tức là giá trị mức xám là 255. Ngược lại, các khối có phương sai lớn hơn ngưỡng đó thì được coi là thuộc đường vân

Ảnh 9: Chuẩn hoá ảnh. (Trái) ảnh gốc. (Phải) ảnh sau khi chuẩn hoá.

và gán giá trị mức xám là 0. Khi đó, các phần đường vân được nổi bật lên kể cả quá trình đo bị mờ và các vùng nền sẽ hoàn toàn trắng. Nhiều phần nhiễu của ảnh cũng bị chìm xuống.

Giả sử mỗi khối có kích thước NxN. Trung bình mức xám của khối trong ảnh I là:

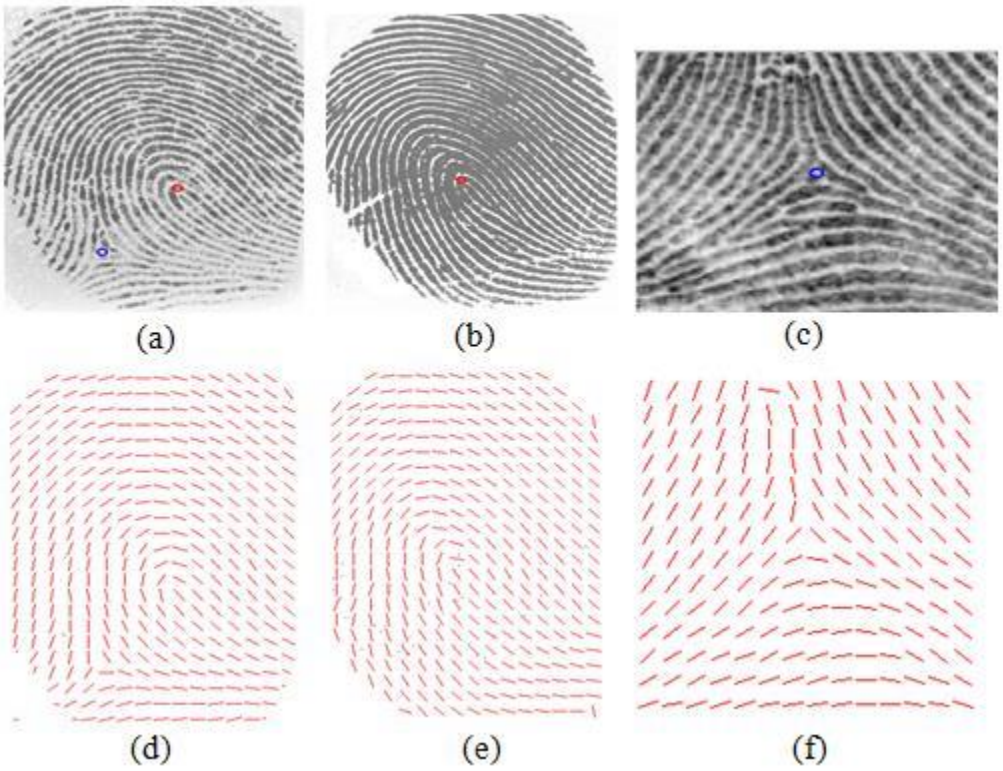
$$\boldsymbol{M(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} I(i,j)}$$

Với I(i,j) là điểm ảnh ở toạ độ (i,j) trong ảnh I. Từ đó ta tính được phương sai:

$$\boldsymbol{VAR(I) = \frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \big(I(i,j) - M(I)\big)^2}$$

3.1.2. Ước lượng bản đồ hướng của đường vân

Bản đồ hướng biểu diễn các thuộc tính nội tại của dấu vân tay. Từ đó xác định toạ độ bất biến cho các đường vân và rãnh trong các vùng cục bộ lân cận. Các đường có hướng biểu diễn trong bản đồ hướng cho biết tại các điểm đó, đường vân sẽ có hướng như thế nào ở ảnh dấu vân tay.



Ảnh 10: Ảnh dấu vân tay(Trên) và bản đồ hướng tương ứng(Dưới).

