A picture containing logo

Description automatically generated

**TRƯỜNG ĐẠI HỌC TÀI NGUYÊN VÀ MÔI TRƯỜNG HÀ NỘI**

**KHOA CÔNG NGHỆ THÔNG TIN**

**------------------------------------------------**

**KỲ THI KẾT THÚC HỌC PHẦN**

**NĂM HỌC 2021-2022**

**Bài tập lớn môn An toàn và bảo mật hệ thống thông tin**

**Đề số : 05**

|  |  |
| --- | --- |
| **Giảng viên hướng dẫn** | **: Nguyễn Văn Hách** |
| **Sinh viên thực hiện** | **: Nguyễn Quang Trường** |
| **Mã sinh viên** | **: 1911061019** |
| **Lớp** | **: ĐH9C5** |
| **Tên học phần** | **: An toàn và bảo mật hệ thống thông tin** |
| **Khóa học** | **: 2019 – 2023** |

***Hà Nội – 2021***

**LỜI MỞ ĐẦU**

Từ trước công nguyên con người đã phải quan tâm tới việc làm thế nào để đảm bảo an toàn bí mật cho các tài liệu, văn bản quan trọng, đặc biệt là trong lĩnh vực quân sự, ngoại giao. Ngày nay với sự xuất hiện của máy tính, các tài liệu văn bản giấy tờ và các thông tin quan trọng đề được số hóa và xử lý trên máy tính, được truyền đi trong một môi trường mà mặc định là không an toàn. Do đó yêu cầu về việc có một cơ chế, giải pháp để bảo vệ sự an toàn và bí mật của các thông tin nhạy cảm, quan trọng ngày càng trở nên cấp thiết. Mật mã học chính là ngành khoa học đảm bảo cho mục đích này. Khó có thể thấy một ứng dụng Tin học có ích nào lại không sử dụng các thuật toán mã hóa thông tin.

An toàn thông tin được định nghĩa là việc bảo vệ chống truy nhập, sử dụng, tiết lộ, sửa đổi, hoặc phá hủy thông tin một cách trái phép. Dưới một góc nhìn khác, an toàn thông tin là việc bảo vệ các thuộc tính bí mật, tính toàn vẹn và tính sẵn dùng của các tài sản thông tin trong quá trình chúng được lưu trữ, xử lý, hoặc truyền tải. An toàn thông tin có thể được chia thành ba thành phần chính: An toàn máy tính và dữ liệu, An ninh mạng và Quản lý an toàn thông tin.

**CHƯƠNG 1. GIỚI THIỆU VỀ PHẦN MỀM**

**1.1. PYTHON**

Python là ngôn ngữ lập trình hướng đối tượng đa năng có cấu trúc dữ liệu cấp cao mạnh mẽ và hệ thống thư viện lớn. Python hoàn toàn tạo kiểu động và sử dụng cơ chế cấp phát bộ nhớ tự động, cú pháp rõ ràng, đơn giản. Lợi ích khi học lập trình Python là gì? Python giúp người học dễ tiếp cận, bao gồm những người chưa từng học lập trình.

Dù rất phát triển trong những năm gần đây, Python lại có một lịch sử lâu đời. Năm 1980, Python bắt đầu được thiết kế bởi Guido van Rossum khi ông muốn tạo ra một ngôn ngữ lập trình đơn giản, dễ hiểu.

Cái tên Python xuất phát từ một chương trình hài cuối những năm 70 mà ông rất thích. Được phát hành lần đầu tiên vào năm 1994, đến nay, Python đã có tổng cộng 6 phiên bản và liên tục nnhiều năm liền năm trong TOP ngôn ngữ lập trình phổ biến nhất.

Ngôn ngữ lập trình nhập môn: Rossum hướng đến sự đơn giản và dễ hiểu nên cấu trúc Python khá rõ ràng, cho phép viết mã lệnh với số lần gõ phím tối thiểu. Ngày nay, Python được đánh giá là ngôn ngữ lý tưởng cho việc bắt đầu học lập trình.

Mã nguồn mở và hoàn toàn miễn phí: Khác với nhiều ngôn ngữ lập trình bậc cao, Python cho phép sử dụng một cách miễn phí tất cả phần mềm, chương trình được viết. Mã nguồn mở sở hữu cộng đồng đông đảo, giúp bạn giải đáp mọi thắc mắc và liên tục cập nhật, cải thiện.

