

# Importancia de los interesados en el problema de la siguiente versión\*

José Antonio Sierra Ibañez<sup>1</sup>, Isabel M. del Águila<sup>1</sup>[0000-0001-9896-7196] and José del Sagrado<sup>1</sup>[0000-0002-5616-056X]

Universidad de Almería, 04120 Almería, [imaguila@ual.es](mailto:imaguila@ual.es), [jsagrado@ual.es](mailto:jsagrado@ual.es)

**Resumen** En la disciplina de la Ingeniería del Software Basada en Búsqueda el problema de la selección de los requisitos a desarrollar en la siguiente versión recoge en su formulación la necesidad de disponer de un peso asignado a cada cliente/interesado que fija su importancia en el proyecto. El cálculo de este peso o cuantificación de los interesados es un problema que debe abordarse antes del problema de búsqueda en sí mismo. Esta definición de los pesos es especialmente importante si se maneja un alto número de interesados, o cuando el mercado requiere tratar con detalle como se cubren las necesidades de todos los clientes. Este trabajo propone la utilización de algoritmos de agrupamiento, para identificar los interesados críticos en el proyecto y estudia, en el caso definido por el conjunto de datos para el proyecto *Replacement Access, Library and ID Card* (RALIC), si la distribución en grupos genera resultados similares a la utilización de los datos de cuantificación de los interesados. Finalmente, se estudia como influye la propuesta en los resultados obtenidos, en especial como los frentes de Pareto obtenidos suponen una alternativa válida para los distintos grupos de interesados.

**Keywords:** Problema de la siguiente versión · Optimización multiobjetivo · Algoritmos de agrupamiento.

## 1. Introducción

El problema de la siguiente versión (NRP) [4,1] consiste en encontrar el conjunto de requisitos o clientes que maximicen la satisfacción de los interesados, teniendo en cuenta que no todos son igualmente importantes para el proyecto, y sin infringir la restricción del límite de esfuerzo de desarrollo disponible. Para ello se formaliza como un problema de optimización monobjetivo el concepto de optimización coste-valor [15]. La formulación bi-objetivo del problema [26] transforma el coste en un segundo objetivo a optimizar, obteniendo soluciones dentro del concepto de la Pareto optimalidad.

A pesar de tratarse de un problema con una amplia cobertura y suficientemente implantado en el ámbito de la Ingeniería del Software Basada en Búsqueda

---

\* Financiado a través de RED2018-102472-T, proyecto PID2019-106758GB-C32 (EML-PA) y Universidad de Almería (UAL)

da, existe un aspecto que, si bien está en su formulación, no ha sido estudiado en la mayoría de las propuestas de solución [21], la importancia de los interesados.

En la formulación del NRP, se da por sentado que se dispone de un peso asignado a cada cliente/interesado que fija su importancia en el proyecto, pero ¿cómo se calcula y quién asigna este valor?. Esta cuestión está en el dominio de la llamada cuantificación de los interesados, donde en base a la información disponible, como puede ser la experiencia o el nivel de autoconfianza de la persona interesada, se establece cuales son las personas claves para definir los requisitos del proyecto [20,2].

En este trabajo estudiamos alternativas para la definición de los pesos/ importancia de los clientes/stakeholders en el ámbito del problema de la selección de requisitos para la siguiente versión (NRP), comprobando como afecta a las soluciones obtenidas por los métodos de optimización. En especial se estudia si las técnicas de agrupamiento o clustering, puesto que son una manera de dividir un conjunto de datos dado en grupos que contengan elementos que sean similares entre sí, permiten identificar los interesados del mismo peso, sobre todo cuando tenemos un alto número de ellos, facilitando a su vez la incorporación de nuevos interesados.

Este trabajo se organiza realizando una revisión de como se tratan los interesados en NRP en la sección 2, para en la sección 3 mostrar las principales enfoques de cuantificación de los interesados. Los conceptos básicos necesarios en este caso para aplicar el agrupamiento a la búsqueda de interesados aparecen en la sección 4, que introduce la sección 5 donde se detalla un caso de estudio sobre el proyecto RALIC. Finalmente se incluyen las conclusiones en la sección 6.

## 2. Gestión de los interesados en NRP

El problema NRP ha supuesto y sigue suponiendo un reto debido a su complejidad, las interacciones entre los elementos que lo componen, el número y diversidad de los implicados, la variedad de variables que necesitan ser revisadas y la incertidumbre de la información en la que se basa [22].

