

Een analyse van Open Source gebruik in het Aangifteprogramma van de Belastingdienst.

© 2008 Peter Roozemaal, MathFox Automatisering.¹

Verspreiding, reproductie en opname in een gegevensbestand in ongewijzigde vorm is toegestaan, onder vermelding van auteursnaam en deze licentie. Het is niet toegestaan toegang tot en verspreiding van dit document bij overeenkomst of met technische middelen te beperken.

Samenvatting

Gebruikers van het Linux besturingssysteem stellen het zeer op prijs dat de Belastingdienst het mogelijk maakt om met een Linux systeem aangifte te doen. Dat de Belastingdienst gebruik maakt van Open Source is zowel uit kosten- als kwaliteitsoogpunt begrijpelijk.

Een klein punt van zorg is dat de belastingdienst in haar haast om de Linux gebruikers van dienst te zijn vergeten lijkt te zijn om, zoals de licentie vereist, de volledige licentietekst voor de OpenSSL bibliotheek over te nemen in haar hulppagina's. Deze omissie moet simpel hersteld kunnen worden.

Vervelender is dat een simpele technische analyse laat zien dat de onderzochte Aangifteprogramma's voor Linux statisch gelinkt zijn met de open-sourcebibliotheek GTK+. Dit betekent dat een aanzienlijke deel van de objectcode uit de GTK+ bibliotheek gekopieerd is in de executables van de Aangifteprogramma's. De standaardlicentie waaronder GTK+ verspreid wordt, de GNU LGPL 2.1, staat het royalty-free gebruik van de bibliotheek en verspreiding van de resulterende executables onder bepaalde voorwaarden toe.

Helaas voldoet de Belastingdienst niet aan sectie 6 van de LGPL, daar zij noch de broncode van de GTK+ bibliotheek, noch objectcode die gebruikt kan worden om met een alternatieve versie van de bibliotheek te linken beschikbaar stelt aan degenen die de executable gedownload hebben. Hiermee wordt verspreiding van het Aangifteprogramma een inbreuk op het auteursrecht van de GTK+ ontwikkelaars.

De Belastingdienst dient verspreiding van betreffende Aangifteprogramma's te staken totdat bron- en objectcode volgens sectie 6 van de LGPL beschikbaar gemaakt is aan allen die een kopie van het Aangifteprogramma in hun bezit hebben. Voor degenen die de programma's gedownload hebben kan dit gebeuren door het aanbieden van een download van een Belasting-

1 Peter Roozemaal werkt sinds 2004 voor Open Source Risk Management en analyseert de risico's verbonden aan het gebruik van Open Source software in commerciële applicaties. Analyses als beschreven in dit rapport voert hij regelmatig uit voor klanten van OSRM. Email: proozemaal@osriskmanagement.com

dienst website. Personen die een kopie van het programma op diskette of CD ontvangen hebben moet de mogelijkheid geboden worden om (tegen kostprijs) de vereiste materialen op CD of diskette te ontvangen. De Belastingdienst kan overwegen om de Aangifteprogrammas voor komende jaren dynamisch met de GTK+ bibliotheek te linken, gebruik makend van de standaard Linux faciliteiten hiervoor.

Feitenoverzicht

Dit rapport is gebaseerd op een analyse van het Programma Aangifte inkomstenbelasting 2007 voor ondernemers, Linux versie, gedownload op 5 februari 2008 als wa2007_linux.tar.gz van mijn.belastingdienst.nl. De analyse heeft zich geconcentreerd op het wa2007ux programma.

