Технически университет - София

Факултет по приложна математика и информатика

Учебна дисциплина „Приложен изкуствен интелект“

Курсов проект

на

Габриела Станчева, фак. № 471221005

Десимира Димитрова, фак. № 471221086

Мертол Бекиров, фак. № 471221092

Тема: **Data obfuscation and cyber security**

София 2023

1. Въведение
   1. Идея

* Основно Понятие за Обфускация на Данни и Киберсигурност:

Обфускването на данните е процесът на преобразуване на чувствителни данни във форма, която е трудна за разбиране и анализ от неоторизирани лица.

В контекста на киберсигурността, обфускацията e важен слой като защита, служи за осигуряване на сигурност на тези дори при компрометиране на системата.

* Свързаност между Обфускация на Данни и Киберсигурност:

Обфускацията включва общата стратегия за киберсигурност, като се интегрира с други техники като криптографията и мрежовата сигурност.

Тя представлява един от многото слоеве на защита, които помагат в борбата срещу кибератаките и защитата на поверителността.

* Приложения в Реалния Свят:

Финансовите власти използват обфускация за защита на клиентската информация, докато здравеопазването прилага обфускация за защита на пациентските данни.

* Значение на Обфускацията в Съвременния Киберсвят:

В днешната ера на непрекъснато развиващи се кибератаки, обфускването играе ключова роля в защитата на ценните данни.

Правилното прилагане на обфускацията може значително да увеличи сигурността на данните и да предпази организациите от финансови и репутационни щети.

**Пример**

Примерът на Uber от 2016 година ясно демонстрира последствията от недостатъчната защита на киберсигурността и липсата на прозрачност при управлението на кризата, предоставена със сигурност на данните. Компанията се изправя пред сериозни финансови и репутационни щети след като хакерите се сдобиват с личните данни на 25 милиона потребители и шофьори. Вместо да докладва инцидент на компетентните органи, Uber избира да плати откуп на атакуващите и да се опита да разкрие случая. Това решение води до глоба от 148 милиона долара, както и до значително накърняване на имиджа на компанията.

Този случай подчертава значението на инвестициите в квалифицирани специалисти по киберсигурност и важността на създаването на надеждни системи за защита на данните. Освен това, той илюстрира нуждата от стриктно спазване на регулациите и протоколите за реагиране при киберсигурностни инциденти, включително незабавното докладване на такива събития на регулаторните органи.

* 1. Терминология

Основни Термини:

* Обфускация (Data Obfuscation): Процесът на преобразуване на чувствителни данни във форма, трудна за разбиране. Това включва промяна на оригиналните данни във вариант, който е безсмислен за неоторизирани лица.
* Криптография (Cryptography): Изкуството на шифроване и дешифриране на информация. При криптографията данните се трансформират в кодиран формат, който може да бъде възстановен обратно само със специален ключ.
* Токенизация (Tokenization): Метод за замяна на чувствителни данни с безопасни символи (токени). Тези токени нямат стойност или значение извън конкретния контекст.
* Шифроване (Encryption): Процесът на превръщане на разбираеми данни (известни като "явен текст") в кодирана форма (наричана "шифрован текст" или "криптиран текст"), която не може да бъде лесно разчетена без специален ключ.
* Маскиране на Данни (Маскиране на данни): Процес на създаване на фиктивна, но реалистична версия на данните, като се заместват оригиналните данни с алтернативни версии.

Различия между Термините:

* Обфускация срещу криптография: Въпреки това, че и двете техники целят защита на информацията, обфукцията винаги не включва обратен процес за възстановяване на оригиналните данни, докато криптографията предполага възможност за дешифриране със специален ключ.
* Токенизация срещу Маскиране на Данни: Токенизацията създава токени, които не съдържат реална информация, докато маскират данните за промяна на съществуващите данни, като запазват структурата и формата им.

Приложение на Термините в Контекста на Киберсигурността:

Токенизацията е широко използвана в сферата на финансовите услуги за защита на платежната информация, като се заменят данните на кредитните карти с безопасни токени.

Маскирането на данните се използва в средата на теста и при разработката на софтуер, където е необходимо да се работи с данни, които изглеждат реалистично, но не разкриват истинската чувствителна информация.

1.3. Подходи и методи

**Стратегически Подходи към Обфускация на Данни**

**Рисково-базиран Подход**: Тази стратегия изисква оценка на риска и определяне на нивото на обфускация според чувствителността на данните. Например, за високорискови данни като банкови сметки или лични идентификационни номера се прилагат по-строги методи на обфускация, като дълбоко маскиране или криптография, за разлика от по-обща информация, където може да се използва базово маскиране.