La formulación más básica del NRP incluye al conjunto de los interesados,  $\mathbf{C} = \{c_1, c_2, \dots, c_k\}$  y un conjunto requisitos candidatos,  $\mathbf{R} = \{r_1, r_2, \dots, r_m\}$ . Los clientes se cuantifican con un peso,  $\mathbf{W} = \{w_1, w_2, \dots, w_k\}$ , puesto que no todos son igualmente importantes en el proyecto. Cada cliente asigna un valor  $v_{ij}$  a cada requisito, que representa el valor añadido que el cliente  $c_i$  estima que obtendrá si el requisito  $r_j$  se incluye en la siguiente versión. Se define por tanto para cada  $r_j$ , la satisfacción  $s_j$ , como:  $s_j = \sum_{i \in C} w_i \cdot v_{ij}$ , que denotamos  $\mathbf{S} = \{s_1, s_2, \dots, s_m\}$ . Además cada  $r_j$  tiene un esfuerzo  $e_j$  que estima el esfuerzo de desarrollo asociado a su incorporación en el producto final,  $\mathbf{E} = \{e_1, e_2, \dots, e_m\}$ .

El problema es seleccionar el conjunto de requisitos a incluir en la siguiente versión de software  $\hat{\mathbf{R}} \subseteq \mathbf{R}$  que maximice las satisfacción total, minimizando el esfuerzo asociado, dentro del límite de esfuerzo  $B$  determinado por la capacidad o velocidad del equipo de desarrollo.

Estos dos objetivos (esfuerzo y satisfacción) están fuertemente implantados en los trabajos de Ingeniería del Software Basada en Búsqueda aplicados a NRP y son pocas las propuestas que añaden objetivos adicionales, pero estos dos siempre están presentes. Algunos trabajos plantean incorporar el número de requisitos y la equidad [7] o un objetivo por cliente para considerar su cobertura [25], o incluso utilizan la penalización además de la satisfacción (asociada al modelo KANO) [19].

Pero el análisis del nivel de los clientes no está expresamente tratado en ninguno de ellos y, con las actuales tendencias de desarrollo de productos software, la llamada ingeniería del software global, es muy habitual tener cientos de clientes con intereses heterogéneos y poco uniformes. O incluso yendo más lejos, los interesados van desde los clientes que pagan por el sistema de software, a los usuarios que interactúan con el sistema, o incluso las autoridades que imponen reglas y normativas sobre el desarrollo y funcionamiento del mismo. Todos ellos plantean requisitos diferentes y que muy probablemente entrarán en conflicto. La visión simplista plantea que los pesos de los clientes son una estimación directa, quizás del equipo de desarrollo, pero la realidad es que la determinación de su importancia es un problema en sí mismo [24,13].

La opción de incluir una función objetivo por interesado [25] no es viable, pensemos por ejemplo en casos como el proyecto *Replacement Access, Library and ID Card* (RALIC) [16], donde se trabaja con entre 50 y 79 personas. La alternativa es mantener la formulación conocida (con más o menos objetivos) y trabajar a nivel de cuantificación de los pesos/importancias para el proyecto en una fase previa de inicialización del problema de optimización NRP.

### 3. Cuantificación de los pesos de los interesados

La información necesaria a inyectar en NRP se obtiene de los documentos y/o artefactos del proyecto software y su cálculo puede ser simplemente una consulta a los datos del requisito, como sucede con el esfuerzo, o puede implicar el cálculo bastante más complejo que conlleve la determinación de los pesos de los clientes, como en el caso de la satisfacción, que a su vez deben ser calculados en base a los datos disponibles sobre los interesados.

Las interacciones y la identificación de los interesados aparecen en diversos trabajos [20], pero su cuantificación implica valorarlos para identificar aquellos que son realmente relevantes para el proyecto. Se han planteado soluciones que aplican un algoritmo genético para identificar los clientes claves [17], también hay aproximaciones basadas en valor, como la que utiliza nueve factores estimados sobre los interesados que son agregados para dar una valoración cuantitativa [2] o la que propone el uso semiautomático de métodos de decisión multicriterio [12].

Algunas técnicas de cuantificación y valoración de los interesados proponen el uso de una serie de factores, como el riesgo, el nivel de conocimientos o la experiencia, que se cuantifican a través de la ponderación de una serie de atributos para asignar un valor y priorizar a los grupos de interés, [2,12]. También

se agrupan a los interesados en función de las responsabilidades, basadas en la estructura funcional del caso de estudio concreto [2]. Sin embargo, otras propuestas buscan la agrupación más genérica basándose en las habilidades, ubicación geográfica o nivel en la jerarquía, aplicando técnicas de agrupamiento [3]. En cualquier caso, las propuestas para la cuantificación implican manejar y, lo que es más difícil, tener disponible, un gran número de datos, estimados o calculados, de diferentes atributos sobre los interesados como son: la experiencia, la habilidad de gestión, el conocimiento del dominio, la práctica sobre el dominio del interesado o la autoestima o la objetividad. Esta disponibilidad, en el contexto de la ingeniería del software global en la que nos movemos actualmente, es casi irreal puesto que con solo tener de orden del cientos de interesados, la mayoría de estos atributos tendrán un valor desconocido.

Si bien no hay uniformidad en las técnicas de cuantificación de los interesados, en lo que si coinciden la mayoría de los autores es en utilizar como base, para la identificación y la cuantificación, tanto de las relaciones, los papeles y la influencia de los interesados entre sí, y en la necesidad de agrupar a todos los interesados de acuerdo con sus implicaciones en el proyecto [20,17,14]. Se han de buscar mecanismos que permitan identificar el conjunto de interesados críticos para el proyecto para poder tomar las decisiones oportunas [12,14].

