

Cálculo de Métricas para Medir el Grado de Entendimiento de una Descripción WSDL

Hernán Bernardis⁽¹⁾, Edgardo Bernardis⁽¹⁾, Mario M. Berón⁽¹⁾, Daniel E. Riesco⁽¹⁾, Pedro Rangel Henriques⁽²⁾, Maria Joao V. Pereira⁽³⁾

⁽¹⁾Departamento de Informática / Facultad Ciencias Físico Matemáticas y Naturales/ Universidad Nacional de San Luis
Ejército de los Andes 950 – San Luis – Argentina
{hbernardis, ebernardis, mberon, driesco}@unsl.edu.ar

⁽²⁾ Departamento de Informática/Universidade do Minho
Braga – Portugal
pedrorangelhenriques@gmail.com

⁽³⁾ Departamento de Informática e Comunicações/ Instituto Politécnico de Bragança
Bragança - Portugal
mjoao@ipb.pt

Resumen

En la actualidad, existe un gran incremento en el desarrollo de Web Services. Cada vez más organizaciones crean servicios web y los publican en la red con el objetivo de que otras organizaciones o personas los utilicen para construir diferentes sistemas de software. Esta tendencia ha generado que la Ingeniería de Software se dedique en los últimos años a especificar su uso definiendo estándares y herramientas que faciliten la tarea de definición, construcción, implementación y uso de los mismos. Como resultado, actualmente existen estándares aceptados a nivel mundial que especifican el uso de los mismos. Aprovechándose de estos estándares altamente abstractos y formales en su definición, ha sido posible definir herramientas que permiten construir servicios Web de manera casi automática. Sin embargo, esta alta estandarización ha provocado que la comprensión de los Web Services (WS)

sea cada vez más difícil y tediosa. Comprensión que se vuelve fundamental a la hora de determinar de una gran cantidad de Web Services, cuál usar, para realizar tareas de mantenimiento sobre un Web Services o sobre un sistema que invoca múltiples Web Services.

En este artículo se describe una línea de investigación centrada en facilitar la comprensión de Servicios Web mediante el análisis de sus especificaciones WSDL. Análisis que utiliza diferentes estrategias para calcular métricas que permiten determinar cualitativa y cuantitativamente la dificultad de comprensión que presenta un WS.

Palabras clave: Web Services, métricas, comprensión, LSP.

Contexto

La línea de investigación descrita en este artículo se desarrolla en el Laboratorio de Calidad e Ingeniería de

Software (LaCIS) de la Universidad Nacional de San Luis; y se encuentra enmarcada dentro del proyecto: *“Ingeniería del Software: Conceptos, Métodos, Técnicas y Herramientas en un Contexto de Ingeniería de Software en Evolución”*, perteneciente a la universidad antes mencionada. Dicho proyecto, es reconocido por el programa de incentivos, y es la continuación de diferentes proyectos de investigación de gran éxito a nivel nacional e internacional.

Introducción

Actualmente, con el auge de internet se están popularizando los Web Services como artefactos de software a partir de los cuales se pueden construir sistemas más complejos. Según la W3C, un Web Service es: *“Una aplicación de software identificada por una URI, cuya interface y enlaces son capaces de ser definidos, descritos y descubiertos como artefactos XML. Un web service soporta interacción directa con otros agentes de software usando mensajes basados en XML intercambiados a través de protocolos basados en internet”*. Actualmente, muchas organizaciones desarrollan Web Services y los publican en internet con el objetivo de que otras personas u organizaciones (o incluso ellas mismas) los usen para construir sus propios sistemas de software. Esto ha modificado la forma mediante la cual se desarrollan actualmente los sistemas debido a que se amplía el rango de opciones en el desarrollo. Cualquier persona u organización en el mundo puede crear un Web Service y publicarlo en internet. De manera equivalente, este WS puede ser utilizado por cualquier otra persona u organización en el mundo en un sistema propio sin necesidad de saber de qué manera está implementado, ni donde se

encuentra alojado físicamente. Solamente se necesita su especificación para que el agente de software propio la interprete y se comunique con el WS. Luego, se utiliza la información retornada por el mismo en el sistema propio de acuerdo a lo que se necesite hacer.

La idea de poder usar un Web Service desarrollado por cualquier otra persona u organización en un sistema propio ha sido posible debido a la creación de estándares y lenguajes formales de definición de los mismos. Sin embargo, esta alta abstracción en la construcción y especificación de los mismos dificulta en gran medida su comprensión a la hora de realizar tareas de mantenimiento (adaptación, arreglo de errores, migración, etc.) sobre ellos mismos o sobre los sistemas de software que los utilizan.

