Open Source License Violation Check for SPDX Files

Georgia M. Kapitsaki\(^1\) and Frederik Kramer\(^2\)

\(^1\) Department of Computer Science, University of Cyprus,
1 University Avenue, Nicosia, Cyprus
gkapi@cs.ucy.ac.cy

\(^2\) Otto von Guericke University, Universitätsplatz 2, D-39106, Magdeburg, Germany
frederik.kramer@ovgu.de

Abstract. The Open Source Software development model has gained a lot of momentum in the latest years providing organizations and software engineers with a variety of software, components and libraries that can be exploited in the construction of larger application systems. Open Source Software is accompanied by licenses that state the conditions under which the intellectual property can be used. Since not all licenses are governed by the same conditions of use, the correct combination of licenses is vital, when different libraries are exploited in newly developed application systems. If this is not adequately handled, license violations might be a consequence of incompatibilities. In this paper we present our work on license violation checking in the framework of Software Package Data Exchange (SPDX). Starting from the modelling of license compatibilities our approach examines potential violations in software package information formatted using the SPDX specification. At the same time alternative solutions in the form of applicable licenses for the software package are proposed. This approach can be a valuable asset for Open Source practitioners in the license decision process assisting in detecting possible violations and in making suggestions on license use.

Keywords: Free Open Source Software, Licensing, License compatibility, Software Package Data Exchange.

1 Introduction

The popularity of Free and Open Source Software (FOSS) among software engineers and enterprises is constantly growing as organizations try to integrate openness in their procedures [5]. FOSS provides new possibilities for engineers to incorporate third party software into their implementations and for organizations to distribute their products as open source embracing open ICT models. In open source software, licenses express how the software is made available and under which conditions it can be used, integrated, modified and redistributed.

© Springer International Publishing Switzerland 2014
The problem that many software vendors often face is how to incorporate third party software in their implementations correctly without causing any license violations; remaining hence legally compliant. License violations are a complex issue in FOSS due to the variety of licenses that state different and often contradicting conditions of use referring mainly to conditions for software modification and re-distribution. When software libraries licensed under different terms are jointly used, this license diversity may lead to license incompatibilities. Licenses cover a range from very permissive licenses, such as the MIT license and the Academic Free License (AFL), to highly restrictive licenses, such as the GNU General Public License (GPL) and the Open Source License (OSL). Licenses are categorised as either permissive or copyleft. Permissive licensing allows the software to be part of a larger product under almost any other license. Copyleft licenses are posing more restrictions. Copyleft is further divided into weak and strong copyleft. If the software used is weak copyleft-licensed, the created work can be distributed under another license as long as no changes are made in this weak copyleft-licensed software used. Strong copyleft requires all derivative work that uses strong copyleft-licensed software to be distributed with the very same license.

The Software Package Data Exchange (SPDX) specification addresses the issue of integrating license information into the distribution of software packages that can be formatted according to SPDX [12]. As such it is vital to examine whether the information contained in the SPDX file regarding the license applied on the software package is correct.

In this work we are addressing license violation detection by using SPDX files. Specifically, we have designed and implemented a compatibility tool that assists in: 1) verifying that the information on the license applied on the software package is correct, 2) identifying any license incompatibilities among the licenses of the software package, and 3) making suggestions for licenses that can be applied without causing violations. A side contribution of our work, that is evolving, lies in the modelling of license compatibilities for a popular license set captured in a license graph. This is the first work that approaches license violations in a specification setting that can offer a global solution for license compatibility enforcement. We hope that it will trigger more research on license compatibility and further promote the use of SPDX, as well as contribute to the tools accompanying the specification.

The rest of the paper is structured as follows. Section 2 briefly presents the Software Package Data Exchange specification focusing on parts that are relevant to our work, whereas section 3 is dedicated to the modelling of license compatibilities presenting the license compatibility graph. Section 4 analyses the SPDX tool. This is further demonstrated in section 5 through a proof-of-concept evaluation on existing open source projects. Section 6 gives a brief overview of related work on open source licensing and compatibilities and, finally, section 7 concludes the paper.