Puesto que el objetivo es identificar grupos de objetos similares, en este caso interesados, el clustering o algoritmos de agrupamiento son una alternativa. Esta técnica no supervisada también puede considerarse como un proceso de análisis exploratorio de datos que ayuda a descubrir patrones ocultos de interés o estructura en los datos, en este caso se busca identificar los pesos de los clientes. No obstante, como señala [6], dado un problema *el agrupamiento está en el ojo del observador* y, a menudo, el algoritmo de agrupación más adecuado debe elegirse experimentalmente. Al tratarse de un trabajo incipiente, cuyo objetivo es buscar grupos de interesados, en lugar de centrarnos en la selección del "mejor" algoritmo de agrupamiento, nos centraremos en la localización de grupos en función de la cercanía empleando el algoritmo  $k$ -medias.

#### 4. Búsqueda de grupos de interesados

El agrupamiento por  $k$ -medias es un algoritmo de aprendizaje no supervisado, ya que no conoce el grupo al que pertenece cada observación, que se emplea para encontrar patrones en los datos (p.ej. segmentos de clientes o productos) y que ya ha sido utilizado en el ámbito de la selección de requisitos [23]. Para construir los grupos con el algoritmo  $k$ -medias, primero hemos de indicar el número  $k$  de grupos en los que queremos agrupar los datos. Existen distintos métodos para determinar el número óptimo de grupos que incluyen:

- Métodos directos, que consisten en optimizar un criterio, como las sumas de cuadrados dentro del grupo (método del codo) o la silueta promedio (método de la silueta).
- Métodos de contraste de hipótesis como el estadístico estadístico de espaciamiento (gap statistic).

Una vez determinado el número de grupos, el algoritmo asigna de forma aleatoria cada observación a uno de los  $k$  grupos y encuentra el centroide de cada uno de ellos. El algoritmo de las  $k$ -medias [11] reparte las observaciones en  $k$  grupos de manera que la suma de los cuadrados de las diferencias entre las observaciones y el centroide del grupo, al que se han asignado, sea mínima.

A la hora de identificar la importancia de los interesados, aplicando  $k$ -medias, es necesario conocer a priori el número de clusters  $k$ , es decir el número de tipos distintos de interesados implicados en un proyecto. Sin embargo este valor no está unificado, [10]. Trabajos clásicos fijan desde 2 (primarios, secundarios) [9,8] a 7 los tipos de stakeholders diferentes [18], pero también puede ser 4 [5]. Por este motivo, consideraremos distintos números de grupos, además del óptimo.

Tomando como referencia el conjunto de datos del proyecto RALIC, que es considerado como el único y completo disponible en este área [12], plantearemos distintas alternativas para la obtención de los pesos de los clientes. Siguiendo las etapas que se describen a continuación:

- Educación de votos y recomendaciones recibidas por los interesados.
- Determinación del número de grupos candidatos.
- Aplicación del algoritmo  $k$ -medias para localizar los grupos de interesados.
- Identificación de la importancia de los grupos.
- Definición y resolución del NRP obtenido con los valores de importancia asignados.

## 5. Estudio de un caso

Para este trabajo se ha utilizado el conjunto de datos RALIC que define objetivos a distinta granularidad, con 10 objetivos, 45 requisitos y de 83 requisitos de específicos, como refinamiento de los 45 del nivel superior, que han sido sugeridos por 76 interesados. La asignación de la satisfacción, tratada en este caso de estudio, es la que se ha realizado utilizando el método de los 100 Puntos, de forma que a cada persona implicada en la valoración de los requisitos se le entregan 100 puntos que tiene que repartir entre los elementos a evaluar.

De esta forma el punto de partida para NRP son los 10 objetivos con su estimación de esfuerzo y las satisfacciones definidas por los 76 interesados, teniendo en cuenta que algunos  $v_{ij}$  estarán vacíos. Para obtener las satisfacciones será necesario identificar los pesos de cada interesado o bien definir grupos de interesados similares a etiquetar con un peso. Se parte de las relaciones entre los interesados que se expresan a través del atributo *nivel de influencia* definido entre ellos. Dados dos interesados  $c_i$  y  $c_j$ , se define la función  $influencia(c_i, c_j)$  en el intervalo  $[1, 8]$ , como la estimación que  $c_j$  (recomendador) fija según su conocimiento de  $c_i$  (recomendado) acerca de lo importante que es el recomendado para el proyecto RALIC. Pero en este caso de los 127 interesados, solo 60 hacen recomendaciones y con una media de 29/127 recomendados, confirmado la falta de estimaciones completas, incluso cuando sólo hay un atributo, en *nivel de influencia* y solo 76 de los 127 recogen valoraciones para los requisitos ( $v_{ij}$ ). Nuestra propuesta es identificar cuales son los interesados importantes para el

problema NRP, que son aquellos de deben ser cubiertos prioritariamente por las soluciones del problema de optimización.