**Контекстно-зависим Подход**: При този подход, техниките за обфускация се адаптират според контекста на данните и изискванията за сигурност в различни сектори. Например, в здравеопазването, където защитата на пациентската информация е критична, се използват разширени форми на обфускация като токенизация, за да се осигури защита на личните здравни записи.

**Методи за Обфускация на Данни**

**Маскиране на Данни**: Техника за създаване на структурно подобни, но фиктивни версии на данни. Пример за това е маскирането на номера на кредитни карти в клиентски бази данни, където последните четири цифри се запазват за идентификация, докато останалите са скрити или заменени.

**Токенизация**: Процесът на замяна на чувствителни данни с безопасни токени, които нямат стойност извън определен контекст. Например, в платежните системи, токенизацията се използва за замяна на данните на кредитните карти с уникални идентификационни кодове при транзакции, което намалява риска от изтичане на финансова информация.

**Шифроване**: Използва се за преобразуване на данни в кодиран формат, който може да бъде дешифриран само със специален ключ. Това е особено полезно при трансфера на данни по интернет или при съхранение на данни в облачни услуги, като се осигурява високо ниво на сигурност и защита срещу хакерски атаки.

Тези стратегически подходи и конкретни методи осигуряват широк спектър от възможности за ефективна обфускация на данни, адаптирани към различни рискове и контексти. Те са основополагащи за разработването на ефективни стратегии за киберсигурност в съвременните организации.

1. Теория

В областта на киберсигурността и обфускацията на данни съществуват множество алгоритми, всеки от които служи за специфични цели и приложения. Изборът на алгоритми за детайлно разглеждане в този документ не е случаен, а е базиран на тяхната широка употреба, важност и представителност в сферата на киберсигурността.

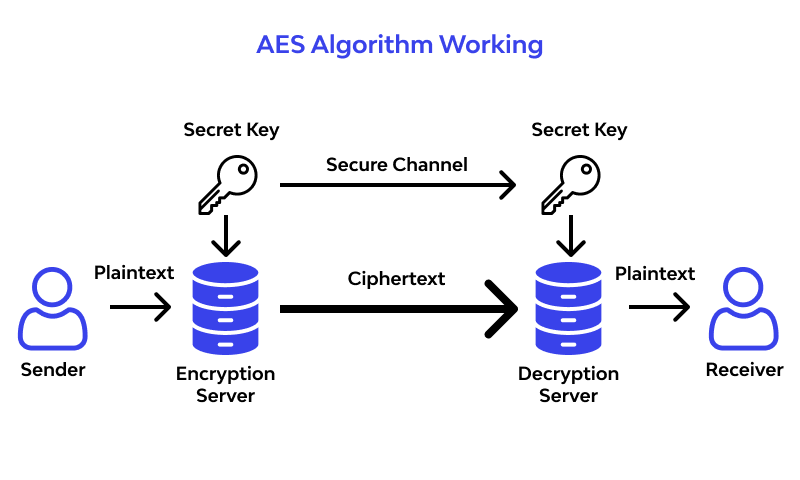
**Теория**: Описание на Алгоритми и Инструменти за Работа с Данни

**Описание на Алгоритми**:

**AES (Advanced Encryption Standard):**

**Описание**: AES е широко използван стандарт за шифроване, приет от правителството на САЩ и много други организации по света. AES използва симетрично шифроване, което означава, че същият ключ се използва за криптиране и декриптиране на данни.

**Реализация**: AES поддържа блокове данни с размер 128 бита и ключове с дължина 128, 192, или 256 бита. Шифроването преминава през няколко рунда, които включват замяна на байтове, пренареждане на редове, смесване на колони и добавяне на ключ.

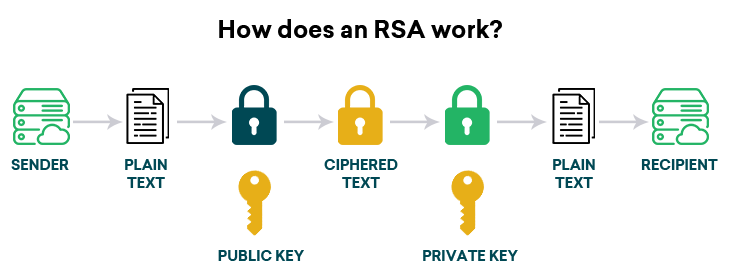
**Приложение**: Използва се за защита на конфиденциални данни в различни приложения, включително електронна поща, файлове и данни, съхранявани в облачни услуги.

**RSA (Rivest-Shamir-Adleman):**

**Описание**: RSA е метод за асиметрично шифроване, използван за защита на съобщения при онлайн комуникации. Той разчита на математическата трудност на факторизацията на големи прости числа.

