（報道発表資料）

2024.10.28
日本電信電話株式会社
学校法人早稲田大学

インジェクション攻撃による被害を防ぐためのソフトウェア修正技術を世界にさきがけて実現

専門知識を持たない開発者でもソフトウェア開発段階で文字列操作の誤りを容易に修正

発表のポイント：

- ◆ ソフトウェアの脆弱性を悪用した攻撃の中でも重大な脅威とされているインジェクション攻撃の主要な原因とされている文字列操作の誤りを修正する技術を開発しました。
- ◆ この技術により、専門知識を持たないソフトウェア開発者でも正規表現に起因する文字列操作の誤りの修正が開発段階で可能となりました。
- ◆ サービスの運用段階におけるソフトウェアの誤りの修正には多大なコストがかかりますが、この成果により、開発段階で誤りが修正できるため、コストの軽減と安全なサービスの実現が期待されます。

日本電信電話株式会社（本社：東京都千代田区、代表取締役社長：島田　明、以下「NTT」）と学校法人早稲田大学（本部：東京都新宿区　理事長：田中愛治　以下、「早稲田大学」）は、情報漏洩やサービス停止の原因となり得るインジェクション攻撃による被害を防ぐための、ソフトウェアを構成するプログラム中の文字列関数を用いた文字列操作の誤り（バグ）を修正する技術を世界で初めて実現しました。

ソフトウェアの脆弱性を悪用した最も大きな脅威の１つであるインジェクション攻撃は、プログラム中の文字列操作の誤りが主要な原因であることが知られています。本技術により、専門知識を持たないソフトウェア開発者でも容易に文字列操作の誤りを開発段階で修正することが可能となります。サービスの運用段階におけるソフトウェアの誤りの修正には多大なコストがかかりますが、開発段階で誤りが修正できるため、コストの軽減と安全なサービスの実現が期待できます。

本技術の詳細は、米国カリフォルニア州サクラメントにて開催されるソフトウェア工学分野の最難関国際会議 IEEE/ACM ASE 2024（※1）にて 2024 年 10 月 30 日（米国時間）に発表します。

1. 背景

ソフトウェアの脆弱性を悪用した攻撃は現代社会における重大な脅威です。最も大きな脅威の１つにインジェクション攻撃があります。インジェクション攻撃は、サーバなどへの攻撃手法のひとつで、サーバで利用しているデータベースなどに対して不正な入力情報を送信して、予期しない動作を引き起こします。この攻撃をもたらす欠陥はインジェクション脆弱性と呼ばれ、プログラム中の文字列操作の誤り（バグ）が主な原因であることが知られています。

ソフトウェアを構成するプログラム中で文字列操作を記述する際には文字列関数が利用されます。

文字列関数は多くのプログラミング言語で提供されており、文字列の抽出・置換検索などの文字列操作を記述するために頻繁に利用されています。

しかし、文字列関数を利用して文字列操作を行う際には文字列関数が利用する正規表現（※2）やその他の入力情報、文字列関数自体の仕様に関する専門的な知識が要求され、適切に記述することは難しいことが知られています。

文字列操作を伴うプログラムはさまざまなソフトウェアで幅広く利用され、Web サイトのフォームにユーザが Web ブラウザ経由で入力した情報を Web サイト側で加工処理するなど多くのサービスを実現しています。文字列操作に誤りがあるとサービスの誤動作を引き起こし、情報漏洩やサービス停止の原因となる場合があり、サービスの運用段階におけるソフトウェアの誤りの修正には多大なコストがかかります。また、この誤動作を意図的に起こそうとするサイバー攻撃も顕在化しており、安全なサービスの実現を脅かすリスク要因となっています。

2. 研究の成果

これまでNTTと早稲田大学は共同で、プログラム中で文字列を扱う操作により発生し得る脆弱性や誤りを自動修正する技術の研究を行ってきましたが、その対象は正規表現に限定されていました（※3、※4）。

本成果では、プログラム中の文字列関数を使った文字列操作の誤りを、ソフトウェア開発者が与える入出力例を基に修正する技術を世界で初めて実現し、正規表現を含む文字列操作にまで修正対象を広げることを可能としました。

役割

NTT：問題の形式化と修正手法の考案。

早稲田大学理工学術院　寺内多智弘教授：NTT が考案した手法の理論的な正確さの検証。

3. 本技術のポイント

本技術は、インジェクション攻撃の主要な原因であるプログラム中の文字列操作の誤りをソフトウェア開発者が与える入出力例を基に修正し、修正結果に誤りがないことを保証する技術です。

本技術のポイントは、以下の通りです（図1）。

① 文字列関数に期待する入出力例を表記する方法を考案

従来的な入出力のみを提示する例を用いた場合と比べ、ソフトウェア開発者が文字列関数に期待する入出力例を入力と出力だけでなく入力のどの部分をどのように変換したいのかも含めて表記できるようになり、適切な修正結果の出力に寄与します。この表記は、プログラムが満たすべき入出力の例を基にプログラムを生成する「Programming by Examples （PBE）」メソッド（※5）を用いて与えます。

② 文字列関数の振る舞いを理論モデルとして厳密に定義

この定義を用いることで、修正対象の文字列関数に与える情報であるパラメータがすべての入出力例を満たすための条件を導き出すことが可能となります。本技術では、Web アプリケーションなどで広く利用されている ECMAScript 2023 （※6） に準拠した文字列関数の振る舞いを厳密に定義しました。これにより、修正結果がすべての入出力例を満たすことを保証できるようになります。