Các hướng của đường vân có thể được tính toán dựa trên hướng đạo hàm tại các điểm ảnh, tính trung bình, bỏ phiếu hoặc tối ưu. Các bước chính để tính toán ra bản đồ hướng đường vân từ ảnh đã được chuẩn hoá là:

Bước 1: Chia ảnh đã được chuẩn hoá thành các khối nhỏ kích thước 8x8.

Bước 2: Tính các đạo hàm hình ảnh $\delta_x(i,j)$ và $\delta_y(i,j)$ của các điểm ảnh.

Bước 3: Tính toán các hướng cục bộ tại mỗi điểm ảnh .

Bước 4: Làm mượt các đường xác định hướng vân bằng cách sử dụng lọc Gaussian. Sau đó cần chuyển đổi sang các vector liên tục bằng:

$$\boldsymbol{\Phi_x(i,j) = \cos(2\theta(i,j))}$$
$$\boldsymbol{\Phi_y(i,j) = sin(2\theta(i,j))}$$

Với $\Phi_x(i,j)$ và $\Phi_y(i,j)$ là các thành phần x, y của vector.

Tất cả các hướng nên được chuyển về 8 hướng trong phạm vi 90 độ tới -67.5 độ. Các số đo được làm tròn đến hướng gần nhất để trông bản đồ hướng được mượt hơn.

Ảnh 11: 8 hướng chính của các đường vân tương tự như 8 hướng địa lý

3.1.3. Phân đoạn đường vân và làm mịn

Sau khi tính được bản đồ hướng đường vân, ta sẽ sử dụng nó để lọc bớt ảnh và làm mượt các đường vân. Ta thấy ảnh gốc do chất lượng cảm biến khiến các đường vân bị nứt nẻ, đứt gãy. Còn ảnh bản đồ hướng thì tần xuất các đường biểu diễn hướng quá nhiều. Mỗi đường chỉ biểu diễn hướng của một hoặc các đường vân nằm trong khối các điểm ảnh khi tính toán, mà các đường vân cạnh nhau thì thường là song song với nhau. Ta làm tròn các tần xuất các đường hướng vân để giảm tần xuất các đường riêng biệt. Sau đó tạo bộ lọc tương ứng để làm mượt ảnh.

Ý tưởng là khi chúng ta di chuyển điểm ảnh dọc theo trục X, thì số các điểm ảnh di chuyển dọc theo trục Y là $x.\tan(\theta)$. Bộ lọc tham chiếu sẽ được đặt trên mỗi khối theo hướng trực giao với hướng chủ đạo của khối. Với kích thước bộ lọc bằng kích thước khối. Trong mỗi ô vuông khối điểm ảnh, các số thập

phân được sắp xếp dọc theo một đường thẳng với góc bằng với hướng của hướng đường vân tương ứng tại đó trên bản đồ hướng đường vân. Kiểm tra các điểm ảnh trong khối, nếu đạt mức giá trị tối thiểu thì sẽ cho bằng 1 và ngược lại sẽ cho bằng 0. Quá trình này gọi là tuyến tính hoá.

3.1.4. Làm mỏng đường vân

Bước cuối cùng trong quá trình xử lý là làm mỏng trước khi trích xuất các đặc trưng. Làm mỏng là một phương pháp hình học làm xói mòn liên tiếp