Inhoud van wa2007_linux.tar.gz

Het wa2007_linux.tar.gz archief (gecomprimeerd tar bestand) bevat de volgende files:

```
wa2007/bin/wa2007d
wa2007/bin/wa2007ux
wa2007/share/belastingdienst.nl/wa2007/wa2007.bdr
wa2007/share/belastingdienst.nl/wa2007/wa2007.efr
wa2007/share/belastingdienst.nl/wa2007/wa2007.gui
wa2007/share/belastingdienst.nl/wa2007/wa2007.mld
wa2007/share/belastingdienst.nl/wa2007/wa2007.pem
wa2007/share/belastingdienst.nl/wa2007/wa2007.pfr
wa2007/share/belastingdienst.nl/wa2007/wa2007.scr
wa2007/share/belastingdienst.nl/wa2007/wa2007.top
wa2007/share/belastingdienst.nl/wa2007/wa2007.zip
wa2007/share/belastingdienst.nl/wa2007/wa2007d.cfg
```

Het wa2007ux Programma

Het wa2007ux programma is een normale Linux ELF binary voor Intel 80386 en compatible processoren. De uitvoer van ldd laat zien dat wa2007ux dynamisch gelinkt² is met een aantal systeembibliotheken:

```
linux-gate.so.1 => (0xffffe000)
```

2 Dynamisch linken betekent dat de bibliotheekcode pas bij de uitvoering van het programma gecombineerd wordt met de programmacode. Dit in tegenstelling tot statisch linken waar de benodigde bibliotheekcode in het programma opgenomen wordt.

```
libdl.so.2 => /lib32/libdl.so.2 (0xf7f32000)
libpthread.so.0 => /lib32/libpthread.so.0
(0xf7f1a000)
libXext.so.6 => /usr/lib32/libXext.so.6 (0xf7f0b000)
libX11.so.6 => /usr/lib32/libX11.so.6 (0xf7e1a000)
libSM.so.6 => /usr/lib32/libSM.so.6 (0xf7e12000)
libm.so.6 => /lib32/libm.so.6 (0xf7ded000)
libc.so.6 => /lib32/libc.so.6 (0xf7ca3000)
/lib/ld-linux.so.2 (0xf7f54000)
libXau.so.6 => /usr/lib32/libXau.so.6 (0xf7ca0000)
libXdmcp.so.6 => /usr/lib32/libXdmcp.so.6
(0xf7c9a000)
libICE.so.6 => /usr/lib32/libICE.so.6 (0xf7c82000)
```

Symbolen in het wa2007ux Programma

De standaard symbooltabel is uit de wa2007ux executable gestript³; de symbooltabel voor dynamisch linken is wel aanwezig zoals `nm -D wa2007ux` laat zien. Wanneer we de C++ functienamen demanglen herkennen we tal van namen uit de GTK+ en wxWidget bibliotheken. De totale lijst van gedefinieerde symbolen, het merendeel C en C++ functies en C++ member functies is 29127 regels lang.

We correleren de functienamen met een lijst van namen uit de standaardbibliotheken die in de gangbare Linux distributies opgenomen zijn; daarvoor hebben we specifiek de Ubuntu bibliotheken voor libgtk1.2 (versie 1.2.10-18) en wxwidgets (libwxgtk2.8, versie 2.8.4.0-0ubuntu3) gedownload en op een testsysteem geïnstalleerd. De belangrijkste gevonden FOSS projecten zijn:

FOSS project	aantal symbolen	bytes code en data
compat-libstdc++-2.96-2.96-139	930	37990
GTK+-1.2	1886	292344
libjpeg-static-6b-39	97	27862
WxWidgets	8707	1235544

Het gehele Aangifteprogramma bevat 29127 symbolen en is 4181490 bytes groot.

³ De term strippen komt van de naam van het Unix/Linux programma strip dat symboolinformatie en de symbooltabel verwijdert.

Erkenning door de Belastingdienst

In de helppagina's bij erkent de belastingdienst het gebruik van Open Source code in het aangifteprogramma. De Belastingdienst noemt in de hulppagina's van het Aangifteprogramma de volgende projecten: gSOAP, OpenSSL, wxWidgets, GTK+, ZLIB, Independent JPEG Group-componenten, Autopackage en NSIS. De licentietekst is beschikbaar voor alle erkende projecten, met uitzondering van de OpenSSL licentie; de hulppagina voor OpenSSL bevat slechts een externe URL van de licentietekst.