③ 理論モデルに従い、例を全て満たす形に文字列関数のパラメータを修正する技術を考案

修正結果のパラメータとなり得る候補を、明らかに入出力例を満たさないものを除外しつつ網羅的に探索する手法を考案しました。この手法では、現実的な時間内に修正処理を終えることが可能となります。また修正結果が修正前のパラメータに対して最小限の変更による修正が可能となり、ソフトウェア開発者が目視で容易に修正結果を確認できるようになります。

本技術により、ソフトウェアの脆弱性を悪用した攻撃の中でも重大な脅威とされているインジェクション攻撃の主要な原因とされている文字列操作の誤りを開発段階で修正可能となるため、修正コストの軽減とソフトウェア開発の品質向上に貢献できるものと期待されます。

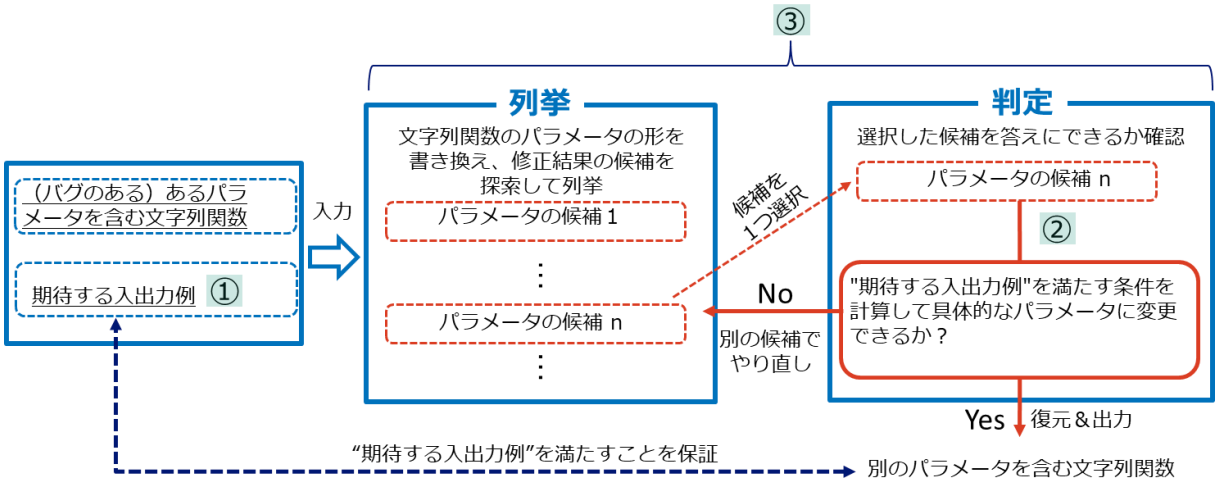

図1　文字列関数に対する処理の例

4. 今後の展開

サービスの運用段階におけるソフトウェアの誤りの修正には多大なコストがかかりますが、本技術が利用されていくことで、ソフトウェアの開発段階で誤りが修正できるため、コストの軽減と安全なサービスの実現が期待できます。また AI を用いたプログラムの自動生成において、非熟練者が AI を用いて作成したプログラムに含まれる誤りにどう対処するのかという新しい問題も生まれています。文字列操作の誤りを修正する本技術は、AI による自動化のメリットを損なうことなくプログラムの安全性向上に寄与できるものと期待されます。今後は、文字列操作に伴う脆弱性そのものを修正する技術の研究を進める予定です。

【用語解説】

※1. ASE・・・Automated Software Engineering は IEEE/ACM によって運営されるソフトウェア工学分野の最難関国際会議。本技術は、2024 年 10 月 27 日～11 月 1 日に開催される IEEE/ACM ASE 2024(39th IEEE/ACM International Conference on Automated Software Engineering)にて、下記のタイトル及び著者で発表されます。

タイトル：Repairing Regex-Dependent String Functions

著者：Nariyoshi Chida (NTT Social Informatics Laboratories), Tachio Terauchi (Waseda University)

URL: https://doi.org/10.1145/3691620.3695005

※2. 正規表現・・・コンピュータで特定の文字の並び（文字列）をルールに基づき簡略化して表現する方法の 1 つで、特定の文字列のパターンを検索・抽出・置換するときに用いられる。

※3. プログラム中の文字列チェック機能の脆弱性を自動修正する技術を世界に先駆けて実現～専門知識をもたない開発者でも ReDoS 脆弱性の修正が容易に～[https://group.ntt/jp/newsrelease/2022/03/23/220323b.html]

※4. プログラム中の文字列抽出機能を自動修正する技術を世界に先駆けて実現～専門知識をもたない開発者でも正規表現の修正が容易に～[https://group.ntt/jp/newsrelease/2023/06/16/230616b.html]

※5. ECMAScript 2023・・・ECMAScript は、Web アプリケーションなどで広く利用されているプログラミング言語 JavaScript の標準規格。本技術は 2023 年に改定された ECMAScript を対象に検証を実施。

※6 Programming by Examples （PBE）メソッド・・・プログラミングの知識を持たないエンドユーザでもプログラムを生成できるようにする手法の 1 つで、プログラムが満たすべき入出力の例を与えると、それを実現するプログラムを自動で生成する技術。

■本件に関する報道機関からのお問い合わせ先
日本電信電話株式会社
サービスイノベーション総合研究所
企画部 広報担当
nttrd-pr@ml.ntt.com

学校法人早稲田大学
広報室広報課
koho@list.waseda.jp