

RESDEC: Un prototipo de herramienta para la selección de configuraciones de despliegue basado en Sistemas de Recomendación

Jorge L. Rodas¹, José A. Galindo², and David Benavides²

¹ Universidad Estatal de Milagro, Facultad de Ciencias de la Ingeniería, Cdla. Universitaria Km 1 1/2 vía Km 26, Milagro - Ecuador,

`jrodass@unemi.edu.ec`

² Universidad de Sevilla, Dept. Lenguajes y Sistemas Informáticos, Av. Reina Mercedes s/n Sevilla - España,
[jagalindo,benavides]@us.es

Abstract. Los sistemas de alta variabilidad son sistemas que representan cientos de configuraciones distintas. En un contexto particular, estas configuraciones pueden ser desplegadas en distintos entornos de despliegue lo cual es una decisión crítica para el correcto funcionamiento de la misma. Por ejemplo, determinar en qué dispositivos móviles se va a ejecutar correctamente nuestra app, es una tarea difícil de resolver. En este artículo presentamos RESDEC, un prototipo de herramienta para asistir al ingeniero de software en la toma de decisiones para el despliegue. Concretamente RESDEC provee algoritmos de recomendación para tres escenarios de despliegue distintos.

Keywords: Sistemas de alta variabilidad, Sistemas de recomendación, Valoraciones, Configuraciones

1 Antecedentes

Los sistemas de alta variabilidad son sistemas de software cuyo comportamiento se puede personalizar de acuerdo con las necesidades específicas de un contexto particular [2]. Los modelos de características [7] se han convertido en el estándar *de facto* para representar características comunes y variables en Variability-intensive System (VIS). Cuando el número de combinaciones en un modelo crece exponencialmente, hace que el análisis y la mejora de dichos sistemas sea una actividad costosa y propensa a errores. Para aliviar el problema, investigadores han propuesto el uso del análisis automatizado de modelos de características [3]. Sin embargo, este problema se vuelve cada vez más difícil cuando hay múltiples y diversas plataformas para implementar un producto. Por ejemplo, para decidir en qué plataformas de implementación es mejor instalar Drupal o Wordpress, o en qué teléfonos móviles debemos instalar una aplicación de Android [5]. En tales casos, la selección de la plataforma de implementación óptima se convierte en una actividad crítica que puede afectar la experiencia del usuario.

En la industria, podemos encontrar propuestas basadas en propiedades estructurales y atributos de calidad que apuntan a seleccionar configuraciones

óptimas para una variedad de casos de uso, como el testing o la predicción del rendimiento. Sin embargo, hay una falta de soporte para beneficiarse de otros artefactos, tales como informes de usuarios o rankings de aplicaciones que evalúan qué tan bien se realizó una configuración en una plataforma de implementación concreta. En este documento, proponemos explotar informes de usuarios para guiar la selección de plataformas de implementación; esto es, para encontrar qué plataformas de implementación se adaptan mejor a un usuario concreto y a una configuración específica.

En consecuencia, hemos encontrado en los sistemas de recomendación una alternativa ideal para guiar al usuario en la búsqueda de configuraciones de implementación que satisfagan sus necesidades. Un sistema de recomendación es un software que, en función de los datos existentes, intenta predecir productos para que se adapten mejor a las necesidades o gustos de un usuario [4]. En este paper presentamos RESDEC(Recommendation System for Selecting Deployment Configurations), un prototipo de herramienta basada en sistemas de recomendación que ayuda al ingeniero de software a seleccionar el mejor entorno de despliegue para configuraciones de VIS basadas en informes de usuarios.

2 RESDEC: Un prototipo de herramienta para recomendar estrategias de despliegue

RESDEC fue diseñado para soportar dos fases tal como se muestra en la Figura 1.

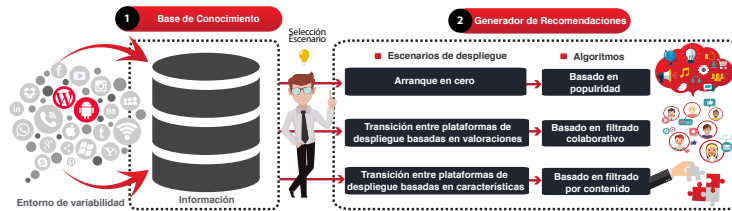


Fig. 1. Fases para establecer recomendaciones en escenarios de despliegue

2.1 Base de conocimiento:

La base de conocimiento contiene la información sobre la satisfacción de los usuarios en despliegues anteriores. Esta información puede representar criterios de evaluación dados previamente por los usuarios en relación a gustos, funcionalidad, errores o desperfectos que puede presentar un producto. En nuestro caso, los valores que se extraen de los informes de usuarios, reflejan los niveles de satisfacción en relación al uso de un producto por parte del usuario.