Todo Web Service posee una especificación que provee la información necesaria para invocarlo. Uno de los estándares de definición más conocido es WSDL (Web Service Definiton Language) [1]. Las especificaciones WSDL son un dialecto XML, con reglas bien definidas para especificar cada componente del WS. Cuántos parámetros recibe y de qué tipo son, qué datos retorna y de qué tipo, qué protocolo de internet usa para su comunicación, qué operaciones posee, son entre otras tantas, características del WS que se encuentran especificadas en su WSDL asociado. Suena lógico entonces comprender un WS mediante el análisis de su correspondiente especificación WSDL.

Aprovechando la alta estandarización presente en el lenguaje WSDL, se pueden definir metodologías de comprensión de los mismos mediante la extracción y análisis de la información presentes en dichas especificaciones [2, 3, 4].

En este trabajo se extrae información aplicando técnicas de compilación, algoritmos de análisis de lenguaje natural y técnicas de cálculo de indicadores sobre su especificación WSDL. Toda esta información se utiliza para definir, calcular y obtener métricas de los WS's que permitan determinar la dificultad de comprensión que poseen [5, 6]. Además, también se utiliza LSP (Logic Scoring of Preference) para definir estructuras de agregación que le asignen pesos a los valores de cada métrica según sea la necesidad del ingeniero de software y, en base a estos pesos, se obtiene un grado de entendimiento global de la especificación WSDL.

La organización de este artículo se expone a continuación. La sección 2 describe la línea de investigación abordada. La sección 3 presenta los resultados obtenidos hasta el momento, junto con aquellos esperados a corto plazo. Finalmente, la sección 4 describe las tareas realizadas por los recursos humanos en formación.

Líneas de Investigación, Desarrollo e Innovación

El análisis de las especificaciones WSDL posee múltiples etapas con su función particular dentro del proceso global de comprensión. En las subsecciones siguientes se describen brevemente dichas etapas.

Extracción de Información

Debido a que las especificaciones WSDL son un dialecto XML, se pueden usar técnicas de compilación sobre las mismas basadas en los parser DOM (Domain Object Model) [7]. Un parser DOM es un parser específico para el lenguaje XML,

el cual construye una representación interna, basada en árboles, del código fuente XML que recibe. Luego, a partir de funciones especialmente diseñadas para recorrer la representación interna construida (funciones transversales) se puede extraer la información deseada. Usando estas funciones transversales, se logran extraer los identificadores de cada componente (*name, type element, etc.*), la documentación y los comentarios presentes en el WSDL.

Cálculo de Métricas

A partir de la información extraída por el parser DOM descrito en la subsección anterior, se procede al cálculo de múltiples métricas. Las métricas definidas en este trabajo de investigación se pueden dividir en dos grupos:

- **Métricas de tamaño:** permiten medir cantidades de componentes de cada tipo, como por ejemplo cantidad de tipos complejos, de parámetros, de operaciones, de mensajes, entre otras. Esto permite tener una idea del tamaño de cada sección particular del WSDL y determinar que tan complejo es, a primera vista, su comprensión. Estas métricas son muy útiles a la hora de contextualizar al ingeniero de software respecto de la dificultad que puede implicar comprender ese WS. También sirven de gran utilidad a la hora de comprender múltiples WS's debido a que fijan valores de comparación. Así, un WS cuya especificación posea 15 tipos complejos posiblemente demande más esfuerzos de comprensión que uno con 2.
- **Métricas de calidad:** permiten medir la calidad semántica de la especificación WSDL. Esto es, que tanta información semántica brinda la especificación WSDL respecto del WS que representa y que tan entendible y

comprensible en sí es dicha especificación. La presencia de identificadores definidos solamente con iniciales, de errores ortográficos y/o gramaticales, de abreviaciones generan complicaciones en la comprensión de las especificaciones WSDL.

Para el cálculo de estas métricas se utilizan algoritmos de procesamiento de lenguaje natural junto con diccionarios de palabras que permiten determinar el grado semántico que poseen los identificadores, comentarios y la documentación presente en la especificación WSDL. Cabe destacar que sobre los identificadores se aplican algoritmos de división y expansión de palabras de manera tal de poder encontrar palabras abreviadas dentro de sus nombres.