Para estudiar la importancia de los interesados tendremos en cuenta, tanto su influencia, como el número de interesados que lo recomiendan. Así, consideraremos que la importancia  $w_i$  de cada cliente  $c_i$  viene dada por la razón de las recomendaciones recibidas entre el número de clientes que lo han recomendado:

$$w_i = \frac{\sum_{c_j} influencia(c_i, c_j)}{\#\{influencia(c_i, c_j) \neq 0\}}. \quad (1)$$

Con esta asignación de pesos, construimos la definición del problema de la siguiente versión calculando los valores de satisfacción asociados a los requisitos (ver Fig. 1). El frente de Pareto obtenido nos servirá como punto de referencia para indagar el efecto de la asignación de pesos basada en el agrupamiento de los interesados.

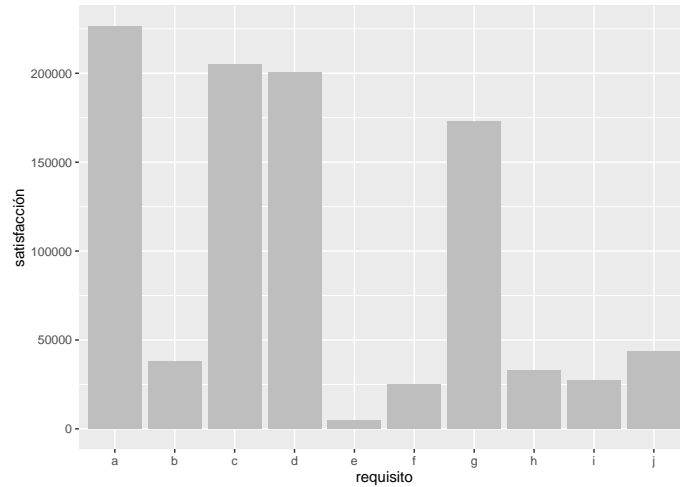


Figura 1: Valores de satisfacción en base a la importancia de cada interesado.

Antes de utilizar el algoritmo k-medias y agrupar a los interesados, debemos indagar el número  $k$  de grupos a emplear. El método del codo y el de la silueta sugieren 2 grupos, mientras que el estadístico de espaciamento sugiere 6. En el estudio fijaremos en 2, 3, 5 y 6 el número de tipos (grupos) de interesados implicados en un proyecto. Descartamos utilizar 4 grupos debido a que en ese punto el método del codo detecta un cambio de tendencia que no se produce en 3 y 5. Los valores fijados se encuentran dentro del rango sugerido en trabajos previos y nos permitirán observar las posibles variaciones que puedan producirse al asignar los valores de satisfacción asociados a los requisitos.

Una vez establecido el número de tipos de interesados implicados en el proyecto, utilizamos el algoritmo de agrupamiento para asignar los interesados a

grupos. La figura 2 muestra los grupos detectados y la tabla 1 el promedio de votos y recomendaciones recibidos por los interesados en cada grupo.

Tabla 1: Detalle de los grupos de interesados según el número  $k$  de tipos usado.

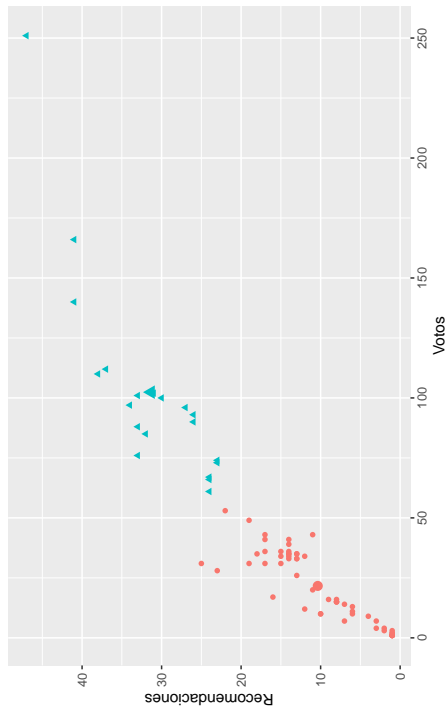
$k$	Grupo	Tamaño	Votos	Recomendaciones
2	1	56	21.64	10.36
	2	20	102.40	31.35
3	1	28	8.00	4.96
	2	33	40.27	16.94
	3	15	113.80	33.93
5	1	26	7.39	4.46
	2	1	251.00	47.00
	3	15	84.60	28.20
	4	4	132.00	39.25
	5	30	34.00	15.47
6	1	10	92.80	30.50
	2	9	55.78	23.00
	3	26	33.04	14.42
	4	1	251.00	47.00
	5	26	7.39	4.46
	6	4	132.00	39.25

Los grupos se ordenarán de forma creciente según su influencia media y tomaremos como peso del grupo la posición que ocupe en esta ordenación. Todos los interesados que pertenezcan al mismo grupo recibirán el mismo peso. Tras fijar los pesos, se obtendrá la satisfacción de los requisitos (ver Fig. 3) y calcularemos el frente de Pareto para el NRP resultante. La Figura 4 recoge los frentes de Pareto obtenidos y la Tabla 2 el número de soluciones que contienen y cuántas de ellas están en el frente de referencia.