**Реализация**: RSA използва два ключа – публичен и частен. Публичният ключ може да бъде споделен свободно, докато частният трябва да остане секретен. Шифроването се извършва с публичния ключ, а дешифрирането – с частния.

**Приложение**: RSA е изключително важен в контекста на сигурността на интернет, включително за защита на онлайн транзакции и осигуряване на безопасността на уеб сайтове.

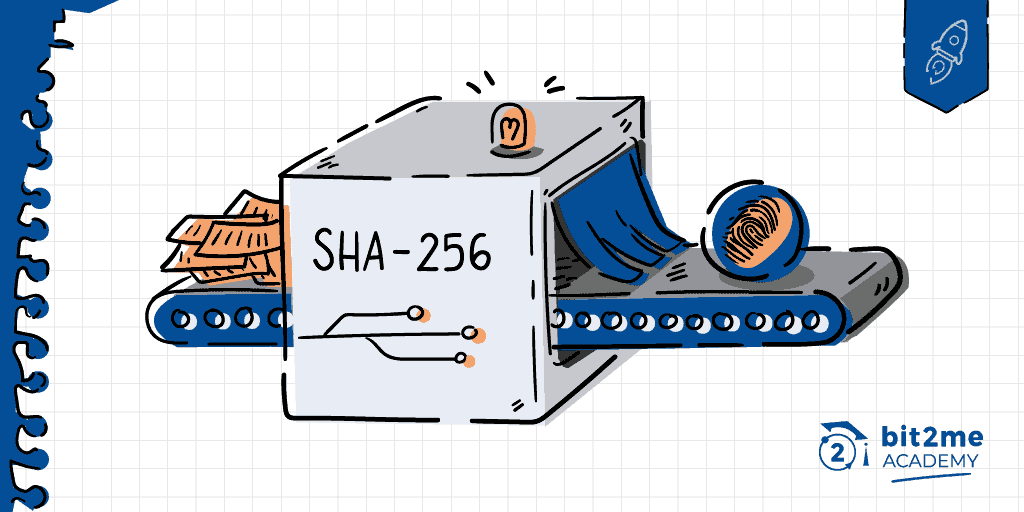


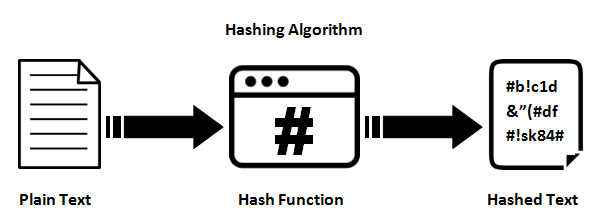
**SHA-256 (Secure Hash Algorithm 256 bit):**

**Описание**: SHA-256 е част от семейството на SHA-2 хеш алгоритми и се използва за създаване на уникален, фиксиран размер хеш от данни. Той е широко използван за проверка на целостта на данни.

**Реализация**: SHA-256 преобразува входни данни в 256-битов хеш, като използва серия от математически операции. Хешът е уникален за всеки набор от данни; дори малка промяна в данните води до коренно различен хеш.

**Приложение**: Използва се в различни приложения за сигурност, включително в криптовалути като Bitcoin, за подписване на дигитални документи и за проверка на софтуерни файлове.





**Инструменти за Работа с Данни**:

В контекста на обфускация и киберсигурност, са необходими специализирани инструменти за работа с данни. Тези инструменти включват:

**Криптографски Библиотеки**: Като OpenSSL, Crypto++, и PyCrypto, които предоставят функции за шифроване, дешифриране и хеширане.

**Платформи за Управление на Данни**: Инструменти като IBM Guardium и Oracle Data Masking, които помагат за защита и управление на чувствителни данни, включително тяхната обфускация и мониторинг.

**Инструменти за Сигурност на Приложения**: Като Veracode или Checkmarx, които анализират и идентифицират уязвимости в софтуера и помагат в интегрирането на сигурностни практики в процеса на разработка.

Тези алгоритми и инструменти са основни компоненти в арсенала на всеки специалист по киберсигурност и дават възможност за ефективна защита и управление на чувствителни данни.

1. Пример

**Описание на Задачата с Използване на RSA Алгоритъм**

**Задача:**

Създайте две конзолни приложения: едно за сървър, който генерира RSA ключове и обработва шифровани съобщения, и едно за клиент, който изпраща шифровани съобщения.

**Спецификации:**

**Сървърно Приложение**:

Генерира RSA ключове (публичен и частен) при стартиране. Съхранява частния ключ и предоставя публичния ключ на клиента. Чака да получи шифровано съобщение от клиента. Дешифрира съобщението използвайки частния ключ. Показва оригиналното съобщение на конзолата.

