ТЕХНИЧЕСКИ УНИВЕРСИТЕТ – СОФИЯ
Факултет по Приложна математика и информатика

Маг. инж. Малинка Спасова Иванова

СИГУРНОСТ НА УЕБ УСЛУГИ – МЕТОДИ И СРЕДСТВА

А В Т О Р Е Ф Е Р А Т

на дисертация за придобиване на образователна и научна степен
"ДОКТОР"

Област: 4. Природни науки, математика и информатика

Професионално направление: 4.6. Информатика и компютърни науки

Научна специалност: Информатика

Научен ръководител: доц. д-р Анна Георгиева Розева

СОФИЯ, 2021 г.

Дисертационният труд е обсъден и насочен за защита от разширен научен съвет на катедра Информатика на ТУ-София на редовно заседание, проведено на 27.04.2021г.

Публичната защита на дисертационния труд ще се състои на 24.06. 2021г. от 15 часа в Конферентната зала на БИЦ на Технически университет – София на открито заседание на научното жури, определено със заповед № ОЖ-4.6-03 от 29.04.2021г. на Ректора на ТУ-София в състав:

1. Доц. д-р Десислава Антонова Иванова
2. Доц. д-р Анна Георгиева Розева
3. Проф. дн Кети Георгиева Пеева
4. Проф. д-р Мария Михайлова Нишева-Павлова
5. Доц. д-р Милен Йорданов Петров

Рецензенти:

1. Доц. д-р Десислава Антонова Иванова
2. Проф. д-р Мария Михайлова Нишева-Павлова

Материалите по защитата са на разположение на интересуващите се в канцеларията на Факултет по Приложна математика и информатика на ТУ-София, блок № 2, кабинет № 2228.

Дисертантът е задочен докторант към катедра Информатика на Факултет по Приложна математика и информатика. Изследванията по дисертационната разработка са направени от автора със съдействието и подкрепата на научния ръководител.

Автор: маг. инж. Малинка Иванова
Заглавие: Сигурност на уеб услуги – методи и средства

Тираж: 30 броя
Отпечатано в ИПК на Технически университет – София

I. ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност

Технологичното развитие на уеб пространството дава възможност за разработване на нови и по-усъвършенствани софтуерни приложения, използващи се в различни отраслови направления. Стремежът е насочен към технологично подпомагане на все повече човешки дейности, извършващи се през Интернет, което да доведе до бързи, резултатни и сигурни транзакции, трансфер на данни и пренос на електронни документи. Архитектурата, основана на услуги е често използвано решение при създаване на уеб приложения, поради голямата гъвкавост и възможност за разширяемост с необходимата конфигурация от уеб услуги. Това води до повишен интерес към тях и непрекъснато разработване с различна функционалност и приложимост. Наред с това се отчита и засилена активност от страна на злонамерени лица, които използват различни видове атаки, за да се възползват от нерегламентирано извършване на зловредни действия. Причинените поражения могат да бъдат малки, но могат да бъдат и огромни, засягащи голям брой служители и потребители. В резултат, обичайната дейност на редица организации може да бъде нарушена, забавена или направо спряна, като намирането на причината и начина на извършване на атаката, както и възстановителните процедури отнемат определено време и значителни ресурси. Изследване на методи и средства за защита на уеб услуги е темата на настоящия дисертационен труд, като акцентът е поставен върху разработване на методология за постигане сигурност на уеб услуги.

Дисертационната работа разглежда и анализира утвърдени на практика методи за защита на уеб услуги, както и търси нови техники, основани на алгоритми за машинно обучение и размита логика. Проведени са експериментални изследвания, като големи масиви от данни, събрани след работата на неатакувана и атакувана уеб услуга, са използвани за създаване на прогнозни и аналитични модели. Превенцията и защитата на уеб услуги е тема, която ще продължава да привлича вниманието на научната общност и на индустрията, както и ще се търси ролята на изкуствения интелект в областта на сигурността.

Цел и задачи

Въз основа на научни изследвания, теории и практически постановки в областта на сигурността на уеб услуги, целта на дисертационния труд е:

Да се разработи методология, осигуряваща сигурни уеб услуги и да се верифицира в определен контекст.

За постигане на поставената цел са дефинирани следните задачи за изпълнение:

1. Анализ на научни публикации и технически доклади в област, касаеща съвременното състояние на сигурността на уеб услуги, обобщаване на съществуващи методи и средства за постигане на определено ниво на сигурност на уеб услуги, както и идентифициране на нерешени проблеми.
2. Разработване на методология за постигане сигурност на уеб услуги въз основа на обобщение и анализ на състоянието до момента, както и дефиниране на критерии, влияещи върху уеб сигурността.
3. Разработване на експериментална среда за симулация на атаки към уеб услуги – идентификация на атаки, въздействия и проблеми за сигурността.
4. Разработване на методология за реализиране на сигурност на уеб услуга чрез прилагане на алгоритми и техники от машинното обучение и размитата логика.

В дисертационния труд са представени оригинални научни изследвания върху теории и съществуващи практики и получени резултати под формата на нови и адаптирани съществуващи знания, модели и решения в областта на сигурността на уеб услуги. Тези резултати са систематично приложени при разработване и верифициране на предложена *методология* за реализиране на сигурност на уеб услуги.

Дисертационният труд е организиран в увод, три глави и заключение с обобщени приноси, получени от представеното изследване.

Първа глава представя творчески обзор - обобщение и анализ на съвременното състояние в областта на сигурност на уеб услуги. Представена е архитектура на уеб услуги, обобщени са техните особености

и изискванията към тях от гледна точка на сигурността. Показани са класификации на заплахи и атаки срещу уеб услуги и са дискутирани както решения за защита на уеб услуги, предложени от изследователи и експерти по сигурността, така и трудни за решаване проблеми. Разгледани са и научни публикации, които предлагат и дискутират методи за изследване на уеб атаки чрез прилагане алгоритми от машинното обучение. Разгледани са механизми и алгоритми за постигане на сигурност на уеб услуги на XML ниво. Обобщени и систематизирани са възможни атаки на ниво XML според спецификациите XML Encryption и XML Signature. Показана е важността от прилагане на разработени спецификации и стандарти по сигурността, което значително може да подобри защитеността на уеб услуги. Главата завършва с разработен мета-модел на съвременни средства и методи за защита на уеб услуги.

Глава втора е фокусирана върху разработване на методология за постигане сигурност на уеб услуги. Изяснено е значението и съдържанието на понятието *методология* и е предложен модел на концептуална рамка за създаване на методологии. Извършено е проучване и сравнителен анализ на съществуващи методологии за проектиране и разработване на уеб услуги и на сигурни уеб услуги от гледна точка на инженерен софтуерен процес. Разработена е методология за постигане на сигурност на уеб услуги, като са взети под внимание значими фактори като: наличие на дейности по сигурността на инженерен софтуерен процес, коректно администриране на уеб услуги и предприемане на защитни мерки, допринасящи за постигане на високо ниво на сигурност, както и наличието и прилагането на възприети политики и стратегии по сигурността в организации.

В трета глава е извършена верификация на предложената методология за реализиране сигурност на уеб услуги. За тази цел е разбработена REST XML базирана уеб услуга, която е атакувана с често прилагана на практика и причиняваща значителни вреди атака *скрипт в уеб приложение* (*Cross-Site Scripting - XSS*). Разработена е *методология* за идентифициране на съхранена XSS атака и осигуряване на съответни защитни мерки на уеб услуга въз основа на използване на алгоритми за машинно обучение и техники от размитата логика. Извършено е експериментално изследване в специално организирана симулационна

среда, като са събрани данни за поведението на уеб услуга, когато е атакувана със съхранена XSS атака и когато не е атакувана. Така генерираните собствени набори от данни са преобразувани във формат за по-нататъшна обработка. Въведена е нова група параметри за избор на свойства, включваща характеристики на средата, в която се извършва XSS атаката. Създаден е *прогнозен модел* въз основа на класификационни алгоритми за машинно обучение за идентифициране на добре известни XSS атаки и *прогнозен модел* на базата на клъстерен алгоритъм за идентифициране на нови XSS атаки. Правила от теориите на размитите множества и размитата логика са приложени за решаване на многокритериална задача при определяне на мерките за сигурност при атакувана и неатакувана уеб услуга. Създаден е *аналитичен модел* на съхранена XSS атака чрез прилагане на техники, характерни за контролирано машинно обучение. Извършена е оценка на точността на разработените прогнозни и аналитични модели.