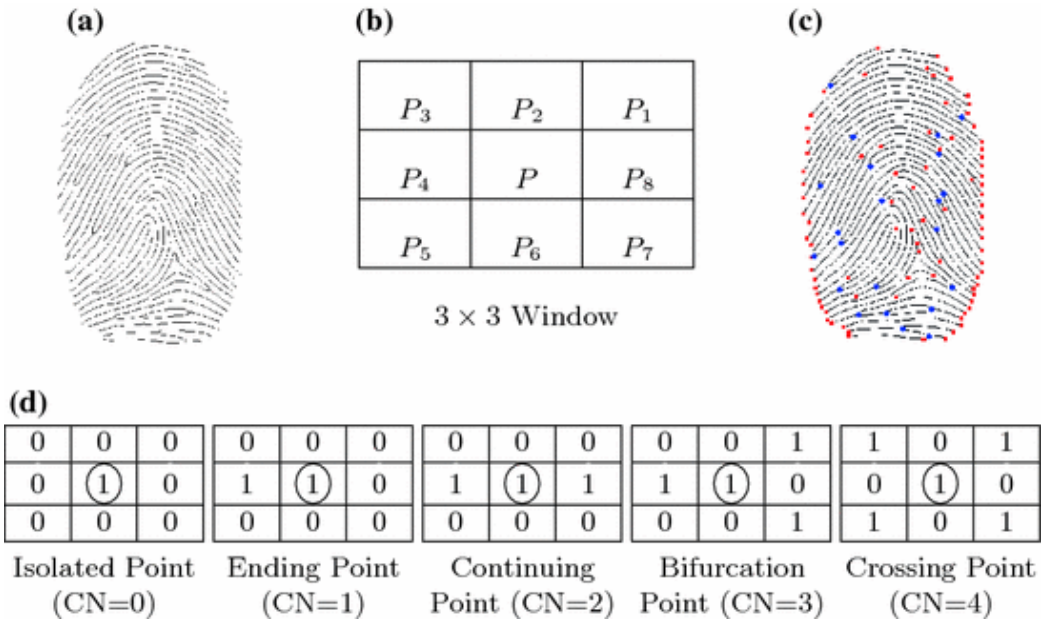
Ảnh 12: Kết quả của quá trình làm mỏng đường vân.

các điểm ảnh nền cho đến khi chúng có độ rộng chỉ bằng một điểm ảnh. Một thuật toán làm mỏng chuẩn được sử dụng, thực hiện thao tác làm mỏng bằng cách sử dụng hai lần lặp phụ. Việc áp dụng thuật toán làm mỏng cho hình ảnh dấu vân tay duy trì sự liên kết của các cấu trúc rãnh vân tay. Hình ảnh bộ xương của dấu vân tay này sau đó được sử dụng trong việc trích xuất các đặc trưng tiếp theo. Sau khi làm mỏng sẽ có một số gai xuất hiện trong hình ảnh nhị phân. Các gai này được loại bỏ bằng cách làm mịn theo hướng.

3.2. TRÍCH XUẤT ĐẶC TRƯNG TỪ DẤU VÂN TAY

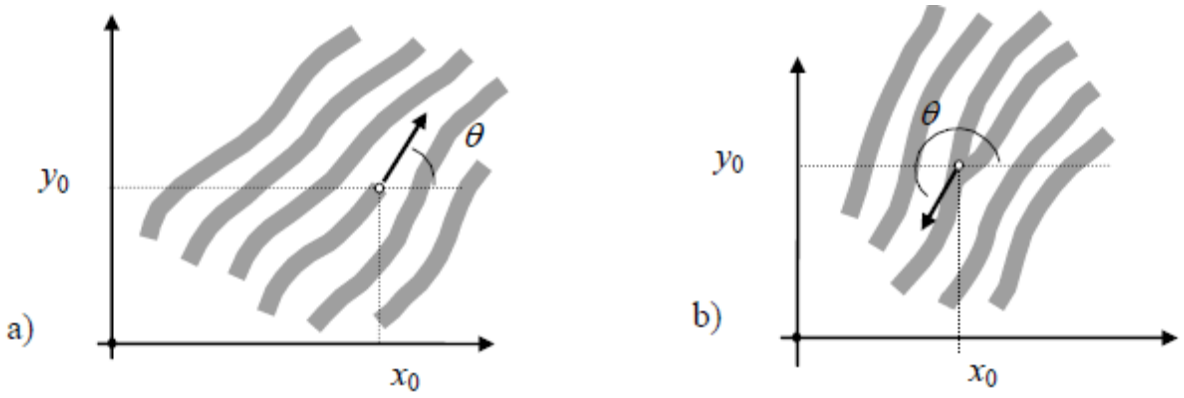
Trong hệ thống xác thực của chúng tôi, các đặc trưng của dấu vân tay dùng để so khớp và xác thực là các điểm kết thúc đường vân và các điểm rẽ nhánh. Sau khi thu được ảnh làm mỏng đường vân ở bước trước, bằng cách sử dụng một số kĩ thuật ta có thể xác định được toạ độ của hai loại điểm này.