2.2 Generador de Recomendaciones:

RESDEC se ejecuta en tres escenarios diferentes de acuerdo a las necesidades específicas del ingeniero de software:

Primer despliegue o arranque en cero. El primer escenario se presenta cuando un usuario nuevo del sistema desea conocer qué configuraciones serían las más adecuadas para un proyecto en particular. Esto ocurre por ejemplo, cuando el ingeniero de software construye un sitio web nuevo y necesita una asistencia guiada en la selección de plugins que podrían ser instalados dentro del sitio.

Para realizar recomendaciones en este escenario hemos empleado un algoritmo basado en popularidad, perteneciente al grupo de algoritmos no personalizados. Aunque el método de recomendación que utiliza este algoritmo parece simple, es uno de los más populares en escenarios reales, tales como: la lista de productos más vendidos en *Amazon*, los vídeos con mayor visitas en *Youtube*, las canciones más escuchadas en *Spotify*, las películas más vistas en *Netflix*, entre otros.

Transición entre plataformas de despliegue basadas en valoraciones. El segundo escenario se presenta cuando un usuario que ha usado previamente una configuración de un producto quiere explorar otras configuraciones para obtener nuevas alternativas de recomendación. Esto ocurre, por ejemplo, cuando el ingeniero de software tiene publicado un sitio web que incorpora una serie de plugins y basado en esa experiencia desea conocer qué otros plugins similares a los instalados podrían también implementarse; para ello, el sistema utiliza información de unos usuarios para producir recomendaciones a otros.

Para realizar este tipo de recomendaciones hemos integrado en RESDEC tres algoritmos de recomendación basados en filtrado colaborativo (FC). Según la clasificación, del grupo Neighborhood-based Collaborative Filtering (Knn) [6] hemos seleccionado el algoritmo "KNNBasic" y el algoritmo "KNNWithMeans"; y del grupo Model-based Collaborative Filtering hemos elegido el algoritmo de factorización de matrices SVD (Singular Value Decomposition).

Transición entre plataformas de despliegue basadas en características. El tercer escenario además de utilizar las valoraciones que recibe una configuración, también considera las características que el usuario define para las configuraciones que el sistema devuelve. Cuando un usuario especifica las características requeridas para una configuración es fácil para el sistema encontrar las configuraciones que mejor se adapten dentro de esos requisitos y de esa manera establecer las mejores recomendaciones.

Esto ocurre por ejemplo, cuando el ingeniero de software está desarrollando un sitio web y necesita instalar un conjunto de plugins que se ajusten a los requerimientos del sitio, en este caso a través de los *tags* asociados a los plugins se podrían encontrar fácilmente plugins para el catálogo de productos, el buscador de productos, la pasarela de pagos, el carrito de compras entre otros componentes que podrían requerirse para el correcto funcionamiento del sitio web.

Para hacer recomendaciones basados en la descripción de cada configuración, hemos integrado en RESDEC dos algoritmos basados en filtrado por contenido, *Rocchio* [1] y *kNN* [8], este último similar al utilizado por el grupo de filtrado colaborativo no obstante su variación radica en la función de similitud utilizada.

2.3 Datasets y validación.

Dataset de dispositivos Android: Para recopilar la información de Android, se realizó una encuesta a 2,099 usuarios de la comunidad de estudiantes y docentes de la Universidad Estatal de Milagro, Ecuador. Esto nos permitió conocer qué aplicaciones son las más utilizadas por los usuarios, los dispositivos que utilizan y las puntuaciones que otorgan a las aplicaciones instaladas en sus dispositivos de acuerdo al funcionamiento que han tenido en el mismo. Después de conocer los dispositivos usados por los usuarios encuestados, obtuvimos información explícita sobre las características de cada uno de ellos a través de información disponible en el sitio web <http://www.gsmarena.com/>.

Dataset de Wordpress: Este dataset contiene información sobre el gestor de contenidos wordpress. Para obtener información de este entorno, se desarrolló un crawler que extrae datos que relaciona los plugins con los tags asociados a los mismos. Esto nos permitió conocer qué plugins son los más utilizados por los usuarios y las puntuaciones que se les otorgan. La búsqueda siguió un proceso sistemático para extraer datos utilizando el sistema de prueba Selenium (<http://www.seleniumhq.org/>).

