

第 1 章

はじめに

正規表現は文字列の検索、置換などテキスト処理を行うための強力な手段である。そのような利便性を持つ正規表現であるが、近年、正規表現の脆弱性を利用した ReDoS(regular expression denial-of-service) 攻撃が度々発生している。ReDoS 攻撃とは、文字列のパターンマッチングに膨大な時間がかかってしまうような正規表現を悪用し、計算資源を占有することで、ウェブサイトやサーバーの可用性を侵害する攻撃のことである。2016 年には Stack Overflow のサーバーが、空白を 2 万個含む投稿によりダウンした [1]。2019 年には Cloudflare が提供する CDN が、ファイヤーウォールに追加した脆弱性な正規表現によりダウンした [2]。

ReDoS 脆弱性に対処するためにはどうしたら良いだろうか。もちろんユーザーが独自に作成した正規表現を使わない、正規表現ライブラリを最新にするなどして予防することはできるだろう。しかし、そのようなことをせずとも正規表現を自動で非脆弱なものに修正できないだろうか。本研究ではこの問いに答えるべく、正規表現の自動修正方法を考案した。

2 章では既存研究、本研究で提案する手法を理解するために必要な知識について説明を行う。具体的には正規表現、NFA、正規表現の等価性判定についての説明を行う。

3 章では ReDoS 脆弱性についての既存研究の紹介を行う。具体的には ReDoS 脆弱性についての定式化、ReDoS 脆弱な正規表現を修正するプログラム REMEDY の紹介を行う。

4 章では ReDoS 脆弱な正規表現の新たな修正方法について提案する。REMEDY と正規表現の等価性判定器 RegEq を使用し、脆弱な正規表現をそれと意味的に等価でかつ非脆弱な正規表現へと修正することを試みる。

5 章では 4 章で提案した手法を用いて、実際に脆弱な正規表現が非脆弱化できるかどうか

か実験する.

6 章では本研究のまとめと今後の課題について述べる.

本論文における貢献

- 脆弱な正規表現の修正

ReDoS 脆弱な正規表現を修正するプログラム REMEDY と正規表現の等価性判定器 RegEq を用いて, 脆弱な正規表現を意味的な等しさを保ったまま非脆弱化することに成功した.

参考文献

- [1] Stack Exchange Network Status — Outage Postmortem - July 20, 2016.
[https://stackstatus.tumblr.com/post/147710624694/
outage-postmortem-july-20-2016](https://stackstatus.tumblr.com/post/147710624694/outage-postmortem-july-20-2016). Accessed: 2023-01-21.
- [2] Details of the cloudflare outage on july 2, 2019. [https://blog.cloudflare.
com/details-of-the-cloudflare-outage-on-july-2-2019/](https://blog.cloudflare.com/details-of-the-cloudflare-outage-on-july-2-2019/). Accessed: 2023-01-21.