Научна новост

Като научна новост се отнасят следните предложени методи и методологии:

- *Мета-модел*, чрез който се концептуализира областта на сигурността на уеб услуги, включващ и методите и средствата за осигуряване на сигурни уеб услуги и посочващ възможностите на машинното обучение за подобряване нивото на сигурност;
- *Концептуален модел* на методология за сигурност на уеб услуги;
- *Методология* за сигурни уеб услуги с дефинирани стъпки и съответни методи, техники, процедури и инструменти за реализация;
- *Модел*, включващ дефиниран набор от критерии и техни параметри, чрез които се извършва оценка на нивото на уеб сигурност;
- *Обобщен* модел за приложение на машинното обучение и размитата логика в областта на уеб сигурността;
- *Модел* на XSS атака чрез машинно обучение с дървовидно-базиран алгоритъм Random Forest с използване на обобщен модел на идентифицираните атрибути;

- *Методология* за идентифициране на съхранена XSS атака и защита на REST уеб услуга чрез машинно обучение и размита логика;
- *Обобщен модел* на идентифицираните атрибути чрез машинно обучение;
- *Размити системи* за оценка на сигурност на XML уеб услуга по методите на Мамдани и Такаги-Сугено-Канг.

Публикации

Основни постижения и резултати от дисертационния труд са публикувани в 5 научни статии, от които три самостоятелни и две в съавторство. Две от статиите са публикувани в чужди списания, като едното е с IF 2.68. Две статии са представени на национални научни конференции, една от които е с международно участие и са публикувани в сборници с трудове. Една от статиите е докладвана на международна научна конференция в чужбина.

Структура и обем на дисертационния труд

Дисертационният труд е в обем от 129 страници, като включва увод, три глави, заключение, списък на основните приноси, списък на публикациите по дисертацията и използвана литература. Цитирани са общо 171 литературни източници, като три от тях са на кирилица, а останалите са на латиница. Работата включва общо 55 фигури и 14 таблици. Номерата на фигурите и таблиците в автореферата съответстват на тези в дисертационния труд.

Благодарности

Изказвам своята огромна благодарност на моя научен ръководител **доц. д-р Анна Розева** за цялостната подкрепа, направените препоръки, ценни съвети и насоки за провеждане на изследванията и прецизиране съдържанието на дисертационния труд.

II. СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

ГЛАВА 1. СИГУРНОСТ НА УЕБ УСЛУГИ – СЪСТОЯНИЕ, МЕТОДИ И СРЕДСТВА, НЕРЕШЕНИ ПРОБЛЕМИ

1.1. Въведение

Целта на първа глава е да анализира и обобщи настоящото състояние в областта на сигурността на уеб услуги, методите и средствата за постигане на високо ниво на сигурност, както и да се дискутират нерешени проблеми по сигурността на тези уеб услуги.

Дефинирани са следните задачи за постигане на поставената цел:

1. Проучване и систематизиране на факти, модели и концепции относно сигурността на уеб услуги и представяне на съществуващи класификации на съвременни заплахи и атаки.
2. Извършване на систематизирано описание, анализ и концептуализиране на заплахи и атаки на уеб услуги на различни нива и разглеждане на методи и средства за тяхното избягване, както и дефиниране на трудни за решаване проблеми по сигурността.
3. Проучване и систематизиране на стандарти за сигурност на уеб услуги от гледна точка цел на сигурността и защитавания компонент.
4. Проучване и анализ относно ролята на машинното обучение за идентифициране на атаки и защита на уеб услуги.

Първа глава е организирана в четири части.

Първа част представя основните характеристики на уеб услугите чрез архитектурата на уеб услуги с фокус върху концепции по сигурността. Обобщени са изискванията за сигурност на уеб услуги, като са систематизирани и показани графично.

Във втора част са разгледани няколко класификации на заплахи и атаки срещу уеб услуги според различни класификационни критерии и фактори. Въз основа на статистически анализи и технически доклади са систематизирани най-често прилаганите заплахи и атаки към уеб услуги и са обобщени графично. Представени са дефиниции на заплахи и атаки, дискутирани са предложени решения от изследователи и експерти по сигурността за защита на уеб услуги, както и трудности и нерешени въпроси

при реализиране на сигурност на уеб услуги. Посочени са възможни приложения на алгоритми и техники за машинно обучение за анализ, прогнозиране и защита на уеб услуги от различни по вид заплахи и атаки.

Третата част адресира спецификата и методите за постигане на сигурност на уеб услуги на XML ниво. Разгледани са механизми и алгоритми за криптиране и цифрово подписване на XML елементи при предаване на съобщения. Обобщени и систематизирани са възможни атаки на ниво XML според спецификациите XML Encryption и XML Signature.

Четвъртата част представя разработени спецификации и стандарти, чието прилагане значително може да подобри сигурността на уеб услуги, тъй като те обобщават най-добрите практики и съвременни решения за уеб сигурност. Разработен е мета-модел, който представя съвременното състояние относно средствата и методите за защита на уеб услуга от заплахи и атаки.

1.2. Обзор на главата

Уеб услугата се разглежда като софтуерен продукт с определена функционалност, който лесно може да се добавя, премахва или заменя в софтуерно приложение. Тя може да бъде част от архитектура, основаваща се на услуги, част от облачни структури или други решения, използващи услуги. Предназначението на уеб услугите е насочено към извършване на специфични действия, когато са извикани от уеб приложения. В зависимост от дизайна на интерфейса, уеб услугите се разделят на две групи: (1) SOAP (Simple Object Access Protocol) - и WSDL (WebService Definition Language) уеб услуги и (2) RESTful (REST - Representational State Transfer) уеб услуги.

Под заплахи и атаки срещу уеб услуги се разбират всички зловредни дейности на атакуващ, насочени срещу мрежовата инфраструктура, технологични решения, уеб приложения и хост компютри, както и всички умишлени или неумишлени дейности на софтуерните разработчици (по време на разработване на уеб приложения/уеб услуги), администратори (по време на администриране на компютри, сървъри, мрежи, софтуер и услуги) и крайни потребители (при използване на технологичната инфраструктура).

Всеки ресурс, който е уязвим и за него не са предвидени мерки за защита, представлява заплаха и опасност за организацията и може да бъде

атакуван. Атаката се свързва с вече започнала или осъществена зловредна дейност от страна на атакуващ, при което са нанесени щети – от въвеждане в заблуждение и затрудняване работата на служители и потребители до модифициране и разрушаване на информация.

