**Algorithm 4** Classify Genre of Document

**Input** Document, Source\_url

Initialize *newsSource* ← […]

Initialize *techSource* ← […]

/\* Calculate the originality and non-article ratio of document \*/

*refernce\_ratio* ← originality\_check(document)

*ioc\_ratio* ← non\_article\_fmt\_check(document)

/\* Judge the genre cascading \*/

**if** *Source\_url* ∈ *newsSource* **or** *refernce\_ratio* >= 30% **then**

**determine type** news

**if** *Source\_url* ∈ *techSource* **or** *ioc\_ratio* >= 30% **then**

**determine type** technical\_report

**if** *multiple used\_malware* in *document* **then**

**determine type** campaign\_report

**else**

**determine type** analysis\_report

**end if**