***Можливості пакету XML***

***Модулі обробки XML:***

*Інтерфейси* ***Python*** *для обробки* ***XML*** *зібрані в пакеті* ***xml****. Проте модулі* ***XML*** *не захищені від помилкових або зловмисно побудованих даних. Модулі в пакеті* ***xml*** *потребуюьб наявності, хоча б , одного доступного SAX-сувмісного парсера* ***XML****. Парсер Expat входить в* ***Python****, тому модуль* ***xml.parsers.expa****t завжди буде доступнийю*

*Документація для пакетів* ***xml.dom*** *і* ***xml.sax*** *являється визначенням прив’язок* ***Python*** *для інтерфейсів* ***DOM*** *і* ***SAX.***

*Підмодулі обробки* ***XML:***

***- xml.etree.ElementTree: API ElementTree,*** *простий і легкий* ***XML****-процесор;*

*-* ***xml.dom****: визначення* ***API DOM;***

***- xml.dom,minidom:*** *мінімальна реалізація* ***DOM;***

*-* ***xml.dom.pulldom:*** *підтримка для створення часткових дерев;*

*-* ***xml.sax :*** *базові класи* ***SAX2*** *і функції зручності;*

*-* ***xml.parsers.expat :*** *прив’язка парсера* ***Expat;***

***Вразливості XML:***

*Модулі обробки* ***XML*** *не захищені від зловмисних даних. Зловмисник може зловживати функціями* ***XML*** *для проведення атак типу “відмова в обслуговуванні”, доступу до локальних файлів, створення мережевих підключень до інших машин або обходу брендмауерів.*

* ***xml.etree.ElementTree*** *не розширює зовнішні об’єкти і викликає* ***ParserError*** *при виникненні сутності.*
* ***Xml.dom.minidom*** *не розширює зовнішні об’єкти і просто повертає дослывно нерозкритий об’єкт;*
* ***xmlrpclib*** *не розширює зовнішні об’єкти і не пропускає їх;*

***Billion laughs:***

Атака Billion laughs – також відома, як експоненціальне вирішення сутності – використовує декілька рівнів вкладених об’єктів. Кожен об’єкт посилається на інший декілька разів, а кінцеве визначення об’єкта містить невелику стрічку. Експоненціальне розширення призводить до декількох гігібайтів тексту і використовує багато пам’яті і процесорного часу.

***Розширення квадратичної роздуваючої сутності:***

Квадратична атака вибухом схожа на атаку “Billion laughs” . Вона також зловживає розширенням сутності. Замість вкладених об’єктів вона повторює однин великий об’єкт з декількома тисячами символів знову і знову.. Атака не стільки ефективна, як показовий випадок, але вона дозволяє уникнути спрацювання контрмеру парсерів, які забороняють глибоко вкладенних об’єктів.

***Пошук DTD:***

Деякі XML-бібліотеки, такі як xml.dom.pulldom Python, отримують визначення типів документів з віддаленого або локального розташування. Ця функція має схожі наслідки з проблемою розширення зовнішніх об’єктів.

***Декомпресивна бомба:***

Декомпресійні бомби (наприклад, ZIP bimb) застосовуються до всіх XML-бібліотек, які можуть аналізувати стиснення XML-потоків, такі як gzipped HTTP-потоки або LZMA-стиснуті файли. Для зловмисника він може зменшити кількість передаючих данних на три величини або більше.

Документація для defusedmxl на PyPL містить додаткову інформацію про всіх відомих векторах атак з прикладами і посиланнями.

**Пакети defusedxml і defusedexpat:**

***defusedxml –***це чистий пакет Python з модифікованими підкласами всіх пасерів XML stdlib, які запобігають будь-яку потенціально небезпечну операцію. Використання цього пакету рекомендується для будь-якого коду сервера, який аналізує недовіренні дані XML. Пакет також поставляється з прикладами експлойтів і розширеною документацією по великому рахунку експлойтів XML, таких як ін’єкція Xpath.

***Defusexpat –*** представляє модифікований libexpat і виправлений модуль pyexpat, який має контрміри проти DoS-атак із розширенням об’єктів. Модуль defusexpat досі допускає розумну і налаштовуючу кількість розширень об’єктів. Модифікації можуть бути включені в деяку майбутню версію Python, але не будуть включенi в які-небудь випуски виправлення помилок Python, оскільки вони порушують зворотнє співвідношення.

**Модуль xml.etree.ElementTree:**

Модуль xml.etree.ElementTree реалізує простий і ефективний API для робору і створення XML-даних.

XML є за своєю природою ієрархічним форматом даних, і найбільш природнім способом його представлення є дерево. У ЕТ є два класи для цього ланцюга: ElementTree представляє весь XML-документ як дерево, а Element являє собою єдиний вузол в цьому дереві. Взаємодія зі всім елементом(зчитування і запис в/з файлів) зазвичай виконується на рівні ElementTree. Взаємодія з одним елементом XML і його піделементами створюється на рівні Element.

**Модуль xml.dom:**

Об’єктна модель документа, або “DOM”, представляє собою API-інтерфейс з декількома мовами із консорціуму World Wide Web(W3C) для доступу і зміни документів XML. Реалізація DOM являє XML-документ у вигляді дерева і дозволяє клієнтському коду будувати таку структуру з нуля. Також він надає доступ до структури через набір об’єктів, які надають добре відомі інтерфейси.

**Модуль xml.dom.minidom:**

Мінімальна реалізація інтерфейсу Document Object Model з API, аналогічним інтерфейсу на інших мовах. Він повинен бути простішим, ніж повний DOM, а також значно менший.

**Модуль xml.dom.pulldom:**

Модуль xml.dom.pulldom представляє “pull parser”, який також може бути запрошений, щоби при необхідності створювати фрагменти документу, доступні в DOM. Основна концепція включає витягування “подій” із потоку вхідного XML і їх обробку. На відміну від SAX, який також використовує модель обробки, керуючу подіями, разом зі зворотніми викликами, користувач парсер-тяга відпоідає за явне витягування подій із потоку, зациклювання цих подій до завершення обробки або виникнення помилки.

**Модуль xml.sax:**

Цей пакет представляє декілька модулів, реалізуючих інтерфейс Simple API для XML(SAX) для Python. Сам пакет представляє виключення SAX і зручні функції, які будуть найбільш широко використовуватись користувачами SAX API.

**Модуль xml.parsers.expat:**

Являється інтерфейсом Python для нечуттєвого XML-аналізатора Expatю Модуль представляє один тип розширення xmlparser, який представляє даний стан аналізатора XML. Після створення об’єкту xmlparser різні атрибути можуть бути встановлені в функції обробника. Коли є XML-документ, тоді подається в ситнтаксичний аналізатор, функції-обробники викликають для символьни даних і розмітки документі XML,

Цей модуль використовує модуль xml.pyexpat для забезпечення доступу до парсера Expat.