

平成 26 年 5 月 29 日現在

機関番号：17401

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：24700029

研究課題名(和文) オープンソースソフトウェア解析に基づくソフトウェアライセンスプロトタイプ作成環境

研究課題名(英文) Building an environment for generating software license prototype based on analysis of open source software

研究代表者

眞鍋 雄貴(眞鍋雄貴)(Manabe, Yuki)

熊本大学・自然科学研究科・助教

研究者番号：20625339

交付決定額(研究期間全体)：(直接経費) 1,400,000円、(間接経費) 420,000円

研究成果の概要(和文)：本研究では、ソフトウェア開発者によるソフトウェアライセンス(ライセンス)の試作支援環境構築を目的とし、ライセンスの観点からオープンソースソフトウェア(OSS)の解析を行った。研究成果として、実際のOSS群において、ソフトウェアのライセンスとソースファイルのライセンスを比較し、ライセンス間の包含関係の抽出を行った。また、比較という観点でコードクローンの理解やソフトウェアの動作予測への発展を行った。

研究成果の概要(英文)：In this study, I analyze Open Source Software in terms of software license to build an environment for generating software license prototype. As a result, I showed an inclusion relations among open source licenses by comparing the license of software to one of source file included in the software. In addition, I expanded this study to code clone comprehension and prediction of future behavior of software.

研究分野：総合領域

科研費の分科・細目：情報学・ソフトウェア

キーワード：ソフトウェア工学 ソフトウェアライセンス 再利用 オープンソースソフトウェア 国際情報交換

1. 研究開始当初の背景

ソフトウェア開発コストを削減するため、既存のオープンソースソフトウェア (OSS) の一部を別のソフトウェアの開発に利用することがある。ソースファイルを利用するためには、そのソースファイルのソフトウェアライセンス(ライセンス)に従う必要がある。ライセンスとは著作物の利用に対する許諾と、その許諾を得るために従わなければならない義務である。ライセンスは自然言語で記述された条件である条項の集合として表現される。

既存研究として、ソースファイルのライセンスを特定する労力を軽減するため、ライセンス特定ツールが開発されている。研究代表者は自動的にライセンスを特定するために、大規模オープンソースソフトウェアの集合を解析し、96%で正しく特定できるツールである Ninka を構築した[1]。また、研究代表者はこのツールを用いて、ライセンスと再利用の関係[2]やソフトウェアにおけるライセンス変化[3]といったライセンスに関する実証的な研究を行ってきた。また、Germanら[4]や、Alsplaughら[5]はライセンスのモデルを提案している。

一方で、ソフトウェア開発者がライセンスを決定する手段はあまり提案されていない。ライセンスの内容を決定する手法を提供しているライセンスとして Creative Commons(<http://creativecommons.jp/>)がある。Creative Commons では作品のクレジットの表示、改変禁止、営利目的の利用の禁止、元の作品と同じ組み合わせの CC ライセンスで公開することの4つの条項を組み合わせることで開発者に応じたライセンスを提供している。しかし、実際には特許などに対する要求などには対応できておらず、単純化しすぎている可能性がある。

しかし、法学知識のないソフトウェア開発者がライセンスを決定することは難しい。その理由として(1)ライセンスは法律文書であり、自然言語の長文を作成することになるから、(2)要求に対してどのような条項を記載しないとイケないかわからない、(3)どのような単語が条項の中で用いられているかわからないという理由からである。

<1章参考文献>

- [1] 真鍋雄貴、ダニエル モラレス ゲルマン、井上克郎、階層的ライセンス知識を用いたライセンス特定ツールの開発、情報処理学会論文誌、Vol.52、No.8、

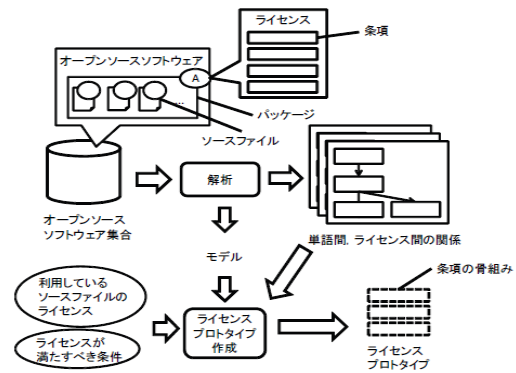


図 1 研究方法の概要

pp.2402-2411、2011

- [2] Yu Kashima、 Yasuhiro Hayase、 Norihiro Yoshida、 Yuki Manabe and Katsuro Inoue、 A Preliminary Study on Impact of Software Licenses on Copy-and-Paste Reuse、 In Proc. IWESSEP2010、 pp. 47-52、 2010
- [3] Yuki Manabe、 Yasuhiro Hayase and Katsuro Inoue、 Evolutional Analysis of Licenses in FOSS、 In Proc. EVOL and IWPSE、 pp. 83-87、 2010
- [4] Daniel M. German and Ahmed E. Hassan、 License Integration Patterns: Dealing with Licenses Mismatches in Component-Based Development、 In Proc. ICSE 2009、 pp. 188-198、 2009
- [5] Thomas A. Alsplaugh、 Hazeline U. Asuncion and Walt Scacchi、 Analyzing software licenses in open architecture software systems、 In Proc. FLOSS 2009、 pp.54-57、 2009