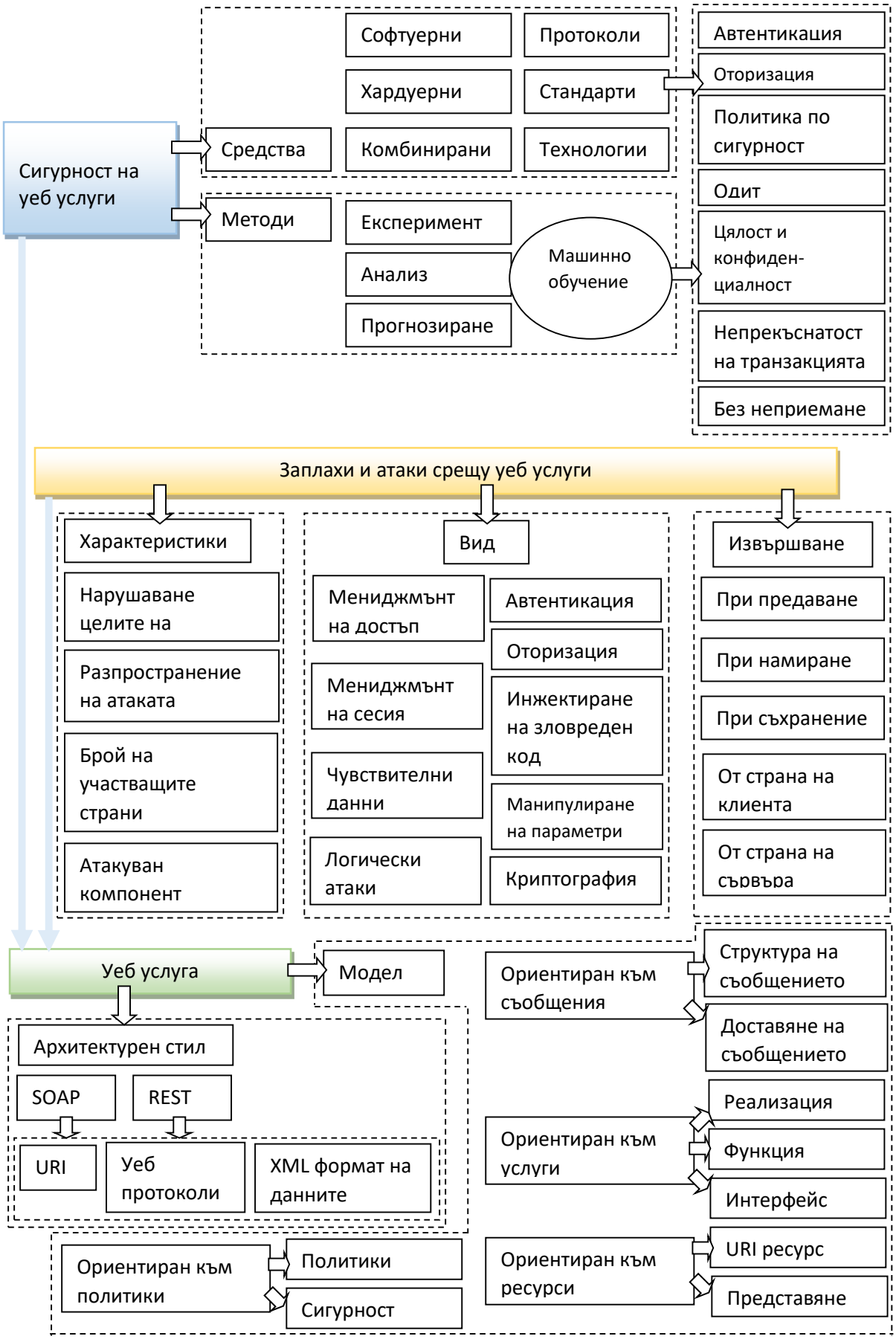
Направени са различни класификации на заплахи и атаки срещу уеб услуги в зависимост от определени признаци, критерии и фактори. Ако трябва да се обобщи, то може да се каже, че атаките се разглеждат и анализират в зависимост от това каква цел на сигурността е нарушена, какъв е броят на атакуващите страни, кой компонент на уеб услугата е атакуван и как се разпространява атаката.

Според технически доклад на Open Web Application Security Project (OWASP) най-прилаганите понастоящем атаки са: инжектиране на зловреден код, нарушаване на автентикацията, нарушен контрол на достъпа, скрипт в уеб софтуер XSS, злоупотреба с XML парсер (XML External Entities (XXE)), разкриване на чувствителни данни, други.

XML играе много важна роля при представяне и обмен на данни в публични мрежи. Това изисква осигуряване на безопасно предаване на XML. Съществуват редица технологии като: XML цифрово подписване, XML криптиране, инфраструктура с публични ключове и други, които спомагат за опазване секретността на информацията. Тези технологии търпят развитие. Много от тях са се превърнали в стандарти, които се прилагат върху различни технологични платформи.

Проучването в първа глава показва, че уеб услугите могат да бъдат сигурни, когато са приложени съответни методи, механизми или стратегии за идентифициране, превенция и защита срещу заплахи и атаки. Вижда се големият брой от възможни заплахи и атаки срещу уеб услуги, който продължава да расте с развитие на технологиите, както и възможни методи и средства за защита, включително техники и алгоритми от областта на машинното обучение.

Следването на препоръките, описани в спецификациите и стандартите по сигурността, може да допринесе за изграждане на инфраструктура, даваща възможност за подобряване на сигурността, както и за създаване на подходящи политики при проектиране и разработване на сигурни уеб услуги.



Фигура 7. Мета-модел на съвременни средства и методи за защита на уеб услуги

Анализът в първа глава на съвременното състояние в областта на сигурността на уеб услуги е концептуализиран чрез разработване на мета-модел, който представя характеристиките, вида и подходите за атакуване на уеб услуги, както и възможните методи и средства за защита (Фигура 7). Показана е значимостта на изкуствения интелект (машинното обучение) за анализ и прогнозиране на заплахи и атаки, както и за идентифициране на съответстващи защитни мерки.

Изводи:

1. Сигурността на уеб услуги е проблем, който е разработван усилено в научната литература от гледна точка на откриване на зловредни атаки и на превенция и защита. Този извод потвърждава актуалността на предмета на изследване на дисертационния труд.
2. Проучването показва, че защитата на XML съобщения е изключително важен компонент за реализиране на сигурност. Следователно изследвания и резултати в тази област ще носят приносен характер за науката и практиката.
3. Налице са изследвания и приложения за прилагане на методи на машинно обучение към решаване на проблеми на сигурността на уеб услуги. По-нататъшно развитие и адаптиране на този подход би било перспективно и с потенциал за тяхното обогатяване с нови приложения и доказана резултатност.

Във връзка с първа глава са публикувани статия и доклад, съответно в International Journal of Computers & Technology и в сборника с трудове от научна конференция с международно участие “Компютърни науки и технологии“ в България.

ГЛАВА 2. РАЗРАБОТВАНЕ НА МЕТОДОЛОГИЯ ЗА ПОСТИГАНЕ НА СИГУРНОСТ НА УЕБ УСЛУГИ

2.1. Въведение

Целта на втора глава е разработване на методология за постигане на сигурни уеб услуги въз основа на извършено изследване на съществуващи научни теории и факти в областта на концептуално и контекстно проектиране на методологии.

За реализиране на поставената цел са дефинирани следните задачи:

1. Проучване и систематизиране на концептуалната основа за разработване на методологии и създаване на модел на концептуална рамка за разработване на методология.
2. Достигане до нови познания и интерпретирането им чрез по-нататъшно развитие и допълнение на съществуващи методологии за разработване на уеб услуги и на сигурни уеб услуги.
3. Създаване на методология за постигане на сигурност на уеб услуги като се отчетат нови значими фактори, оказващи влияние върху сигурността, въз основа на съвременни научни изследвания и най-добра практика.

Глава втора е структурирана в две части.

В първа част се изяснява значението и съдържанието на понятието *методология*, разглежда се методология за създаване на методологии като комплексен процес, състоящ се от последователни стъпки. Въз основа на тези проучвания е предложен модел на концептуална рамка за създаване на методологии.

Във втора част на базата на анализ на съществуващи методологии за проектиране и разработване на уеб услуги и на сигурни уеб услуги като инженерен софтуерен процес е разработена методология за постигане на сигурност на уеб услуги. Взети са под внимание значими фактори от софтуерното инженерство, оказващи влияние върху процеса по проектиране и разработване на сигурни уеб услуги, дейности по администриране на уеб услуги, допринасящи за постигане на високо ниво на сигурност и значимостта на възприети политики и стратегии за осъществяване на мерки по сигурността от страна на определена организация.

2.2. Обзор на главата

Изследванията направени в тази глава се използват за проектиране на методология за реализация на сигурни уеб услуги. Той обобщава спецификата на три области: разработване на сигурен програмен код, вземане мерки за сигурност при администриране на уеб услуги и политиките по сигурност, възприети в дадена организация. Методологията

за сигурни уеб услуги представя съществуващи теории и практики с цел подобряване на познанието, след обобщаване и анализиране на въпроси от следните области: 1). Разработване на програмен код, по такъв начин, че да се елиминират грешки и бъгове, както и настъпването на непредвидени събития, което може да бъде предпоставка за появата на уязвимост или заплаха, както и да улесни осъществяването на атаки.; 2). Коректно конфигуриране и администриране на уеб средата, сървърите и приложенията с минимален брой грешки от страна на ИТ администраторите; 3). Наличие на политики в организациите, които обхващат стандарти и правила за определена култура и поведение от страна на служителите, както и стандарти, свързани със сигурността.

Методологията за сигурни уеб услуги включва методи, процедури, техники и инструменти, засягащи:

(1) Дейности за осигуряване на периодична проверка и контрол на генерирания програмен код, неговата реализация и внедряване от гледна точка на сигурността, както и дейности по подготовка на модели на възможни заплахи и атаки и препоръки за използване на софтуера;

(2) Съответстващо администриране на уеб средата на дадена организация и на мрежовата инфраструктура;

(3) Политики по сигурността, включващи клаузи по превенция и защита от заплахи и атаки, както и мениджмънт на риска в организацията в случаи на атаки.