Tương thích nhiều nền tảng: Python dễ dàng tương thích với nhiều nền tảng như Windows, MacOS, Linux. Thậm chí là di chuyển qua lại giữa các nền tảng cũng không gặp nhiều khó khăn.

Khả năng nhúng và mở rộng: Có thể kết hợp cùng các ngôn ngữ lập trình khác để phát triển các ứng dụng phức tạp.

Tự động chuyển đổi code: Khi dùng Python, người sử dụng hoàn toàn không phải lo các vấn đề về quản lý bộ nhớ, dọn dẹp dữ liệu… Python sẽ tự động chuyển code sang ngôn ngữ máy tính có thể hiểu.

Thư viện lớn: Nhờ số lượng thư viện khổng lồ của mình Python hoàn toàn đáp ứng được mọi nhu cầu lập trình khác nhau.

Ngôn ngữ lập trình hướng đối tượng: Dù được đơn giản hóa cho người mới học song Python vẫn giữ bản chất hướng đối tượng. Giải quyết các vấn đề trong lập trình một cách trực quan, dễ hiểu.

Tại sao nên dùng Python?

Có rất nhiều lý do để lựa chọn Python làm ngôn ngữ lập trình của bạn. Python rất dễ tiếp cận, dễ học với cú pháp đơn giản, cấu trúc rõ ràng.

Mã nguồn mở và hoàn toàn miễn phí, cộng đồng đông đảo là ưu điểm không thể chối cãi. Python còn sở hữu số lượng thư viện chuẩn lớn đáp ứng nhiều nhu cầu trong nhiều lĩnh vực khác nhau. Có thể kể đến các lĩnh vực từ lập trình, phân tích dữ liệu, khoa học máy tính…

Viết các ứng dụng web: Công việc lập trình các ứng dụng web ngày nay càng trở nên phổ biến. Các nền tảng tạo ứng dụng web khá đình đám sử dụng Python như Django, Flask, Pyramid.

Khoa học và phân tích số liệu: Hiếm có ngôn ngữ nào có được sức mạnh lớn như Python khi có thể sử dụng trong cả Data Analytics, Machine Learning, Computing Science. Bằng số lượng thư viện chuẩn lớn và các plugin như Numpy, SciPy và Panda, Python hoàn toàn đáp ứng tốt công việc phân tích và tính toán.

Tạo nguyên mẫu hay bản thử của phần mềm: Một tính năng rất hay dành cho các nhà lập trình, giúp tạo bản mẫu để thử nghiệm trước khi đưa vào thực tiễn.

Dạy và học lập trình: Python hiện đã và đang được đưa vào chương trình giảng dạy cho người lần đầu tiếp cận ngôn ngữ lập trình tại nhiều quốc gia.

**CHƯƠNG 2: ỨNG DỤNG MẬT MÃ THAY THẾ CEASAR VÀ MẬT MÃ HILL TRONG BẢO MẬT THÔNG TIN**

**2.1. Tổng quan về an toàn và bảo mật thông tin**

Khi nhu cầu trao đổi thông tin dữ liệu ngày càng lớn và đa dạng, các tiến bộ về điện tử - viễn thông và công nghệ thông tin không ngừng được phát triển ứng dụng để nâng cao chất lượng và lưu lượng truyền tin thì các quan niệm ý tưởng và biện pháp bảo vệ thông tin dữ liệu cũng được đổi mới. Bảo vệ an toàn thông tin dữ liệu là một chủ đề rộng, có liên quan đến nhiều lĩnh vực và trong thực tế có thể có rất nhiều phương pháp được thực hiện để bảo vệ an toàn thông tin dữ liệu. Các phương pháp bảo vệ an toàn thông tin dữ liệu có thể được quy tụ vào ba nhóm sau:

- Bảo vệ an toàn thông tin bằng các biện pháp hành chính.

- Bảo vệ an toàn thông tin bằng các biện pháp kỹ thuật (phần cứng).

- Bảo vệ an toàn thông tin bằng các biện pháp thuật toán (phần mềm).

Ba nhóm trên có thể được ứng dụng riêng rẽ hoặc phối kết hợp. Môi trường khó bảo vệ an toàn thông tin nhất và cũng là môi trường đối phương dễ xân nhập nhất đó là môi trường mạng và truyền tin. Biện pháp hiệu quả nhất và kinh tế nhất hiện nay trên mạng truyền tin và mạng máy tính là biện pháp thuật toán.

An toàn thông tin bao gồm các nội dung sau:

- Tính bí mật: tính kín đáo riêng tư của thông tin

- Tính xác thực của thông tin, bao gồm xác thực đối tác( bài toán nhận danh), xác thực thông tin trao đổi.