Tabla 2: Número de grupos de interesados tenidos en cuenta para fijar la importancia en el problema de la siguiente versión, número de soluciones del frente de Pareto y coincidencias con el frente de referencia que contiene 40 soluciones.

$nro. grupos$	$nro.soluciones$	$coincidencias$
2	49	37
3	45	33
5	41	38
6	45	34

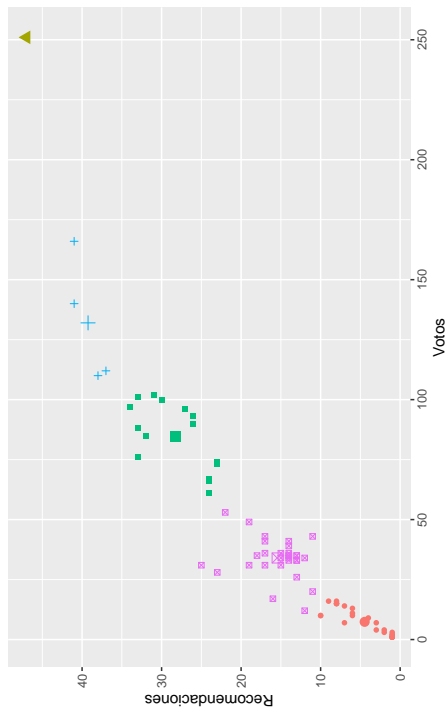
La importancia de los interesados influye sobre la satisfacción del problema y sus soluciones proporcionarán distintas coberturas de los interesados. La definición de los pesos en base a la identificación de la importancia persigue fo-



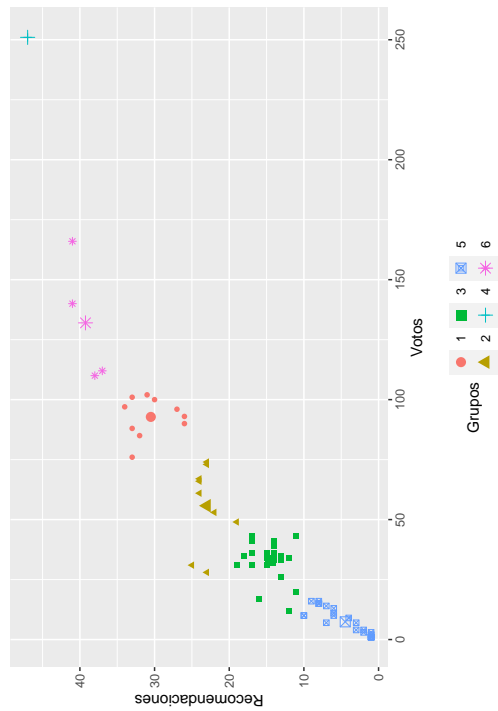
(a) Dos grupos de interesados.



(b) Tres grupos de interesados.



(c) Cinco grupos de interesados.



(d) Seis grupos de interesados.

Figura 2: Distintas divisiones de los interesados en grupos.



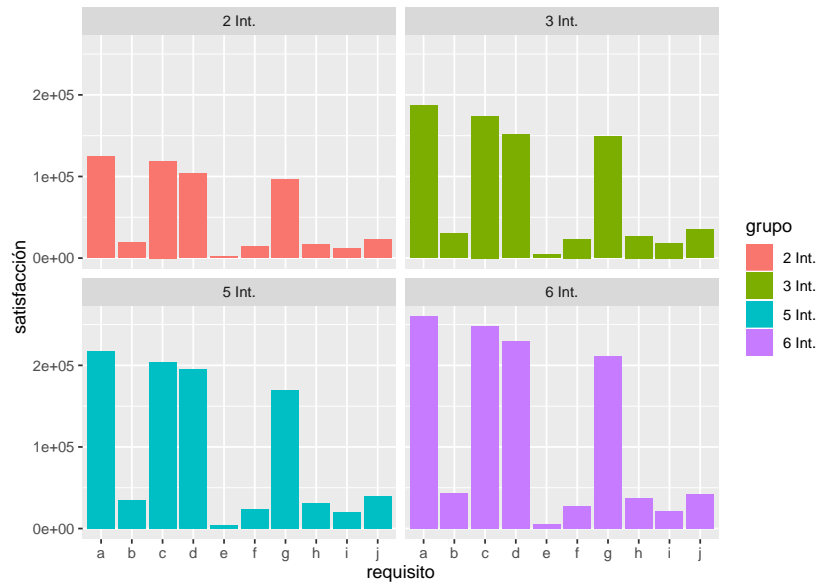


Figura 3: Valores de satisfacción obtenidos para los diferentes grupos de interesados considerados.