2. 研究の目的

本研究では、開発者が容易に自分の要求を満たすソフトウェアライセンスを作成する労力軽減を目的とし、ライセンスの素材となるソフトウェアライセンスプロトタイプ(ライセンスプロトタイプ)を作成する環境を作る。ライセンスプロトタイプとは、本研究で決定するライセンスのモデルの形で表現された条項の集合である。この研究を通じて、条項に含まれる単語がどのような関係にあるか、その単語間の関係をどのようなモデルで表現できるか、モデルを用いることでライセンスプロトタイプを作成できることを明らかにする。

3. 研究の方法

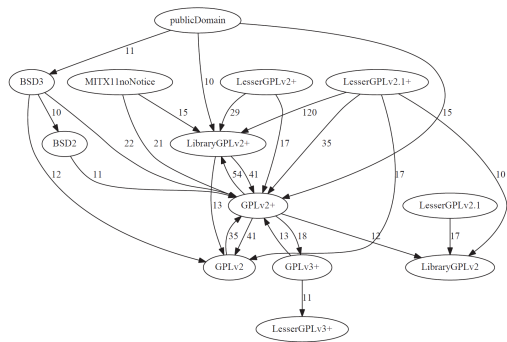


図2 Fedora Core 17のパッケージから得られたライセンス間の包含関係

ソフトウェアライセンス試作支援環境の概要を図1に示す。ソフトウェアライセンス支援環境は、大きく分けてソフトウェアライセンス間の関係を解析し、ライセンス間の関係を解析することとライセンスプロトタイプ作成手法の開発によりライセンスプロトタイプ作成環境の構築を行う。

ソフトウェアライセンス間の関係解析では、実際にオープンソースソフトウェアで用いられている用法を調べ、オープンソースソフトウェア全体のライセンスはその中に含まれるソースファイルのライセンスより同程度が厳しいものでなくてはならないという関係を用いることで、ライセンスを直接比較することなく、ライセンス間の関係を得る。次に、ライセンス間の関係とでライセンス条文間での差分から、条項に含まれる単語間の関係を得る。また、既存のソフトウェアライセンスを解析することで、モデルの構築を行う。

ライセンスプロトタイプ作成手法の開発では、作成したライセンス間の関係と現在開発中のソフトウェアに含まれるソースファイルのライセンスの対応関係を調べ、ソースファイルのライセンスに違反しない条項の骨組みを作成する。骨組みとは作成したモデルで表現された条項のことである。次に、その他の条件により、その集合をフィルタリングし、ユーザーが求める条項の骨組みの集合を得ることで、ライセンスプロトタイプを作成する。

4. 研究成果

2年間の研究期間において、研究計画のうちライセンス間の関係解析までを行った。その結果、オープンソースライセンス間の関係を事例に基づいて解析した成果を主に、比較という観点でコードクローンやソフトウェアの動作予測への展開を行うという成果が得られた。以下に、各年度の研究成果について概要を述べる。

平成24年度の成果として、(1)オープンソースソフトウェアにおける事例に基づき、ライセンス間の関係の解析を行った。また、

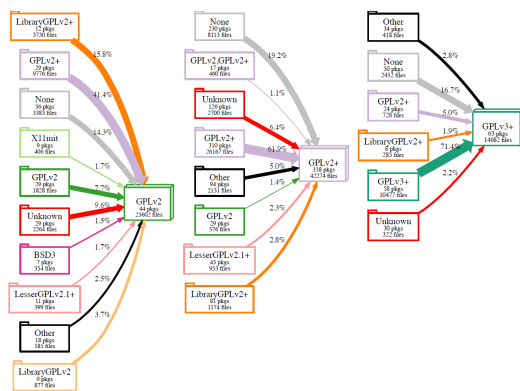


図3 主要なオープンソースソフトウェアライセンスに含まれるソースファイルのライセンスの調査結果(一部)

ライセンス間の関係解析における枠組みを応用し、(2)過去の動作と現在の動作を比較することによって、今後の振る舞いを予測するという手法を考案した。以下、各成果について説明する。