Методологията за сигурни уеб услуги е логически организирана в стъпки, подстъпки и възможни методи, процедури, техники и инструменти, необходими за тяхното осъществяване (Таблица 7).

Първата стъпка е сканиране на уеб и мрежовата среда на организацията и установяване на коректно и/или некоректно работещи услуги, приложения и бази данни, както и определяне спецификата на мрежовия трафик. Също така е необходимо да бъдат разгледани съществуващите политики и стратегии по сигурността, за да се установи настоящото положение в организацията.

Втората стъпка включва определяне на нуждите и приоритетите на организацията, като се отчетат установените уязвимости, заплахи и атаки в контекста на технически и нетехнически въпроси, определяне и разбиране

на вътрешната и външната комуникация, определяне на отговорните и заинтересованите лица, както и ограниченията и границите на уеб средата на организацията.

Третата стъпка представлява изясняване на проблемите по сигурността – идентифициране на източниците на тези проблеми и разбиране на тяхната същност, подготовка на списък с приоритизирани проблеми по сигурността, ако е възможно класифициране и групиране на проблемите, за да бъде улеснено решаването им.

Четвъртата стъпка е фокусирана върху обобщаване и анализиране на цялата информация, събрана в предните стъпки и изработване на сценарий, показващ уязвимости, заплахи, атаки и рискове за организацията. Съгласно този сценарий, трябва да се изготвят препоръки за осигуряване или подобряване на сигурността (пета стъпка).

В шестата стъпка, създаденият сценарий трябва да бъде подобрен, като се вземат под внимание приетите препоръки от предната стъпка.

В седмата стъпка, сценарият трябва да бъде тестван и верифициран в средата на организацията. Ако се установят някакви отклонения, предишните стъпки трябва да бъдат извършени отново.

Таблица 7. Методологията за реализация на сигурни уеб услуги

Стъпки	**Подстъпки**	**Методи, процедури, техники, инструменти**
Сканиране средата на организацията	• Идентифициране на уеб услуги, които се използват коректно, но предизвикват неочакван резултат • Идентифициране на некоректно администриране на услугите • Сканиране на мрежовия трафик и потребителските акаунти • Сканиране на използваните приложения и бази данни	• Процедура за докладване на бъгове и неподходящ код • Техники и инструменти за мрежово сканиране и мониторинг

	• Разбиране наличието на политики по сигурността	
Определяне на потребностите	• Определяне на уязвимости, възможни заплахи и атаки • Разбиране на техническите въпроси • Разбиране на нетехническите въпроси • Разбиране на вътрешната и външната комуникация • Идентифициране на отговорни и заинтересовани лица • Идентифициране на ограниченията	• Методи и инструменти за определяне на възможни заплахи • Процедури за разбиране на техническите и нетехническите въпроси • Процедури за идентифициране на отговорностите и техните граници
Идентифициране на проблеми по сигурността	• Идентифициране на източника и естеството на проблемите • Приоритизиране на проблемите • Класифициране и групиране на проблемите	• Метод и процедура за идентифициране на проблемите • Метод и процедура за приоритизиране и класифициране на проблемите
Създаване на сценарий	• Моделиране на текущата ситуация – заплахи, атаки, рискове и влиянието им върху организацията	• Метод и инструмент за моделиране на заплахите • Метод за оценка на риска
Създаване на препоръки	• Свързване на проблемите с препоръките • Приоритизиране на препоръките • Приемане или отхвърляне на препоръките	• Метод за създаване на препоръки и тяхното приоритизиране • Процедура за приемане/отхвърляне
Създаване на нов сценарий, включващ въпроси по сигурността	• Елиминиране на проблеми, свързани с приетите препоръки	• Процедура за създаване на нов сценарий
Верифициране на сценария	• Верифициране на създадения нов сценарий	• Процедура за верификация

Методологията е проектирана на концептуално ниво и затова по-надолу в тази глава тя се детайлизира чрез разработен модел, включващ дефиниран набор от критерии и техни параметри. Критериите и параметрите допринасят за извършване оценка на нивото на уеб сигурност. Те са показани на Фигура 16 под формата на йерархични зависимости. Глобалният критерий GC: Сигурност на уеб услуги и уеб приложения се състои от три критерия: C1: Сигурен програмен код, C2: Коректно администриране и C3: Политика по сигурността. Критерий C1 включва осем параметъра: P01÷P08, критерий C2 се формира от девет параметъра: P09÷P17 и критерий C3 се определя от шест параметъра: P18÷P23.

Коректен код – Грешки в програмния код могат да доведат до несигурен софтуер, да се даде възможност за изпълнение на зловреден код, да се достигне до неочаквано поведение на софтуера или неочакван резултат. Лошо софтуерно изпълнение води до уязвимости, заплахи и атаки.

Процес за разработка с мерки за сигурност – Процесът по разработка на софтуер трябва да включва дейности по сигурността, например при определяне изискванията към софтуера трябва да е предвидена дейност по сигурността, чрез която да се дефинират целите на сигурността или при разработването на софтуера да се извършва проверка и контрол върху кода относно сигурността.

Криптиране и цифрово подписване – Използване на криптографски алгоритми за криптиране на кода с цел невъзможност за достъп до него от неоторизиран потребител. Може да се използва и цифрово подписване на кода, за да се удостовери авторът на софтуера и да се гарантира, че програмният код не е подменен или променен.

Сигурно внедряване – Внедряването на кода да бъде извършено в сигурна среда с контрол върху достъпа, за да не се допусне неоторизирано вмешателство.

Списък с бъгове – Да бъде налице списък с установени бъгове в софтуера, с което да се алармират разработчиците за отстраняването им, както и да се даде възможност за уведомяване на потребителя.

Прилагане на стандарти за сигурност – За да се избегнат възможни рискове в сигурността и да се минимизират заплахите и атаките, се прилагат

стандарти за сигурност, които позволяват прилагане на техники, методи и алгоритми, касаещи сигурността.

Тест за сигурност – Тестовете относно сигурността на разработения софтуер водят до откриване на бъгове, грешки в кода, неочаквано поведение или резултат.

Упътвания относно сигурността – Софтуерната реализация на уеб услугата да бъде придружена с упътвания и препоръки за сигурно доразработване, внедряване, администриране и използване.

Софтуер – При администриране на уеб услугата да се идентифицира дали е сигурна – дали не води до компрометиране на базите данни, други уеб услуги или уеб приложения, до използване без необходимото разрешение на лични данни, до компрометиране и застрашаване сигурността на онлайн средата като цяло.

Транзакции – При предаване на информация и лични данни да се установи защитеността на трансмисионния канал между клиента и сървъра – криптиране на данните и кода и/или използване на цифрово подписване, използване на защитени уеб протоколи.

Комуникационни съобщения – Установяване сигурността на протоколите за предаване на съобщение при функциониране на уеб услуги.

Лични данни – Трябва да се установи дали уеб услугата използва или обработва лични данни и да се определи какви точно.

Конфигуриране на сигурността – Конфигурирането на сигурността на уеб услугите да се извърши на три нива: на ниво SOAP съобщение, в което данните са цифрово подписани или криптирани (могат да се включат и токъни за автентикация); на транспортно ниво - използване на SSL за постигане на сигурна връзка между клиентското приложение и уеб услугите; контрол на достъпа – осигуряване на различни нива на достъп до уеб услугите в зависимост от ролята на потребителя или група потребители.