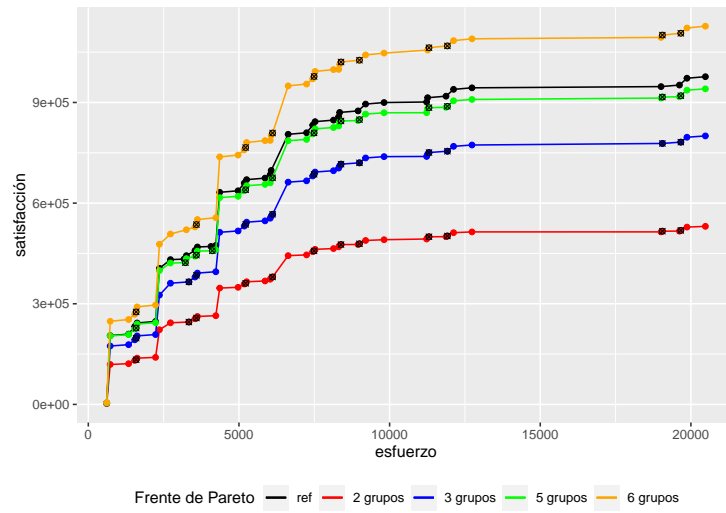


Figura 4: Frentes de Pareto obtenidos para los diferentes grupos de interesados considerados. Las soluciones marcadas no están en el frente de referencia (en negro).

mentar la cobertura de aquellos interesados que tengan una mayor importancia. Definimos la cobertura que una solución  $s$  da a un determinado interesado  $c_i$  como

$$cob(s, c_i) = \frac{|\{r_j | r_j \in s \wedge v_{ij} \neq 0\}|}{|\{r_j | v_{ij} \neq 0\}|}, \quad (2)$$

la proporción de requisitos propuestos por el interesado que están incluidos en la solución. Esta cantidad está acotada entre 0 y 1 y la representaremos como un porcentaje. Así, cuanto más importante sea un interesado, mayor debería ser el valor de cobertura proporcionado por las soluciones. Las tablas 3, 4, 5 y 6 muestran, para los distintos grupos de interesados (el índice del grupo indica la importancia del mismo), el porcentaje medio de cobertura (y la desviación típica) de las 10 soluciones con mayor satisfacción en los frentes de Pareto construidos en base a los distintos números tipos de interesados (2, 3, 5 y 6) empleados. Las soluciones diferentes, que no están siempre entre las 10 mejores, se marcan con subrayado. La columna "*sin grupos*" indica el valor de cobertura de la solución en el NRP inicial en el que no se emplea ninguna valoración de la importancia (peso) de los interesados.

Cuando se emplean dos grupos de interesados ( $k = 2$ ) se consigue que las coberturas se diferencien. Así, para el *grupo 2* identificado como más importante (los interesados en este grupo reciben más votos y recomendaciones) las coberturas (en promedio) son mayores que las obtenidas en el NRP inicial y mayores que las del *grupo 1*. Las coberturas para este grupo se sitúan por debajo de las del problema de referencia. La Figura 5a recoge las situaciones comentadas. Sin embargo, este efecto se pierde a partir de tres grupos de interesados y la cobertura de las soluciones por grupos de interesados se entremezcla, tal y cómo puede observarse en la Figura 5b. Una posible explicación puede encontrarse en la división del grupo de interesados con más importancia que se produce conforme crece el número de grupos y a la mezcla de interesados en los grupos con menos importancia (ver Figura 2).

Tabla 3: Diez soluciones con máxima satisfacción en el frente de Pareto obtenido para dos grupos de interesados.

<i>soluciones</i>										<i>cobertura interesados</i>		
<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>sin grupos</i>	<i>grupo 1</i>	<i>grupo 2</i>
1	1	1	1	1	1	1	1	1	1	100.00±0.00	100.00±0.00	100.00±0.00
1	1	1	1	0	1	1	1	1	1	98.88±4.31	98.73±4.66	99.29±3.19
1	1	1	1	1	1	1	1	0	1	95.79±11.56	95.48±11.64	96.67±11.60
<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	96.61±9.51	95.99±10.13	98.33±7.45
1	1	1	1	0	1	1	1	0	1	94.67±11.94	94.21±12.06	95.95±11.82
<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	95.49±10.33	94.73±11.02	97.62±7.95
1	1	1	1	1	1	1	0	1	1	94.05±10.76	93.63±11.44	95.20±8.73
1	1	1	1	0	1	1	0	1	1	92.92±12.01	92.36±12.68	94.49±10.03
1	1	1	1	1	1	1	0	0	1	89.83±14.67	89.11±15.18	91.87±13.30
<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	90.66±13.76	89.63±14.64	93.54±10.72