LSP

LSP es un método de evaluación multicriterio basado en un árbol de criterios, una estructura de agregación y un conjunto de funciones de criterios elementales definidos [8]. Combinando sistemáticamente estos elementos, este método permite obtener un nivel de satisfacción que indica, en el caso de esta investigación, el grado de entendimiento que posee la especificación WSDL de un WS. Pero, para lograr obtener este grado de entendimiento, se deben definir las subestructuras de evaluación de cada componente dentro de la especificación WSDL. En la Figura 1 se muestra, a modo ilustrativo, la estructura de agregación específica para medir el grado de entendimiento de los mensajes presentes en un WSDL. Dicha estructura evalúa, para cada mensaje, 3 criterios:

calidad de la documentación, calidad del nombre y entendimiento de cada componente part. A cada uno de estos criterios se le asigna un valor numérico obtenido por las métricas definidas en esta investigación. Luego, los criterios son evaluados usando diferentes operadores propios de LSP con diferentes pesos para cada evaluación (pesos que se muestran en los arcos de la estructura de agregación). El lector interesado en profundizar más sobre estos conceptos puede leer [9, 10].

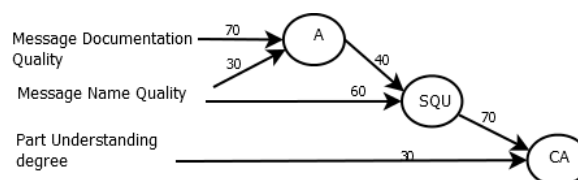


Figura 1. Estructura de agregación para medir el grado de entendimiento de los mensajes.

Resultados y Objetivos

Algunos de los resultados más destacados obtenidos por esta investigación son:

- Se establecieron criterios y métodos para la extracción de información desde las especificaciones WSDL de los WS.
- Se definieron y calcularon diferentes métricas que permiten medir la complejidad de los WS. Dichas métricas miden desde cantidad de componentes de cada tipo presentes en el WSDL como calidad semántica de la especificación mediante el análisis semántico de identificadores, documentación y comentarios existentes en la especificación WSDL.
- Se utilizó LSP para construir una estructura de agregación que permita relacionar los valores de todas las

métricas mediante operadores para obtener un grado de entendimiento global del WS.

- Se construyó WSDLUD, una herramienta que realiza la extracción de información, cálculo de métricas, el cálculo del grado de entendimiento del WS usando LSP y la visualización de la información de múltiples WS de manera automática.

Entre los objetivos planteados a corto y largo plazo se pueden mencionar:

- Detectar y analizar automáticamente relaciones entre WSDLs mediante el estudio de sus espacios de nombres.
- Realizar un análisis más exhaustivo sobre los tipos complejos usados en cada WS.

Formación de Recursos Humanos

Las tareas realizadas en el contexto de la presente línea de investigación están siendo desarrolladas como parte de trabajos para optar al grado de Magister en Ingeniería de Software. En el futuro se piensa generar diferentes tesis de maestría y doctorado a partir de los resultados obtenidos de los trabajos de licenciatura en curso.

Referencias

[1] WSDL Specification for W3C <https://www.w3.org/TR/wsdl>.

[2] N. Gold and K. Bennett. "Program comprehension for web services". In Program Comprehension, 2004. Proc. 12th IEEE International Workshop on. June 2004.

[3] L. O'Brien Lero and D. Smith. "Working session: program comprehension strategies for web service and service oriented archi-

tures". Proc. of 12th IEEE International Workshop on Program Comprehension. 2004.

[4] H. El Bouhissi, M. Malki, and D. Bouchiha. "A reverse engineering approach for the web service modeling ontology specifications". In Sensor Technologies and Applications 2008. SENSORCOMM '08. Second International Conference on, pages 819–823, Aug 2008.

[5] C. Mao. "Towards a data complexity metric set for web service composition". In Computer and Information Technology (CIT), 2011 IEEE 11th International Conference on, pages 127–131, Aug 2011.

[6] Fangfang Liu, Yuliang Shi, Jie Yu, Tianhong Wang, Jingzhe Wu. "Measuring Similarity of Web Services Based on WSDL". IEEE International Conference on Web Services ICWS. 2010.

[7] Parser DOM specification for W3C. <https://www.w3.org/DOM>.

[8] Olsina, L., Rossi, G.: Measuring Web Application Quality with WebQEM. IEEE MultiMedia, 2002 09(4), 20–29 (2002).

[9] Mario M. Berón, Hernán Bernardis, Enrique A. Miranda, Daniel E. Riesco, Maria João Pereira, Pedro Rangel Henriques. "WSDLUD: a Metric to Measure the Understanding Degree of WSDL Descriptions". Proceedings of the 2015 Symposium on Languages, Applications and Technologies, SLATE'15. Madrid, España 2015.

[10] Bernardis, Hernán; Berón Mario; Bernardis, Edgardo; Riesco, Daniel; Henriques, Pedro. "Extracción de información y cálculo de métricas en WSDL 1.1 y 2.0". II Congreso Nacional de Ingeniería Informática / Sistemas de información (CoNaIISI). Argentina. 2014.