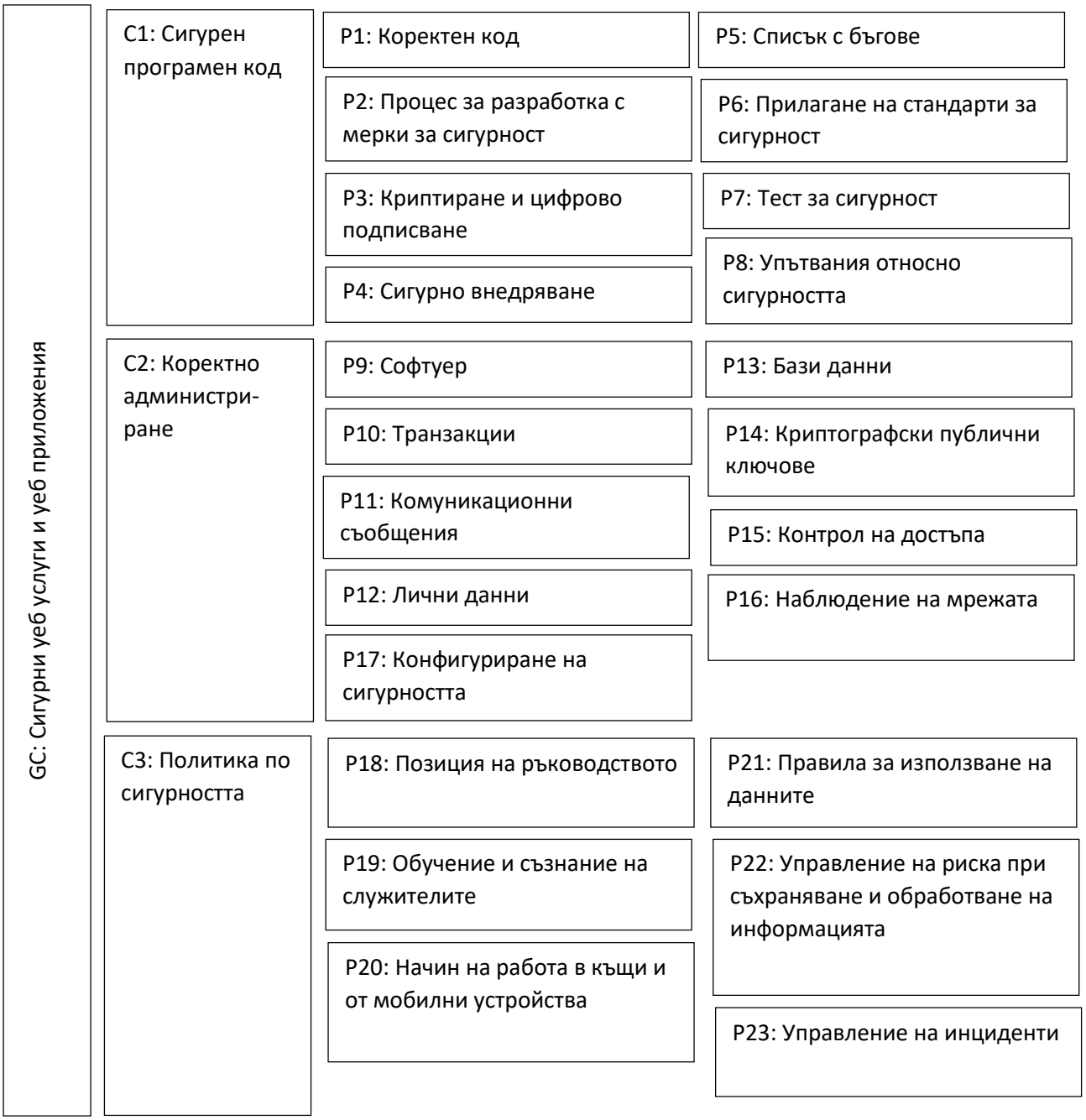
Бази данни – При достъп на уеб услуга до бази данни, сигурността на базата данни зависи от сигурността на уеб услугата. Ако уеб услугата не е сигурна (например съдържа зловреден код), тогава и базата данни е уязвима.

Криптографски публични ключове – Да се установи дали е гарантирана сигурността на криптографските публични ключове.

Контрол на достъпа – Да се установят нива за достъп до уеб услуги на оторизирани потребители. Да се следи за неоторизиран достъп.

Наблюдение на мрежата – Конфигуриране на защитната стена, проследяване на трафика и получаване на разбиране относно правилното функциониране на уеб услугите.

Начин на работа в къщи и от мобилни устройства – Необходими са инструкции или правилник за сигурна работа в организационната онлайн среда, когато служителите се намират извън сградите на организацията.



Фигура 16. Модел с критерии и параметри за оценка сигурността на уеб услуги и уеб приложения

Правила за използване на данните – Проследяване и спазване на определени правила и/или политики за опериране с данни на организацията или с лични данни на служителите.

Управление на риска при съхраняване и обработване на информацията – Идентифициране на рискове, съотносими към организираната онлайн среда и предприемане на технически и организационни мерки за реализиране на уеб сигурност на високо ниво.

Управление на инциденти – При възникнали инцидентни ситуации, за които не са създадени правила или политики, да се реагира своевременно и релевантно с цел постигане на сигурност в краткосрочен и дългосрочен аспект.

Позиция на ръководството – Ръководството на организацията трябва да има изградено виждане относно важността на сигурността на уеб средата, която използва и да съдейства за въвеждане на правила и политики за постигане на високо ниво на сигурност.

Обучение и съзнание на служителите – Провеждане на обучения, семинари и срещи сред служителите на организацията с цел разясняване на политики и практики в областта на уеб сигурността.

Изводи:

1. На база на задълбочено проучване и анализ на методологии за разработване на софтуер, за които е установено не достатъчно застъпване на мерки за сигурност е съставен *концептуален модел* на методология за сигурност на уеб услуги.
2. На базата на представения концептуален модел е разработена методология за сигурни уеб услуги с дефинирани стъпки и съответни методи, техники, процедури и инструменти за реализация.
3. Разработената методология е детайлизирана чрез модел, включващ дефиниран набор от критерии и техни параметри, чрез които се извършва оценка на нивото на уеб сигурност.

Резултатите от тази глава са обобщени в статия, публикувана в Journal of Advances in Electronics and Computer Science с IF 2.68 и в доклад, представен на Национална конференция по информатика, посветена на 80г. от рождението на професор Петър Бърнев.

ГЛАВА 3. ВЕРИФИЦИРАНЕ НА МЕТОДОЛОГИЯТА ЗА ИДЕНТИФИЦИРАНЕ НА XSS АТАКА И ЗАЩИТА НА УЕБ УСЛУГИ – ПРАКТИЧЕСКА РЕАЛИЗАЦИЯ ЧРЕЗ МАШИННО ОБУЧЕНИЕ И РАЗМИТА ЛОГИКА

3.1. Въведение

Целта на трета глава е да се представи практическа реализация на разработена методология за идентифициране на постоянна (съхранена) атака *скрипт в уеб приложение* (*Cross-Site Scripting – XSS*) и защита на уеб услуги чрез прилагане на техники и алгоритми от областта на машинното обучение и теориите на размитите множества и размитата логика.

За осъществяване на поставената цел са дефинирани следните задачи:

1. Проучване, анализ и концептуализиране на възможностите на машинното обучение и теориите на размитите множества и размитата логика за идентифициране на XSS атаки и за посочване на съответни защитни мерки.
2. Разработване на REST XML уеб услуга с JAVA API, която да бъде атакувана с XSS атака, подготовка на симулационна среда за тестване на уеб услугата и атаката и създаване модел на постоянна XSS атака.
3. Разработване на аналитичен и прогнозен модел за защита на уеб услуга чрез решаване на класификационна и клъстерна задача чрез машинно обучение, както и решаване на многокритериална задача чрез размити множества и размита логика.

Трета глава е организирана в четири части, както следва:

В първа част е представен анализ и обобщение на съвременни публикации на научната общност относно прилагане на алгоритми и техники от машинното обучение и размита логика в областта сигурност на уеб услуги и по-конкретно за идентифициране на XSS атаки и защита на уеб услуги. Разработен е обобщен модел на приложението на машинното обучение и размитата логика в областта на уеб сигурността.

Втора част разглежда теоретични постановки, свързани с алгоритми за машинно обучение и техни параметри, както и дефиниции, използвани за изграждане на размити системи и достигане до решение/заключение чрез прилагане теориите на размитите множества и размитата логика. Предложена е методология за разработване на прогнозни и аналитични

модели чрез машинно обучение. Разгледана е и методология за разработване на размита система.

Трета част представя разработена REST уеб услуга и нейната функционалност. Описана е изградената симулационна среда, в която е тествана работоспособността на уеб услугата, когато не е атакувана и когато е атакувана чрез съхранена XSS атака.