Het wa2007d Programma

Een analyse van de gedefinieerde symbolen in het wa2007d programma levert een sterk vermoeden van de aanwezigheid van GTK+ code in de executable. We vinden in dit programma eveneens sterke aanwijzingen voor de aanwezigheid van OpenSSL code

Aangifte Particulieren, ib2007ux

Ook een analyse van het ib2007ux programma uit een recente ib2007_linux.tar.gz download van www.belastingdienst.nl wekt een gerechtvaardigd vermoeden van de aanwezigheid van GTK+ en OpenSSL code in de executable.

Analyse

Deze analyse gaat in detail in op de Linux versie van de Aangifteprogramma's voor ondernemers. De gesignaleerde problemen kunnen ook aanwezig zijn in andere Aangifteprogramma's, maar deze zijn niet in voldoende detail onderzocht om een gemotiveerd oordeel te vormen.

Discrepanties tussen erkenning en gevonden code

De door de Belastingdienst genoemde tool Autopackage is niet gevonden in de onderzochte Linux tarball . De Belastingdienst distribueert ook een alternatieve, met Autopackage gemaakte package en het is waarschijnlijk voor die versie dat Autopackage genoemd wordt. De analyse vond ook geen gSOAP, OpenSSL, ZLIB of NSIS code. Dit kan komen doordat de code niet (meer) aanwezig is in het Aangifteprogramma, of door beperkingen van de code die de matches maakt, zoals het ontbreken van Open Source projecten in de referentiedatabase.

Het package compat-libstdc++-296-2.96 bevat objectcode voor de (ISO) standaard C++ libraries, zoals die meegeleverd werden met GNU C++ versie 2.96. Deze bibliotheek valt onder de GNU GPL, met een uitzondering die

gebruik van de code als onderdeel van een gecompileerd programma toestaat. Dit is analoog aan de uitzondering die de WxWidgets auteurs maken.

In de volgende twee secties worden de gevonden problemen met het Open Source gebruik door in de Aangifteprogramma's beschreven. Verdere problemen zijn tot dusver niet door de auteur gevonden.

De OpenSSL licentie

Zowel de OpenSSL licentie als de SSLeay licentie bevatten de volgende bepaling:

```
* 2. Redistributions in binary form must reproduce the above copyright
*    notice, this list of conditions and the following disclaimer in
*    the documentation and/or other materials provided with the
*    distribution.
```

Verspreiders van OpenSSL code moeten zich aan beide licenties houden en de volledige licentietekst meeleveren in de documentatie bij het programma. OpenSSL code wordt niet gebruikt in het wa2007ux programma, maar vanwege de aanwezigheid in het wa2007d programma dient de Belastingdienst zich aan deze voorwaarde in de OpenSSL licenties te houden. Alleen een link naar de licentievoorwaarden aanbieden is onvoldoende.

GTK+ en de LGPL

De GTK+ bibliotheek wordt verspreid onder de voorwaarden van de LGPLv2. Deze licentie zegt het volgende over het verspreiden van bibliotheekcode als onderdeel van een programma:

6. As an exception to the Sections above, you may also combine or link a "work that uses the Library" with the Library to produce a work containing portions of the Library, and distribute that work under terms of your choice, provided that the terms permit modification of the work for the customer's own use and reverse engineering for debugging such modifications.

You must give prominent notice with each copy of the work that the Library is used in it and that the Library and its use are covered by this License. You must supply a copy of this License. If the work during execution displays copyright notices, you must include the copyright notice for the Library among them, as well as a reference directing the user to the copy of this License. Also, you must do one of these things:

a) Accompany the work with the complete corresponding machine-readable source code for the Library including whatever changes were used in the work (which must be distributed under Sections 1 and 2 above); and, if the work is an executable linked with the Library, with the complete machine-readable "work that uses the Library", as object code and/or source code,

so that the user can modify the Library and then relink to produce a modified executable containing the modified Library. (It is understood that the user who changes the contents of definitions files in the Library will not necessarily be able to recompile the application to use the modified definitions.)

b) Use a suitable shared library mechanism for linking with the Library. A suitable mechanism is one that (1) uses at run time a copy of the library already present on the user's computer system, rather than copying library functions into the executable, and (2) will operate properly with a modified version of the library, if the user installs one, as long as the modified version is interface-compatible with the version that the work was made with.

c) Accompany the work with a written offer, valid for at least three years, to give the same user the materials specified in Subsection 6a, above, for a charge no more than the cost of performing this distribution.

d) If distribution of the work is made by offering access to copy from a designated place, offer equivalent access to copy the above specified materials from the same place.