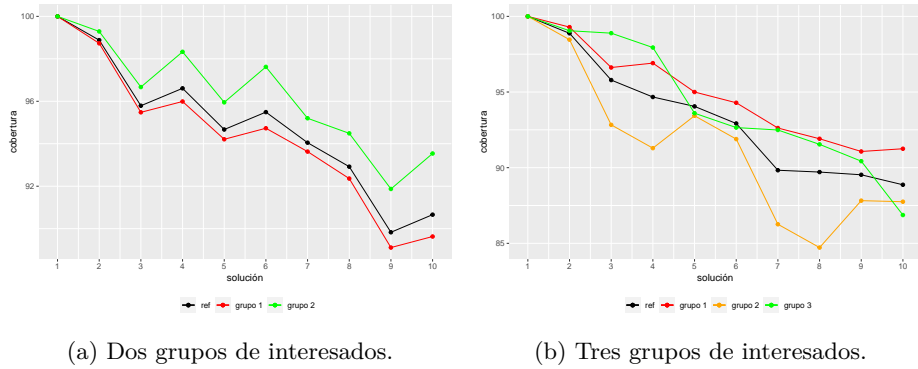


Figura 5: Cobertura promedio para las 10 soluciones con mayor satisfacción en los frentes de Pareto obtenidos para el NRP considerando dos o tres grupos de interesados.

Tabla 4: Diez soluciones con máxima satisfacción en el frente de Pareto obtenido para tres grupos de interesados.

<i>soluciones</i>										<i>cobertura interesados</i>			
<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>sin grupos</i>	<i>grupo 1</i>	<i>grupo 2</i>	<i>grupo 3</i>
1	1	1	1	1	1	1	1	1	1	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00
1	1	1	1	0	1	1	1	1	1	98.88±4.31	99.29±3.78	98.46±5.01	99.05±3.69
1	1	1	1	1	1	1	1	0	1	95.79±11.56	96.62±8.74	92.83±14.97	98.89±4.31
1	1	1	1	0	1	1	1	0	1	94.67±11.94	96.91±9.34	91.29±15.05	97.94±5.46
1	1	1	1	1	1	1	0	1	1	94.05±10.76	95.00±10.07	93.43±12.00	93.60±11.05
1	1	1	1	0	1	1	0	1	1	92.92±12.01	94.29±10.41	91.89±13.80	92.65±11.05
1	1	1	1	1	1	1	0	0	1	89.83±14.67	92.62±12.38	86.26±17.61	92.49±9.78
<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	89.71±15.30	91.91±12.51	84.72±18.27	91.54±11.09
<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>0</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	89.53±14.68	91.07±12.70	87.82±17.12	90.43±12.69
<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	<u>0</u>	<u>1</u>	<u>1</u>	88.87±18.00	91.25±14.36	87.75±21.72	86.87±15.48

## 6. Conclusiones

En este trabajo se comprueba como para asignar los pesos a los interesados en el problema NRP la utilización de algoritmos de agrupamiento es una alternativa viable. De esta forma, la incorporación de nuevos interesados es más sencilla, puesto que existen unas categorías ya definidas, facilitando la obtención de las satisfacciones. También se comprueba como este agrupamiento mantiene la cobertura de los clientes de mayor importancia en el proyecto.

Los resultados obtenidos indican que buscar tipos de interesados y agrupar los interesados en grupos según éstos, tiene un impacto directo sobre la cobertura que ofrecen las soluciones del problema. No obstante, el número de tipos de interesados ha de escogerse cuidadosamente. Un número adecuado de tipos, hace

Tabla 5: Diez soluciones con máxima satisfacción en el frente de Pareto obtenido para cinco grupos de interesados.

<i>soluciones</i>										<i>cobertura interesados</i>					
<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>sin grupos</i>	<i>grupo 1</i>	<i>grupo 2</i>	<i>grupo 3</i>	<i>grupo 4</i>	<i>grupo 5</i>
1	1	1	1	1	1	1	1	1	1	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00
1	1	1	1	0	1	1	1	1	1	98.88±4.31	99.23±39.22	98.30±5.24	99.05±3.69	100.00±0.00	100.00
1	1	1	1	1	1	1	0	1	1	95.79±11.56	97.44±9.06	93.78±13.41	95.56±13.31	100.00±0.00	100.00
1	1	1	1	1	0	1	1	1	1	96.61±9.51	96.54±8.67	95.52±11.38	97.78±8.61	100.00±0.00	100.00
1	1	1	1	0	1	1	1	0	1	94.67±11.94	96.67±9.66	92.08±13.61	94.60±13.48	100.00±0.00	100.00
1	1	1	1	0	0	1	1	1	1	95.49±10.33	95.77±9.22	93.83±12.46	96.83±9.12	100.00±0.00	100.00
1	1	1	1	1	1	0	1	1	1	94.05±10.76	95.58±9.60	91.94±12.74	96.27±8.01	95.00±10.00	80.00
1	1	1	1	0	1	1	0	1	1	92.92±12.01	94.81±10.03	90.24±14.43	95.32±9.97	95.00±10.00	80.00
1	1	1	1	1	1	0	0	1	1	89.83±14.67	93.01±12.28	85.72±16.77	91.83±14.35	95.00±10.00	80.00
1	1	1	1	1	0	1	0	1	1	90.66±13.76	92.12±12.56	87.47±16.13	94.05±10.98	95.00±10.00	80.00

Tabla 6: Diez soluciones con máxima satisfacción en el frente de Pareto obtenido para seis grupos de interesados.