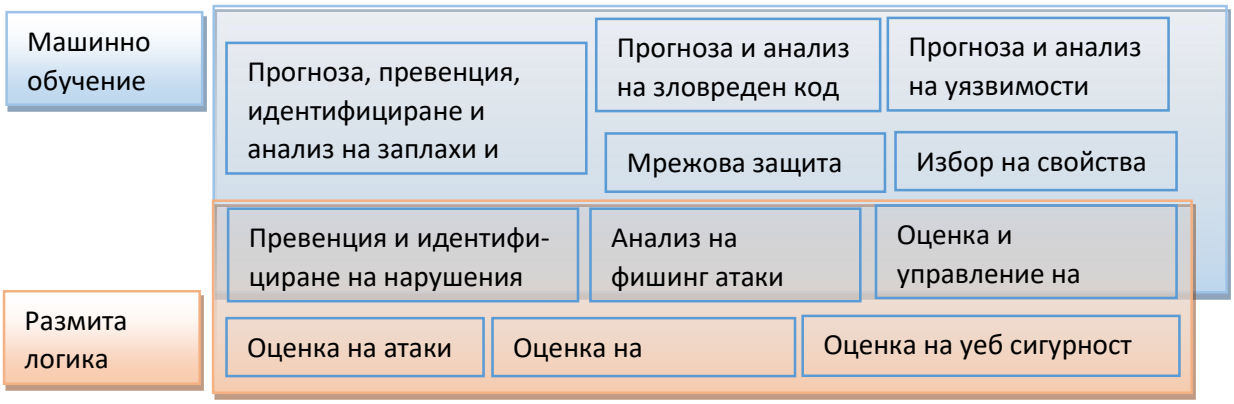
Четвърта част включва представяне на разработен модел на съхранена XSS атака чрез методите на машинното обучение, както и резултатите от извършена оценка на модела. Представена е методология за идентифициране на съхранена XSS атака и защита на REST уеб услуга. Показани са разработени модели за избор на атрибути, за идентифициране на съхранена XSS атака, за идентифициране на нови атаки чрез клъстериране, както и модели на размити системи, чрез които се решава многокритериална задача.

3.2. Обзор на главата

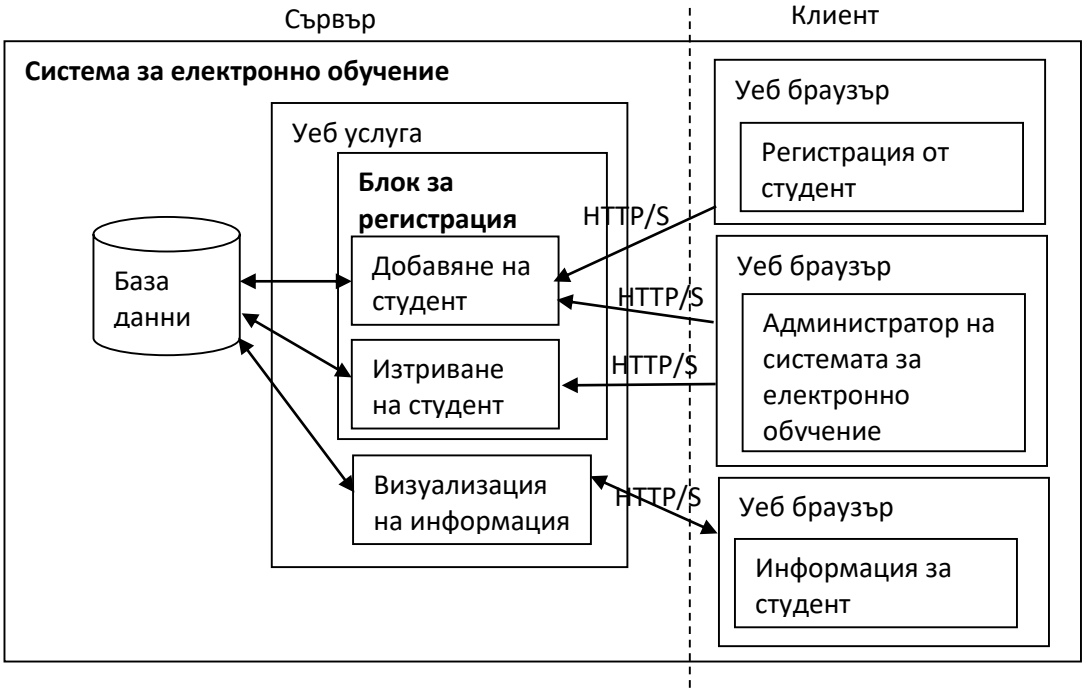
В тази глава е представена разработена методология за откриване на съхранена XSS атака и защита на уеб услуга, По-конкретно е разгледана атаката *скрипт в уеб приложение* (*Cross-Site Scripting (XSS)*), тъй като тя е една от най-често срещаните в съвременното уеб пространство, причиняваща много големи вреди, като може да засегне огромен брой потребители. Атаката XSS се характеризира с трудна откриваемост и разрушителни последици и затова се търсят различни методи за идентифициране, предпазване и защита.

Въз основа на направения литературен обзор е създаден обобщен модел на приложението на машинното обучение и размитата логика в областта на уеб сигурността (Фигура 21).

Разработената XML уеб услуга е с REST архитектурен стил и с JAVA API. Използва се в система за електронно обучение, чиято техническа архитектура е клиент/сървър. Уеб услугата функционира чрез протокола HTTP/HTTPS, като подаваната заявка и получаваният отговор са в XML формат. Функционалността на REST уеб услугата включва: регистриране и премахване на потребител в системата и получаване на информация за потребител (Фигура 28). В резултат от работата й върху уеб страница се визуализира исканата информация за потребител на системата.



Фигура 21. Обобщен модел на приложението на машинното обучение и размитата логика в областта на уеб сигурността



Фигура 28. Функционалност на REST уеб услуга

В подготвената тестова среда се проверява работоспособността на разработената REST уеб услуга, когато не е атакувана и когато е атакувана чрез съхранена XSS атака. Организираната тестова среда включва:

(1) Софтуерът IntelliJ IDEA за разработване и стартиране на Java уеб услугата и уеб сървъра;

(2) Платформата Postman, използвана като виртуален клиент, изпълняващ заявки за създаване или премахване на потребителски профил;

(3) Уеб браузър за стартиране и визуализиране резултата от потребителската заявка.

Разработен е ***модел на съхранена XSS атака***, като за първи път това е извършено чрез прилагане на дървовидно-базиран алгоритъм за контролирано машинно обучение Random Forest. Този модел е подобрен и оценен. Резултатът от оценката е показан чрез стойностите на параметрите, представени в Таблица 8. Постигната точност на модела е висока 98.71%, а абсолютната грешка е малка 0.301+/-0.102.

Таблица 8. Параметри на подобрения модел на съхранена XSS атака

	Random Forest
Accuracy	98.71%
Classification error	1.29%
Kappa	0.981
Weighted mean recall	98.61%
Weighted mean precision	98.72%
Absolute error	0.301+/-0.102
Relative error	30.13%+/-10.22%
Root mean squared error	0.318

Разработена е ***методология***, която дава възможност да се извършат изследвания за идентифициране на съхранена XSS атака и защита на REST уеб услуга, която се основава на прилагане на методи и техники от машинното обучение и теориите на размитите множества и размитата логика (Фигура 37). Методологията включва следните процедури:

(1) Формиране на набори с данни въз основа на експерименти в симулационна среда, когато уеб услугата не е атакувана и когато е атакувана;

(2) Избор на релевантни атрибути, които се използват за създаване на прогнозни и аналитични модели;