Samengevat in het Nederlands: De Belastingdienst dient code beschikbaar te maken zodat de ontvanger van de Aangifteprogramma's deze kan linken met gewijzigde versies van de GTK+ bibliotheek. Verder dient de Belastingdienst de broncode van de GTK+ bibliotheek beschikbaar te maken

Op de website waarvandaan het Aangifteprogramma gedownload is, is noch de vereiste broncode, noch linkbare objectcode niet te vinden. Er is ook geen schriftelijk aanbod om deze code te leveren. Zoals eerder vermeld is de GTK + bibliotheek statisch gelinkt, kortom de Belastingdienst voldoet niet aan de eisen van de LGPL en pleegt auteursrechtinbreuk.

Om zo goed mogelijk in overeenstemming met de LGPL te komen zal de Belastingdienst aan alle ontvangers/downloaders van de software alsnog een aanbod voor code waarmee de gebruiker van een Aangifteprogramma zijn eigen versie kan samenstellen moeten doen⁴. Dit geldt ook voor alle Aangifteprogramma's uit voorgaande jaren die statisch gelinkt zijn met de GTK+ bibliotheek. De LGPL stelt daarbij expliciet dat voor programma's die op een medium (floppy of CD) verspreidt zijn de linkbare code ook op een medium aangeboden dienen te worden.

Als alternatief voor het aanbieden van linkbare code staat de LGPL het gebruik van een shared library systeem toe. Immers wanneer de LGPL bibliotheek als shared library, een apart bestand, aangeboden wordt kan de gebruiker van het systeem het bibliotheekbestand vervangen door een bestand naar eigen keuze. Linux heeft de infrastructuur om dit mogelijk te

4 De details omtrend welke code precies geleverd dient te worden staan in de LGPL.

maken. Dit alternatief kan echter niet met terugwerkende kracht ingevoerd worden en is (ook vanwege vereiste testen) alleen relevant voor de Aangifteprogramma's van komende jaren.

Conclusies

De Belastingdienst heeft een aantal kleinere, maar vermijdbare fouten gemaakt bij het samenstellen en distribueren van de Aangifteprogramma's.

Ten eerste mist de volledige tekst van de OpenSSL licentie in de helppagina's en de Belastingdienst dient deze omissie te herstellen.

De wijze waarop GTK+ gebruikt wordt in de Aangifteprogramma's is niet in overeenstemming met de LGPL, de Belastingdienst dient haar gemaakte fout te herstellen door het leveren van code waarmee de gebruiker van een Aangifteprogramma zijn eigen versie kan samenstellen. Deze code zal, met terugwerkende kracht, aangeboden moeten worden aan alle ontvangers en downloaders van Aangifteprogramma's die GTK+ code bevatten. Voor de Aangifteprogramma's in komende jaren kan de Belastingdienst overwegen om dynamisch met de GTK+ bibliotheek te linken, gebruik makend van de standaard faciliteiten voor het delen van bibliotheekcode in Linux.

Zolang de Belastingdienst niet aan alle van toepassing zijnde licentievoorwaarden voldoet, is er sprake van auteursrechtinbreuk en verspreiding van de Aangifteprogramma's behoort gestaakt te worden totdat aan alle licentievoorwaarden voldaan kan worden. Ook zal de Belastingdienst in het verleden gemaakte fouten moeten rechtzetten, dit betekent met terugwerkende kracht een zelfbouwpakket voor de Aangifteprogramma's 2006, 2005 (en mogelijk eerder) aanbieden.

Het verdient aanbeveling dat de Belastingdienst maatregelen treft om vergelijkbare problemen in de toekomst te voorkomen. Dit betekent dat er licentiebewustzijn gekweekt wordt bij ontwikkelaars en dat er een controle op het navolgen van de licentievoorwaarden plaatsvindt voordat met de verspreiding van software wordt gestart.