論 文 内 容 の 要 旨

博士論文題目　Characterizing the Quality of Third-Party Libraries through Runnability and Risk Assessment in the Open Source Ecosystem

氏　　名　BODIN CHINTHANET

Third-party libraries become an important part of software development within the open source community. These library packages provide developers with useful features without the need to "reinvent the wheel," with each package often depending on several others. It is challenging for developers to choose high-quality packages and to quickly respond to their security vulnerability fixes. The goal of this thesis is to help developers (i) choosing high-quality packages and (ii) understanding lags from the adoption and propagation of vulnerability fixes. In particular, the first part of this thesis investigates features that make a good Node.js package by investigating users, contributors, and runnability of packages. The results from the survey and the empirical study show that both users and contributors share similar views on which features they use to assess package quality, especially by using the runnability of packages. Also, the runnability can be predicted by using the package features. The second part of this thesis investigates the vulnerability fixes and their lags from both package-side and client-side. The results from the empirical studies show that most of the commits in the fixing release are not related to the fix and lags in the adoption and propagation of those fixes depend on factors such as the severity of the vulnerability. To help developers to prioritize the development task and reduce the lags, I propose two prototype tools for detecting the vulnerable codes in the software project and identifying the reachability of those vulnerable codes. Overall, this thesis shows practical implications to help developers choose a good package and quickly mitigate the risk of vulnerability from their dependency packages.

氏　名	BODIN CHINTHANET

（論文審査結果の要旨）

本論文は，オープンソースソフトウェアのエコシステムにおいて，ソフトウェアパッケージにおける品質評価と脆弱性伝搬に着目し，ソフトウェア開発者支援の高度化を図るものである．ソフトウェアパッケージは，汎用性の高いプログラム部品の集合体であり，その用途や処理対象を同じくするものがライブラリとして更にまとめられ，エコシステムにおいて広く配布，共有，活用されている．特に，第三者が作成したライブラリはサードパーティライブラリと呼ばれる．エコシステムに多数存在するサードパーティライブラリやそこに含まれるパッケージは品質のばらつきが大きく，開発者が高品質なパッケージを選択することは容易ではないとされている．また，パッケージの脆弱性に関しては，その配布元で脆弱性が発見され，コード修正などの対策が施されても，パッケージを共有し活用している開発者（配布先）にその対策がすぐに伝搬し脆弱性が解消されるとは限らない．脆弱性の発見から解消までの遅延にはばらつきがあり，かつ，伝搬や遅延の状況を開発者が把握することは容易でないとされている．

ソフトウェアパッケージの品質評価については，広く普及している JavaScript 環境 Nodo.js を対象として，サードパーティライブラリの作成者（エコシステムにおける貢献者）とその利用者（エコシステムにおける一般の開発者）の双方を対象とした調査を行うとともに，同ライブラリを構成するパッケージの特性値の定量化を実施した．その結果，パッケージ品質評価の観点は，パッケージの作成者と利用者の間で差異はなく，特に，パッケージの実行可能性を重視していることが分かった．更に，パッケージの特性値を用いてその実行可能性を予測することで，開発者が高品質なパッケージを選択できるよう支援することが可能であることを示した．

ソフトウェアパッケージの脆弱性伝搬については，開発プラットフォーム GitHub 上で最大級と言われている npm エコシステムを対象として，サードパーティライブラリについて 2009 年 4 月から 2020 年 8 月までに提出された脆弱性報告の分析を実施した．その結果，脆弱性伝搬における遅延は，脆弱性の深刻度，および，脆弱性発見からの経過時間，に大きく依存していることが分かった．更に，パッケージに内在する脆弱性コードを検出するとともに，パッケージ利用時に脆弱性コードが実行されるかどうか（到達可能性）を評価する２つのプロトタイプツールをパッケージ利用者向けに提案した．同ツールを用いることで，パッケージ利用者は脆弱性修正に係る作業に優先順位をつけ，脆弱性伝播における遅延を低減できるようになる．

以上のとおり，本論文は，ソフトウェアエコシステムで広く用いられているサードパーティライブラリとそこに含まれるパッケージを，品質評価と脆弱性伝搬の観点で分析することを通じて，ソフトウェア開発者が高品質なパッケージを容易に選択することを可能にし，脆弱性伝搬の遅延低減にもつながる明確で具体的な知見を明らかにしている．その手法や得られた知見は，広くソフトウェア開発者支援の高度化，そして，ソフトウェア工学研究の発展に大きく貢献することから，博士（工学）論文として価値あるものと認める．