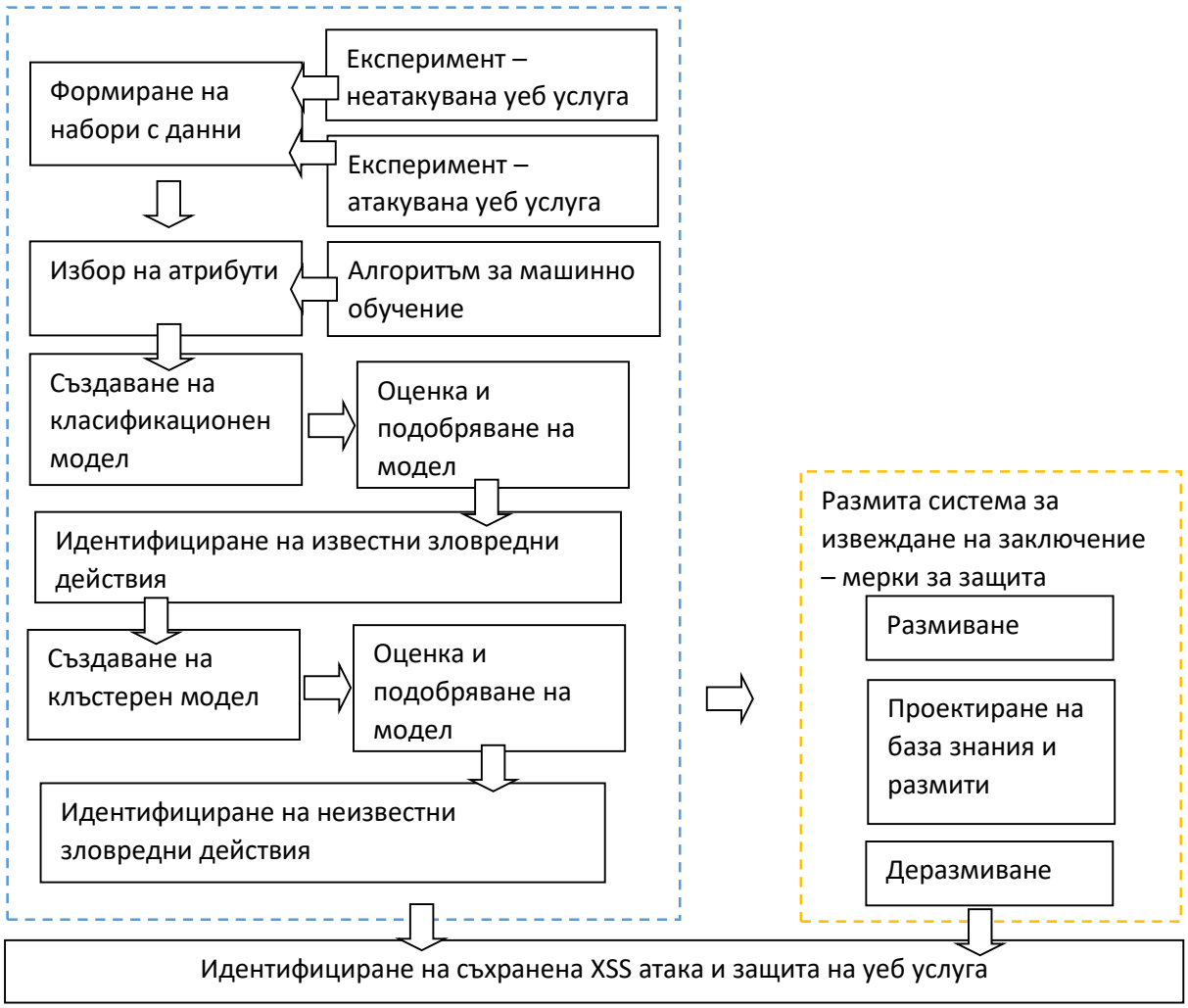
(3) Разработване на модели – класификационен и клъстерен за идентифициране на съхранена XSS атака;

(4) Оценка и подобряване на моделите;

(5) Създаване на размита система за извеждане на решение/извод за степента на повредите в следствие на атаката и възможни мерки за защита.

Изборът на атрибути е процес, в който се извършва редуциране на атрибути, ръчно или автоматизирано, с цел достигане до най-релевантните атрибути за моделите, които се създават и за проблемите, които се решават. В дисертационния труд е възприет автоматизиран подход за избор на най-важните атрибути чрез прилагане на алгоритъм за машинно обучение.

Наборите с данни са формирани, след като данните са събрани въз основа на проведени експерименти в тестовата среда в два случая: (1) когато уеб услугата не е атакувана и работи нормално и (2) когато уеб услугата е атакувана чрез съхранена XSS атака. За улесняване процеса по избор на атрибути, всички идентифицирани атрибути са класифицирани в три групи: (1) структурни; (2) поведенчески; (3) дължащи се на средата. За първи път в този дисертационен труд е въведена третата група атрибути, свързани със свойства, дължащи се на средата. Чрез тях се описват и се вземат под внимание важността на свойствата на средата от страна на клиента и от страна на сървъра, където XSS атаката е идентифицирана.



Фигура 37. Методология за идентифициране на съхранена XSS атака и защита на REST уеб услуга

Идентифицираните 171 атрибути и тяхната взаимосвързаност е показана чрез създаден обобщен модел, прилагайки алгоритъм за контролирано машинно обучение RandomForest. Наборът с данни съдържа

511 записа. 70% от набора с данни е използван за обучение, а 30% от него за тестване. Извършена е оценка на класификатора, като резултатът е обобщен в Таблица 9.

Таблица 9. Параметри на модел с идентифицирани атрибути

	Random Forest
Accuracy	89.04%
Classification error	10.96%
Kappa	0.216
Weighted mean recall	56.82%
Weighted mean precision	94.42%
Absolute error	0.142+/-0.214
Relative error	14.19%+/-21.42%
Root mean squared error	0.257

За да бъдат избрани най-важните от тези 171 свойства, е приложен алгоритъмът *Определяне на тегловни коефициенти* (Weight by Information Gain) в RapidMiner Studio. За по-нататъшно машинно обучение са избрани 30 атрибута от трите групи.

Избраните свойства и свързаните с тях данни са използвани като вход към контролирано машинно обучение, като са приложени следните дървовидно-базирани алгоритми: Random Forest, Random Tree, Decision Tree и Gradient Boosted Tree.

Използваният набор с данни за машинно обучение съдържа 4116 записа, включващи данни както за нормално функциониране на REST уеб услугата, така и когато е атакувана със съхранена XSS атака. Извършена е оценка на класификаторите, като резултатът е показан в Таблица 10.

Таблица 10. Оценка на класификаторите

	Random Forest	**Random Tree**	**Decision Tree**	**Gradient Boosted Tree**
Accuracy	100%	100%	93.52%	100%
Classification error	0%	0%	6. 84%	0%
Kappa	1	1	0.824	1
Weighted mean recall	100%	100%	88.20%	100%
Weighted mean precision	100%	100%	95.90%	100%
Absolute error	0.004+/-0.004	0.000	0.127+/-0.208	0.149+/-0.068
Relative error	0.41%+/-0.44%	0%	12.74%+/-20.82%	14.88%+/-6.83%
Root mean squared error	0.006	0.000	0.244	0.164

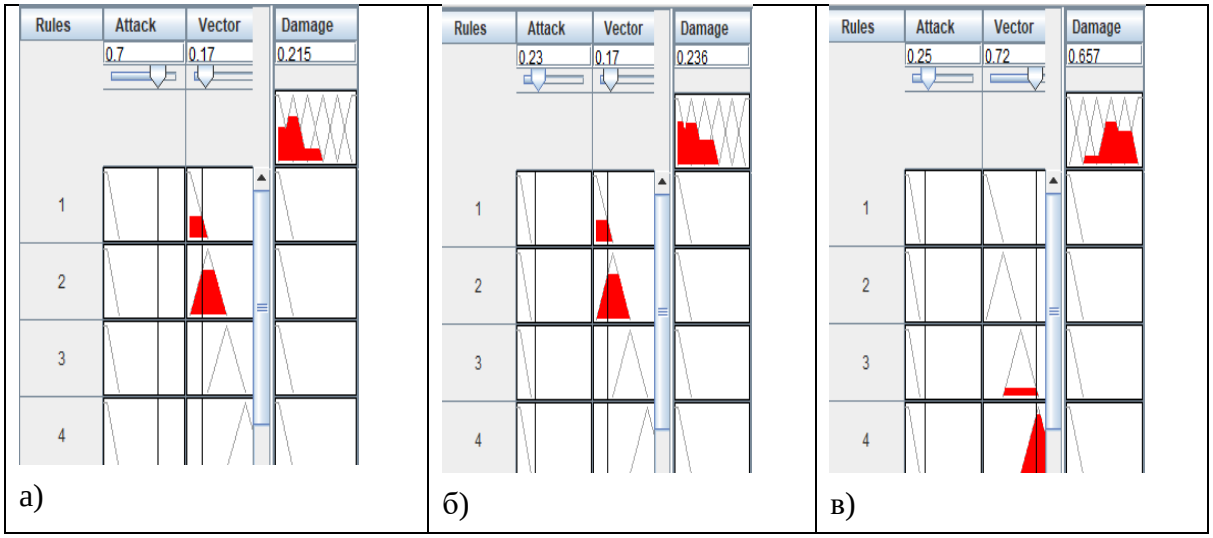
За целите на това изследване е приложен алгоритъмът за клъстериране k-Means за прогнозиране на нови атаки. Създаденият клъстерен модел използва разширен набор с данни. Извършени са редица

експерименти при различна стойност на предварително зададения брой на клъстерите k, за да се избере най-подходящата стойност за по-нататъшни изследвания.