Resultados: Nuestras pruebas incluyeron la generación automática de recomendaciones de entre un conjunto de datos de 832 dispositivos móviles en Android y 634 plugins de Wordpress.

Acceso al código y demostración del prototipo. El código fuente de RESDEC se puede descargar desde el sitio web del proyecto <https://github.com/RESDEC>. La herramienta junto con el conjunto de datasets presentados en este trabajo, está disponible en el sitio web de RESDEC <http://resdec.com/>.

References

1. Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6):734–749, 2005.
2. Muhammad Ali Babar, Lianping Chen, and Forrest Shull. Managing variability in software product lines. *Software, IEEE*, 27(3):89–91, 2010.
3. David Benavides, Sergio Segura, and Antonio Ruiz-Cortés. Automated analysis of feature models 20 years later: A literature review. *Information Systems*, 35(6):615–636, 2010.
4. Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. Recommender systems survey. *Knowledge-Based Systems*, 46:109–132, 2013.
5. José A Galindo, Hamilton Turner, David Benavides, and Jules White. Testing variability-intensive systems using automated analysis: an application to android. *Software Quality Journal*, pages 1–41, 2014.
6. Anil K Jain. Data clustering: 50 years beyond k-means. *Pattern recognition letters*, 31(8):651–666, 2010.
7. Kyo C Kang, Sholom G Cohen, James A Hess, William E Novak, and A Spencer Peterson. Feature-oriented domain analysis (foda) feasibility study. Technical report, DTIC Document, 1990.
8. Michael Pazzani and Daniel Billsus. Learning and revising user profiles: The identification of interesting web sites. *Machine learning*, 27(3):313–331, 1997.

Apéndice A: Tiempo planeado para la presentación de la herramienta

5min – Presentación del problema y alcance de la herramienta.

5min – Presentación del problema al que se dirige RESDEC con un breve ejemplo de la aplicación.

5min – Demostración de los tipos de recomendación que RESDEC emplea para diferentes escenarios.

5min – Demostración de las capacidades de la herramienta para seleccionar y priorizar configuraciones en un conjunto de datos de dispositivos móviles Android.

5min – Conclusiones del trabajo

Apéndice B.- Capturas de pantalla de la herramienta:
Caso ANDROID



Fig. 2. Pantalla principal de RESDEC, en la parte izquierda un menú flotante muestra las opciones agrupadas por el entorno de despliegue y el gestor de recomendaciones.











Products					
1 2 3 4 5 6 7 8 9 10					
	Id	Brand	Model	State	
	4946205726	APPLE	IPHONE 4	Activo	
	4946215804	SAMSUNG	GALAXY S5	Activo	
	4946232631	SAMSUNG	GALAXY J5	Activo	
	4946234750	SAMSUNG	GALAXY S6	Activo	

Fig. 3. La ventana de productos permite la gestión (creación, modificación y eliminación) de los productos (dispositivos) pertenecientes al entorno de variabilidad mediante el cual el sistema establece las recomendaciones.



Users		
Id	Name	
1	B612	
2	4SHARED App	
3	Adobe Acrobat Reader App	
4	Amino app	
5	Animusic PodCast App	
6	Apple iTunes Store	
7	Ask	
8	Asphalt Airborne 8 App	
9	Banca Movil App	
10	Biblia Santa Valera App	
11	Biblioteca del Evangelio de la Iglesia App	
12	Body Human Anatomy App	

Fig. 4. La ventana de Apps, permite la gestión (creación, modificación y eliminación) de las aplicaciones que fueron desplegadas en los dispositivos registrados en la opción productos.



Fig. 5. Muestra los resultados obtenidos cuando se ejecuta el primer escenario de nuestra propuesta, en este caso se observa las recomendaciones basadas en el top de popularidad.

RESDEC
Recommendation System

Configuration Transition Based on Ratings	
Method:	Coeff. Pearson
Base user:	1 B612
Base Product:	494633786 SAMSUNG - GALAXY S4 MINI
Number of records:	5
Sort:	Falling
Calculate	

SAMSUNG - GALAXY S4	SAMSUNG - ACE 3	SAMSUNG - GALAXY S3 MINI
Coeff. Pearson -0.49999999999999994	Coeff. Pearson 0.5	Coeff. Pearson 0.5

Fig. 6. Muestra los resultados obtenidos cuando se ejecuta el segundo escenario de nuestra propuesta, en este caso se observa las recomendaciones basadas en el algoritmo de filtrado colaborativo.



Fig. 7. Muestra los resultados obtenidos cuando se ejecuta el tercer escenario de nuestra propuesta, en este caso se observa las recomendaciones basadas en el algoritmo de filtrado por contenido.