- Tính trách nhiệm: đảm bảo người gửi thông tin không thể thoái thác trách nhiệm về thông tin mà mình đã gửi.

Để đảm bảo an toàn thông tin dữ liệu trên đường truyền tin và trên mạng máy tính có hiệu quả thì điều trước tiên là phải lường trước hoặc dự đoán trước các khả năng không an toàn, khả năng xâm phạm, các sự cố rủi ro có thể xảy ra đối với thông tin dữ liệu được lưu trữ và trao đổi trên đường truyền tin cũng như trên mạng. Xác định càng chính xác các nguy cơ nói trên thì càng quyết định được tốt các giải pháp để giảm thiểu các thiệt hại.

Có hai loại hành vi xâm phạm thông tin dữ liệu đó là: vi phạm chủ động và vi phạm thụ động. Vi phạm thụ động chỉ nhằm mục đích cuối cùng là nắm bắt được thông tin (đánh cắp thông tin). Việc làm đó có khi không biết được nội dung cụ thể nhưng có thể dò ra được người gửi, người nhận nhờ thông tin điều khiển giao thức chứa trong phần đầu các gói tin. Kẻ xâm nhập có thể kiểm tra được số lượng, độ dài và tần số trao đổi. Vì vậy vi pham thụ động không làm sai lệch hoặc hủy hoại nội dung thông tin dữ liệu được trao đổi. Vi phạm thụ động thường khó phát hiện nhưng có thể có những biện pháp ngăn chặn hiệu quả. Vi phạm chủ động là dạng vi phạm có thể làm thay đổi nội dung, xóa bỏ, làm trễ, xắp xếp lại thứ tự hoặc làm lặp lại gói tin tại thời điểm đó hoặc sau đó một thời gian. Vi phạm chủ động có thể thêm vào một số thông tin ngoại lai để làm sai lệch nội dung thông tin trao đổi. Vi phạm chủ động dễ phát hiện nhưng để ngăn chặn hiệu quả thì khó khăn hơn nhiều.

Một thực tế là không có một biện pháp bảo vệ an toàn thông tin dữ liệu nào là an toàn tuyệt đối. Một hệ thống dù được bảo vệ chắc chắn đến đâu cũng không thể đảm bảo là an toàn tuyệt đối.

**2.2 Mật mã Ceasar:**

Trong mật mã học, Mật mã Caesar (hay còn được gọi là Mật mã của Caesar, Mật mã chuyển vị, Mã của Caesar hay Chuyển vị Caesar) là một trong những kỹ thuật mã hóa đơn giản và phổ biến nhất. Đây là một dạng mật mã thay thế, trong đó mỗi ký tự trên văn bản thô sẽ được thay bằng một ký tự khác, có vị trí cách nó một khoảng xác định trong bảng chữ cái.

Bước mã hóa được thực hiện trong mật mã Caesar thường được kết hợp như một phần của các dạng mã hóa phức tạp hơn, chẳng hạn như mật mã Vigenère, hiện nay vẫn được áp dụng cho mã hóa ROT13. Cũng giống như tất cả các dạng mật mã thay thế một bảng chữ cái khác, mật mã Caesar rất dễ bị phá giải và về cơ bản không đáp ứng đủ khả năng bảo mật thông tin liên lạc trong cuộc sống hiện đại.

Không thể đánh giá chính xác mức độ hiệu quả của mật mã Caesar vào thời điểm đó, nhưng có thể nó đã mang lại sức bảo mật đáng kể, đặc biệt là vì hầu hết kẻ thù của Caesar đều mù chữ và số còn lại thì cho là các thông điệp được viết bằng một thứ ngoại ngữ không xác định. Lúc này, vẫn chưa có bất cứ ghi chép nào về kỹ thuật phá giải các loại mật mã thay thế đơn giản. Những bằng chứng sớm nhất còn sót lại về phương pháp thám mã Caesar có niên đại từ khoảng thế kỷ thứ 9, đó là các công trình của Al-Kindi trong thế giới Ả Rập, với việc khám phá ra phương pháp phân tích tần suất.

Mật mã Caesar với phép dịch chuyển một vị trí được sử dụng ở mặt sau của mezuzah để mã hóa tên của Chúa. Đây có thể là vật được lưu giữ từ trước khi người Do Thái không được phép có mezuzah. Bản thân các chữ cái của mật mã đã chứa đựng "tên vị thần" mang ý nghĩa tôn giáo mà tín ngưỡng Do Thái giáo chính thống cho là có thể kiểm soát được các thế lực quỷ dữ.