Теориите на размитите множества и размитата логика са приложени за решаване на многокритериален проблем. За тази цел, получените резултати от процесите класификация и клъстериране, както и избрани свойства от групата на свойствата, дължащи се на средата, са използвани за построяване на размита система, извеждаща заключения. Размитият модел е прогнозен и въз основа на вероятност за настъпване на атака и доколко тя е опасна за потребителя прогнозира степента на повреденост. За реализиране на размитата система е избран методът на Мамдани.

Верификация на създадения модел на интелигентната размита система е извършена в средата на софтуера FisPro, като изходното заключение в неговата размита и точна форма, след като са взети под внимание всички размити правила, са представени на Фигура 48.

Фигура 43. Функционалност на размитата система



Фигура 48. Изходно заключение при различни стойности на входните размити променливи

Изводи:

1. Създадена е симулационна среда за идентификация на атаки срещу XML уеб услуга тип REST.
2. Проектирана е методология за идентифициране на атака и защита на XML уеб услуга.
3. Създаден е модел за откриване на XSS атака чрез класификация с насочено машинно обучение.
4. Проектирана е размита система за оценка на сигурност на XML уеб услуга с интерпретация в работна среда.
5. Практически е верифицирана общата методологията за сигурност на уеб услуги с илюстративни примери.
6. Извършен е сравнителен анализ между метода на Мамдани и метода на Такаги-Сугено-Канг в областта на сигурност на уеб услуги.

Резултатите, представени в тази глава, са обобщени в статията “Detection of XSS Attack and Defence of REST Web Service - Machine Learning Perspecive”, докладвани на 5th International Conference on Machine Learning and Soft Computing (ICMLSC 2021), Виетнам (виртуална конференция).

Приноси

Научни и научно-приложни приноси

1. ***Научни и научно-приложни приноси***
   1.1. Разработен е мета-модел на съвременни средства и методи за защита на уеб услуги.
   1.2. Създаден е концептуален модел на методология за реализиране на сигурни уеб услуги.
   1.3. Разработен е модел с критерии и параметри за оценка сигурността на уеб услуги и уеб приложения.
   1.4. Създаден е обобщен модел за приложение на машинното обучение и размитата логика в областта на уеб сигурността.

2. Приложни приноси

2.1. Създадена е симулационна среда за идентифициране на атаки срещу уеб услуги и защита.

2.2. Разработен е модел на XSS атака и практическа реализация с дървовидно-базиран алгоритъм Random Forest в платформа за машинно обучение с използване на обобщен модел на идентифицираните атрибути.

2.3. Разработена е размита система за оценка на сигурността на REST уеб услуга.

Списък на публикациите по дисертационния труд

- **M. Ivanova**, Security of Web Services: Methods and Contrivance, International Journal of Computers & Technology, 14(11), 2015, 6229-6239, doi: 10.24297/ijct.v14i11.6374, ISSN 2277-3061.
- **М. Иванова**, Уеб услуги: сигурност на ниво XML, Трета научна конференция с международно участие "Компютърни науки и технологии" 25-26 септември, 2015 г. Варна, България, 2015, 107-114, ISSN 1312-3335.
- **М. Иванова**, Методологии за проектиране и разработване на сигурни уеб услуги: съвременно състояние, Национална конференция по информатика, посветена на 80 г. от рождението на професор Петър Бърнев, София, 2016, ISBN 978-954-8986-45-8111-120, http://www.math.bas.bg/infres/cib80/book/cib80-p13.pdf.
- **M. Ivanova**, A. Rozeva, Methodology for Realization of Secure Web Services, International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835, 4(8), 2017, 57-62, IF 2.68.
- **M. Ivanova**, A. Rozeva. 2021. Detection of XSS Attack and Defense of REST Web Service – Machine Learning Perspective. In 2021 The 5th International Conference on Machine Learning and Soft Computing (ICMLSC'21), January 29-31, 2021, Virtual Event, Vietnam. ACM, New York, NY, USA, 13 Pages.

SUMMARY

WEB SERVICE SECURITY - METHODS AND TOOLS

MSC ENG. MALINKA SPASOVA IVANOVA

The technological development of the web space provides for the development of new and more advanced software applications used in various industries. The aim is to provide technological support to more and more human activities carried out over the Internet, leading to fast, efficient and secure transactions, data transfer and transmission of electronic documents. The service-based architecture is a commonly used solution when creating web applications, due to the great flexibility and scalability with the required configuration of web services. This leads to increased interest in them and their continuous development with different functionality and applicability. In addition, there is increased activity by malicious individuals who use various types of attacks to take advantage of unregulated malicious actions. The damage caused can be small, but it can also be huge, affecting a large number of employees and users. As a result, the normal activities of a number of organizations can be disrupted, delayed or even stopped, and finding the cause and manner of the attack, as well as recovery procedures, takes some time and significant resources. Research of methods and means for protection of web services is the topic of the present dissertation, whereas the emphasis is placed on the design of a methodology for achieving security of web services.

The dissertation examines and analyzes proven methods for the protection of web services, as well as seeks new techniques based on machine learning algorithms and fuzzy logic. Experimental research has been conducted, and large data sets collected after the operation of an unattacked and attacked web service have been used to create predictive and analytical models. The prevention and protection of web services is a topic that will continue to attract the attention of the scientific community and industry, as well as the role of artificial intelligence in the field of security.

Based on research, theories and practical statements in the field of web service security, the aim of this PhD thesis is:

Design of methodology, providing secure web services and its verification in a specific context.

For the achievement of the main goal, the following research tasks are defined:

1. Review of scientific publications and technical reports in the field of web service security and analysis of the current situation, summary of existing methods and tools for achieving a certain level of web service security, as well as identification of unresolved issues.
2. Design of methodology for achieving security of web services based on summarizing and analysis of the current situation, as well as, definition of criteria that affect web security.
3. Development of an experimental environment for simulation of attacks on web services - identification of attacks, impacts and security issues.
4. Development of methodology for realizing security of a web service by applying algorithms and techniques from machine learning and fuzzy logic.

The dissertation presents original research on theories and existing practices and the results obtained are in the form of new and adapted existing knowledge, models and solutions in the field of web service security. These results are systematically applied in the development and verification of the proposed methodology for the implementation of web service security.

The dissertation is organized in an introduction, three chapters and a conclusion with summarized contributions received from the presented research.

The first chapter presents a creative overview - summary and analysis of the current state of web service security. The architecture of web services is presented, summarizing their features and requirements from a security point of view. Classifications of threats and attacks against web services are shown, and with both web service protection solutions proposed by researchers and security experts as well as difficult-to-solve problems are discussed. Scientific publications, which offer and discuss methods for investigating web attacks by applying machine learning algorithms are also considered. Mechanisms and algorithms for achieving security of web services at XML level are specifically analyzed. Possible XML-level attacks are summarized and systematized

according to the XML Encryption and XML Signature specifications. The chapter concludes with a developed meta-model of modern tools and methods for the protection of web services.

Chapter two focuses on developing a *methodology* for achieving security of web services. The meaning and content of the concept of methodology is clarified and a model of a conceptual framework for creating methodologies is proposed. A study and comparative analysis of existing methodologies for design and development of web services and secure web services from the point of view of the software engineering process has been performed. A methodology for achieving security of web services has been developed, taking into account important factors such as: availability of security activities for software engineering process, correct administration of the web, as well as, the existence and implementation of adopted security policies and strategies in organizations.

In the third chapter the verification of the proposed methodology for implementing security of web services is performed. For this purpose, a REST XML-based web service has been developed, which is attacked with a commonly used and causing significant damage script attack in a web application (Cross-Site Scripting - XSS). A *methodology* has been developed to identify a stored XSS attack and to provide appropriate security measures to a web service based on the use of machine learning algorithms and fuzzy logic techniques. An experimental study was performed in a specially organized simulation environment, collecting data on the behavior of a web service when it is attacked with a stored XSS attack and when it is not attacked. The own datasets thus generated are converted into a format for further processing. A new set of property selection parameters has been introduced, including characteristics of the environment in which the XSS attack takes place. A predictive model based on machine learning classification algorithms for identifying well-known XSS attacks and a predictive model based on a cluster algorithm for identifying new XSS attacks have been developed. Rules from fuzzy set theory and fuzzy logic are applied to solve multicriteria task in determining security measures to an attacked and non-attacked web service.