**Клиентско Приложение:**

Получава публичния ключ от сървъра. Чете текстово съобщение от файл, чийто път е предоставен от потребителя. Шифрова съобщението използвайки публичния ключ. Изпраща шифрованото съобщение на сървъра.

**Изисквания към Данните**:

Съобщение за Шифроване: Текстово съобщение, съхранено във файл. Потребителят въвежда пътя до файла в клиентското приложение.

**Ключове:**

Публичен и частен ключ, генерирани от сървърното приложение.

**Технически Детайли:**

Използвайте python с криптографски библиотеки с поддръжка на RSA. Комуникацията между клиент и сървър може да се осъществи чрез прост TCP сокет.

**Очакван Резултат**: При изпълнение на клиентското и сървърното приложение, потребителят трябва да може да въведе пътя до файл със съобщение в клиентската конзола. След това, съобщението ще бъде прочетено от файла, шифровано с публичния ключ и изпратено на сървъра. Сървърът след това ще дешифрира съобщението и ще покаже оригиналния текст на своята конзола.

1. Експеримент:

**Инструменти:** Криптографска Библиотека: cryptography, използвана за генериране на RSA ключове и за шифроване/дешифриране на съобщения.

**Комуникационни Протоколи:** TCP сокети, използвани за проста и директна комуникация между клиент и сървър.

**Функционалност на Клиента и Сървъра:**

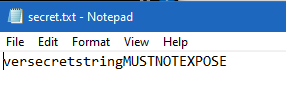
**Сървър:** Генерира чифт RSA ключове, изпраща публичния ключ на клиента, след това чака, получава и дешифрира шифрованото съобщение.

**Клиент:** Свързва се със сървъра, получава публичния ключ, чете съобщение от файл, зададен от потребителя, шифрова го и го изпраща обратно на сървъра.

**Резултати от Експеримента:** Успешно шифроване на съобщение от клиента и неговото дешифриране от сървъра. Валидация на коректността на шифроването и дешифрирането чрез сравняване на оригиналното съобщение и полученото след дешифрирането съобщение на сървъра.

**Анализ и Изводи:** RSA алгоритъмът е ефективен за шифроване на малки по обем съобщения, като в този случай се използва за шифроване на текст от файл. Взаимодействието между клиента и сървъра демонстрира основни принципи на асиметричната криптография.

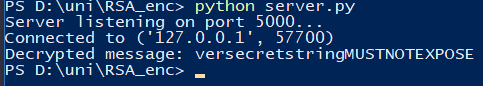
Съдържание на нашия файл:



Клиент:



Сървър:



**Заключение и Ограничения:** Този експеримент е основно учебно демонстрация и не трябва да се използва директно в производствена среда.

**Ограничения**: Размер на Съобщението: RSA е ограничен по отношение на размера на съобщението, който може да шифрова, в зависимост от размера на ключа.

Сигурност на Ключовете: В реални приложения, управлението и съхранението на ключовете трябва да бъде извършвано с повишено внимание. Този експеримент служи като основа за разбирането на ключови концепции в криптографията и може да бъде разширен и адаптиран за по-сложни и сигурни приложения.

**Заключителни Мисли и Обобщение**

В този документ, ние изследвахме важни аспекти на темата "Обфускация на данни и киберсигурност", като разгледахме нейните ключови компоненти, алгоритми и приложения.

Въведение: Разгледахме основната идея зад обфускацията на данни и киберсигурността, като обърнахме внимание на съответната терминология, различни подходи и методи. Това постави твърда основа за разбирането на концепциите и тяхната практическа значимост.

Теория: Представихме детайлно описание на избрани алгоритми, като AES, RSA и SHA-256, и разгледахме тяхната реализация в контекста на основни инструменти за работа с данни. Това допринесе за по-дълбоко разбиране на техническите аспекти на обфускацията на данни и киберсигурността.

Пример: Представихме конкретен пример за задача, включваща шифроване и дешифриране с RSA алгоритъм. Обсъдихме данните, обработката на данни и моделирането на задачата, което демонстрира как теоретичните концепции могат да бъдат приложени на практика.

Експеримент: Проведохме експеримент, като използвахме разработените примери за клиент и сървър. Анализирахме резултатите и направихме изводи, които подчертават важността на криптографията и обфускацията на данни в защитата на информацията.

В крайна сметка, този документ осветли сложността и важността на киберсигурността и обфускацията на данни. Откроиха се както техническите аспекти, така и практическите приложения на тези концепции. Въпреки че разгледахме само част от темата, ясно стана, че обфускацията на данни и киберсигурността са от ключово значение в съвременния свят, където защитата на информацията е от първостепенно значение.