Vào thế kỷ 19, phần quảng cáo cá nhân trên các tờ báo đôi khi là nơi người ta sử dụng để trao đổi những thông điệp mã hóa, sử dụng các bộ mã đơn giản. Kahn (1967) mô tả về các trường hợp mà những cặp đôi tham gia đối đáp bí mật với nhau bằng cách dùng mật mã Caesar trên tờ The Times. Ngay cả vào cuối năm 1915, mật mã Caesar vẫn được sử dụng khi quân đội Nga thay thế nó cho những dạng mật mã phức tạp hơn, vốn tỏ ra quá khó đối với họ; thế nhưng các nhà phân tích mật mã người Áo và người Đức lại rất dễ dàng phá giải các tin nhắn mã hóa của quân đội Nga.

Cấu trúc của hai đĩa quay với mật mã Caesar có thể được sử dụng trong cả việc mã hóa lẫn giải mã.

Ngày nay, chúng ta có thể tìm thấy mật mã Caesar trong các trò chơi dành cho trẻ em, chẳng hạn như vòng giải mã bí mật. Thuật toán ROT13 là một mật mã Caesar với độ dịch chuyển 13, nó được sử dụng để làm xáo trộn các văn bản được tìm thấy trên Usenet và dùng để che mờ đoạn văn (như đoạn cuối của câu chuyện cười hay phần tiết lộ nội dung câu chuyện), nhưng không được dùng như một phương pháp mã hóa nghiêm túc.

Mật mã Vigenère sử dụng mật mã Caesar với độ dịch chuyển khác nhau tại mỗi vị trí trong văn bản; giá trị của mỗi khóa mã được xác định bằng một từ khóa lặp lại. Nếu từ khóa có dung lượng bằng tin nhắn, được chọn ngẫu nhiên, không bị người khác biết đến và không bao giờ được sử dụng lại, thì đây là dạng mật mã một lần, được chứng minh là không thể phá giải. Điều kiện trên khó tới mức gần như không bao giờ đạt được. Những từ khóa ngắn hơn tin nhắn (ví dụ: "Complete Victory" được Liên minh miền Nam sử dụng trong Nội chiến Hoa Kỳ), tạo ra một dạng mật mã tuần hoàn có thể phá giải bằng cách thực hiện phân tích tần suất với thống kê nâng cao.

Text

Description automatically generated

Nhận xét: Có thể bị tấn công theo kiểu vét cạn (brute-force) bằng cách thử hết tất cả 25 khóa

\* 3 đặc điểm chính để áp dụng tấn công theo kiểu brute-force trong thuật toán này

- Giải thuật mã hóa và giải mã được biết trước

- Số khóa để thử rất ít

- Ngôn ngữ của bản rõ được biết trước và dễ dàng nhận ra

\* Giải quyết vấn đề (tổng quát)

- Sử dụng nhiều khóa

- Bản rõ có thể được nén lại (Huffman, ZIP) để cho người đọc khó nhận ra ngôn ngữ sử dụng

**2.3. Mật mã Hill**

**CHƯƠNG 3. XÂY DỰNG THUẬT TOÁN MÔ TẢ CHO CÁC MẬT MÃ TRÊN**

**3.1. Thuật toán mã hóa Caesar mô tả bằng ngôn ngữ python:**

**Text

Description automatically generated**

**3.2. Kết quả thuật toán mã hóa Caesar mô tả bằng ngôn ngữ python:**

-Nhập chuỗi cần mã hóa: NguyenQuangTruong

-Khóa K: 13

-Kết quả: AthlraDhnatGehbat

Text

Description automatically generated

**3.3. Thuật toán giải mã Caesar mô tả bằng ngôn ngữ python:**

**Text

Description automatically generated**

**3.4. Kết quả thuật toán giải mã Caesar mô tả bằng ngôn ngữ python:**

-Nhập chuỗi cần giải mã: AthlraDhnatGehbat

-Khóa K: 13

-Kết quả: NguyenQuangTruong

Text

Description automatically generated

**3.5. Thuật toán mã hóa Hill mô tả bằng ngôn ngữ python:Text

Description automatically generated**

**3.6. Kết quả thuật toán mã hóa Hill mô tả bằng ngôn ngữ python:**

-Nhập chuỗi cần mã hóa:

-Khóa K: 13

-Kết quả: NguyenQuangTruong

Graphical user interface, application

Description automatically generated