(1)では、あるライセンスの成果物を別のライセンスで配布できるとき、それらの中にある関係を包含関係と定義し、既存のオープンソースソフトウェア内に現れる包含関係を抽出し、まとめた。本調査では、Linuxディストリビューションの一つであるFedora Core 17で管理されるソースパッケージ集合を既存のオープンソースソフトウェアの集合とみなし、多数のパッケージに登場する包含関係を抽出した。得られた結果を図2に示す。この成果によって、ライセンスの条項間に出現する単語間の関係を解析する際の基盤となる知見を得ることができ、以降のライセンスのモデル化やライセンスプロトタイプの構築に貢献できると考えている。また、この知見は、法学知識のないソフトウェア開発者がソースファイルやライブラリを選択する際、ライセンスを考慮するコストを削減できる可能性を持つ。

(2)では、オブジェクト指向プログラムを対象として、記録した実行履歴と現在の実行とを比較することでオブジェクトの振る舞いを予測する手法を提案し、実際のソフトウェアで評価した。その結果、提案手法によって、多くの場合において、呼び出し文の系列が予測可能であることを示した。本研究課題では、事例に基づきライセンス間の関係を解析することが目的であるが、本成果はこのアプローチがアプリケーションの振る舞いに応用することができることを示しており、本研究課題で得られた知見が他の分野に適用できる可能性を示している。

平成25年度の成果は(1)ライセンス間の包含関係のうち、主要なライセンスが関係するものについてより詳細に調査した、(2)重複しているソースコード片であるコードクローンの周囲にある要素の比較方法とそれをういた調査であった。以下、各成果につ

いて説明する。

(1)では、平成24年度に行った包含関係の解析を、最新のFedora Core19のソースパッケージを使用し、また、海外の研究者と共同で行うことにより洗練させた。本年度は、各主要なOSSライセンスで配布されているパッケージにはどのようなライセンスのソースファイルが含まれるかを調査した。調査結果の一部を図3に示す。その結果、様々なライセンスが含まれる場合がある一方、ほぼ単一のライセンスのソースファイルのみ持つような場合があった。

(2)では、他のコードと重複するコード片であるコードクローンに対する周辺コード(コードクローン中に存在する変数のデータフロー上に出現するフィールド変数、引数を持たないメソッド呼び出し、コードクローンを含むメソッドの仮引数(以上、外部要素)とコードクローンを含む制御ブロックのうち、齋内にあるブロック)を抽出、表示するツールの作成と、これを用いた7つのOSSの調査を行った。その結果、コードクローンであっても外部要素が異なっている場合や異なるコード片であっても外部要素が同一である場合を示した。

これらの研究成果では、ライセンスプロトタイプ構築までは至らなかった。しかし、ライセンス間の関係を実際の用例に基づいて、より詳細に調査することができ、また、メタデータを対象とする点で研究範囲を広げることが出来た。これらの成果は今後のOSS開発をより発展させるために有用であったと考える。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 5 件)

Yuki Manabe, Daniel M. German, and Katsuro Inoue, Analyzing the Relationship between the License of Packages and Their Files in Free and Open Source Software, The 10th IFIP WG2.13 International Conference on Open Source Systems OSS 2014, 2014.5.7, San Jose, Costa Rica
DOI:10.1007/978-3-642-55128-4_6

眞鍋雄貴, オープンソースライセンスへのソフトウェア工学からのアプローチ, SES2013 ワークショップ オープンソースソフトウェア工学 招待講演, 2013.9.11, 東洋大学白山キャンパス, 東京都
<https://sites.google.com/site/sesjp20>

13/workshop#ws-5

バンネメフ オドファー, 眞鍋雄貴, 伊達 浩典, 石尾 隆, 井上 克郎, コードクローンの動作を比較するためのコードクローン周辺コードの解析, 情報処理学会ソフトウェア工学研究会, 2013.5.27, 東京大学弥生キャンパス内向ヶ丘ファカルティハウス, 東京都
<http://ci.nii.ac.jp/naid/110009579620>

眞鍋雄貴, 井上克郎, オープンソースソフトウェアにおけるソフトウェアライセンス間の包含関係の調査, 電子情報通信学会ソフトウェアサイエンス研究会, 沖縄県石垣市市民会館, 沖縄県, 2013.1.10,
<http://webcat.nii.ac.jp/naid/120005373378>

脇阪 大輝, 眞鍋雄貴, 石尾 隆, 井上 克郎, DOPGを用いたオブジェクトの振舞い予測手法, ソフトウェア工学の基礎XIX, 第19回ソフトウェア工学の基礎ワークショップ (FOSE2012), 2012.12.14, ゆふいん山水館, 大分県
<http://sel.ist.osaka-u.ac.jp/~lab-db/betuzuri/contents/ja/887.html>

〔図書〕(計 0 件)

〔産業財産権〕
出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕
なし

6. 研究組織

(1)研究代表者

眞鍋雄貴(MANABE, Yuki)
熊本大学・大学院自然科学研究科・助教
研究者番号: 20625339

(2)研究分担者

なし

(3)連携研究者

なし