Các đặc trưng được trích xuất bằng cách quét các khối nhỏ trong hình ảnh bằng có kích thước 3x3. Phương pháp trích xuất được sử dụng phổ biến nhất là so sánh số chéo(Crossing Number - CN). Giá trị CN sau đó được tính toán và xác định bằng một nửa tổng số chênh lệch giữa các cặp pixel liền kề trong tám vùng lân cận. Sử dụng các đặc điểm của CN, điểm ảnh có vân tay sau đó có thể được phân loại là điểm kết thúc đường vân, điểm chia đôi hoặc không phải là đường vân. Ví dụ: một điểm ảnh đường vân với CN = 1 tương ứng với một điểm kết thúc của đường vân và CN = 3 tương ứng với một điểm chia đôi.



Ảnh 13: Phương pháp Crossing Number tìm toạ độ các đặc trưng.

Ngoài toạ độ x, y của hai loại điểm đặc trưng trên, người ta còn lưu thêm một thông tin về góc nghiêng θ cho biết hướng của đường vân tại điểm đó.



Ảnh 14: Góc nghiêng ϑ đường vân tại điểm đặc trưng. (Trái) Với điểm đặc trưng là kết thúc đường vân. (Phải) Điểm đặc trưng là rẽ nhánh của đường vân.

Góc này được lấy từ bản đồ hướng đường vân được tạo ra ở phần trước. Có thêm thông tin này sẽ khiến cho việc so khớp đấu vân tay có độ chính xác hơn rất nhiều.

3.3. SO KHỚP HAI DẤU VÂN TAY

So khớp là quá trình so sánh hai mẫu vân tay có là cùng thuộc một người hay không. Đây là một quá trình bắt buộc đối với các hệ thống xác thực sử dụng dấu vân tay. Độ chính xác của thuật toán so khớp sẽ ảnh hưởng trực tiếp tới hiệu quả bảo mật của hệ thống. Mặt khác, việc so khớp cũng phải đảm bảo hiệu năng, tốc độ để nhanh chóng xác thực người dùng thực hiện giao dịch.

Sau bước trích xuất dặc trưng, với ảnh đầu vào được thiết lập có kích thước tối da là 500x500. Với hai mẫu ảnh dấu vân tay cần so khớp, ta trích xuất ra được hai tập vector đặc trưng của hai mẫu là:

$$T = \{m_1, m_2, m_3, \dots, m_p\}$$

$$Q = \{m'_1, m'_2, m'_3, \dots, m'_q\}$$

Với n, k là số điểm đặc trưng trích xuất được của hai mẫu dấu vân tay cần so khớp. Các vector đặc trưng của hai mẫu là m_i, m_j' có:

$$m_i = (x_i, y_i, \theta_i), i = 1,2, \dots, p$$

$$m_j' = (x_j, y_j, \theta_j), j = 1,2, \dots, q$$

với:

$$x, y \in [0,500]$$

$$\theta \in [0,360]$$

Để so khớp hai mẫu vân tay trên, ta so sánh sự tương quan giữa các điểm đặc trưng với nhau. Hai điểm đặc trưng được coi là khớp với nhau nếu thoả mãn hai điều kiện:

$$sd_{m_i,m_j'} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \leq r_0$$

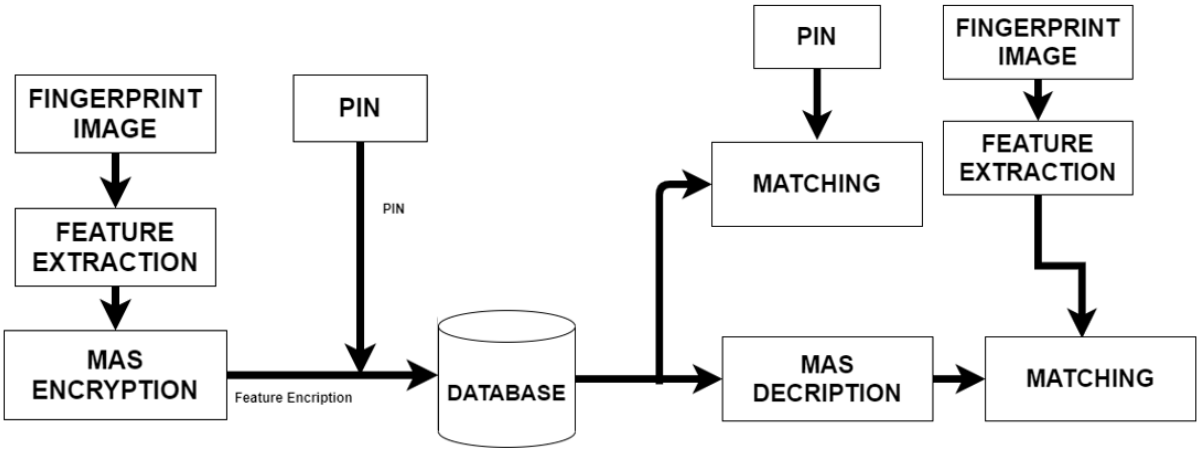
$$dd(m_i, m_j') = \min(|\theta_j' - \theta_i|, 360 - |\theta_j' - \theta_i|) \leq \theta_0$$

Với $sd_{m_i,m_j'}$ và $dd(m_i, m_j')$ là độ lệch về vị trí và góc của hai điểm đặc trưng. Và r_0, θ_0 là hai mức ngưỡng cho trước.

Quá trình lấy mẫu dấu vân tay đăng ký thường yêu cầu người dùng quét dấu vân tay nhiều lần nhằm mục đích lấy nhiều mẫu khác nhau. Từ đó khi xác thực sẽ so khớp với các mẫu đã quét. Rồi từ tỉ lệ khớp sẽ cho biết xấu vân tay đó có phải từ cùng một người hay không.

CHƯƠNG 4: HỆ MẬT SINH TRẮC DỰA TRÊN DẤU VÂN TAY VÀ MẬT MÃ

Trong phần này, tôi đề xuất một hệ mật sinh trắc học với dữ liệu đầu vào là một dấu vân tay của người dùng. Quá trình như hình bên dưới:



Ảnh 15: Tổng quan hệ mật sinh trắc học

Với các đặc trưng trích xuất là các điểm đặc trưng được mô tả ở phần trước. Chúng ta sẽ sử dụng hệ mật MAS để mã hoá từng thành phần của mỗi vector đặc trưng. Sau đó khi cần xác thực thì giải mã rồi so khớp. Do việc xác thực bằng dấu vân tay có tỷ lệ sai sót do các mức ngưỡng là thủ công và chưa thể đạt chính xác cao, thường chỉ ở mức khớp 30% là có thể kết luận khớp. Nên đề xuất thêm mã PIN để nâng cao tính bảo mật.

Gọi thành phần của vector đặc trưng cần mã hoá là u, với u là một số nguyên có giá trị trong khoảng 0 đến 500 nếu là toạ độ, hoặc 0 đến 360 nếu là góc của đường vân.

KẾT LUẬN

Qua quá trình nghiên cứu, hệ mật này đã đáp ứng được các mục tiêu ban đầu đề ra. Hệ mật sinh trắc học dựa trên dấu vân tay có độ bảo mật cao, tiện lợi cho người dùng. Các hệ mật sinh trắc học khác thường lưu trữ hình ảnh của dấu vân tay trực tiếp trên hệ thống mà không có mã hoá. Hoặc có mã hoá bằng các hệ mật như AES. Tuy nhiên, hệ mật sinh trắc học sử sụng dấu vân tay và mật mã MAS đề xuất trên đã cải tiến và cho tốc độ mã hoá cũng như giải mã nhanh hơn AES và DES. Tốc độ rất quan trọng trong việc thực hiện các giao dịch trực tuyến và yêu cầu tính an toàn, bảo mật.

TÀI LIỆU THAM KHẢO

1. Nguyen Dinh Han, Longzhe Han, Dao Minh Tuan, Hoh Peter In, Minho Jo, “A scheme for data confidentiality in Cloud-assisted Wireless Body Area Networks”, Information Sciences, 2014, 158 - 164
2. Weiping Chen and Yongsheng Gao, “A Minutiae-based Fingerprint Matching Algorithm Using Phase Correlation”, Digital Image Computing Techniques and Applications, IEEE, 2007, 234-237