<i>soluciones</i>										<i>cobertura interesados</i>						
<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>	<i>j</i>	<i>sin grupos</i>	<i>grupo 1</i>	<i>grupo 2</i>	<i>grupo 3</i>	<i>grupo 4</i>	<i>grupo 5</i>	<i>grupo 6</i>
1	1	1	1	1	1	1	1	1	1	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00±0.00	100.00
1	1	1	0	1	1	1	1	1	1	98.88±4.31	99.23±39.22	99.23±39.22	96.56±6.85	98.57±4.52	100.00±0.00	100.00
1	1	1	1	1	1	1	0	1	1	95.79±11.56	97.44±9.06	92.82±14.19	94.44±16.67	98.33±5.27	100.00±0.00	100.00
1	1	1	0	1	1	1	0	1	1	94.67±11.94	96.67±9.66	92.05±14.32	91.01±16.78	96.91±6.55	100.00±0.00	100.00
1	1	1	1	0	0	1	1	1	1	95.49±10.33	95.77±9.22	94.62±12.32	94.97±10.41	95.24±11.00	100.00±0.00	100.00
1	1	1	1	1	1	0	1	1	1	94.05±10.76	95.58±9.60	91.35±13.31	98.15±5.56	94.41±9.39	95.00±10.00	80.00
1	1	1	0	1	1	0	1	1	1	92.92±12.01	94.81±10.03	90.58±14.51	94.71±11.53	92.98±11.67	95.00±10.00	80.00
1	1	1	1	1	1	0	0	1	1	89.83±14.67	93.01±12.28	84.17±17.30	92.59±16.90	92.74±9.76	95.00±10.00	80.00
1	1	1	1	0	1	1	0	0	1	89.71±15.30	92.24±12.45	83.40±17.92	89.15±18.56	91.31±11.75	95.00±10.00	80.00
1	1	1	1	0	0	1	0	1	1	89.53±14.68	91.35±12.67	85.96±17.72	93.13±13.71	89.64±13.97	95.00±10.00	80.00

que se obtengan buenas coberturas. Pero aumentar el número de tipos no tiene por qué traducirse en una mejor cobertura. Al menos esto es lo que indica el caso estudiado, siendo esta la principal amenaza, que queda como trabajo futuro, puesto que habría que validar la posible generalización de estos resultados sobre otros proyectos.

## Agradecimientos

Esta investigación ha sido financiada por el Ministerio de Ciencia, Innovación y Universidades en el marco de la red de investigación en ingeniería de software basada en búsquedas (RED2018-102472-T) y del proyecto PID2019-106758GB-C32 (EML-PA); cuenta con el apoyo parcial del grupo de investigación de Ingeniería de Datos, Conocimiento e Software (DKSE) (TIC-181) de la Universidad de Almería.

## Referencias

1. Van den Akker, M., Brinkkemper, S., Diepen, G., Versendaal, J.: Determination of the next release of a software product: an approach using integer linear programming. In: CAiSE short paper proceedings (2005)

2. Babar, M.I., Ghazali, M., Jawawi, D.N.A., Zaheer, K.B.: Stakemeter: Value-based stakeholder identification and quantification framework for value-based software systems. *PLOS ONE* **10**(3), 1–33 (03 2015)
3. Babar, M.I., Ghazali, M., Jawawi, D.N.: A bi-metric and fuzzy c-means based intelligent stakeholder quantification system for value-based software. In: *SoMeT*. pp. 295–309 (2014)
4. Bagnall, A., Rayward-Smith, V., Whittley, I.: The next release problem. *Information and Software Technology* **43**(14), 883 – 890 (2001)
5. Ballejos, L.C., Montagna, J.M.: Method for stakeholder identification in inter-organizational environments. *Requirements engineering* **13**(4), 281–297 (2008)
6. Estivill-Castro, V.: Why so many clustering algorithms: A position paper. *SIGKDD Explor. Newsl.* **4**(1), 65–75 (2002)
7. Finkelstein, A., Harman, M., Mansouri, S.A., Ren, J., Zhang, Y.: A search based approach to fairness analysis in requirement assignments to aid negotiation, mediation and decision making. *Requirements engineering* **14**(4), 231–245 (2009)
8. Freeman, R.E.: *Strategic management: A stakeholder approach*. Cambridge university press (2010)
9. Freeman, R.E., Reed, D.L.: Stockholders and stakeholders: A new perspective on corporate governance. *California management review* **25**(3), 88–106 (1983)
10. Glinz, M., Wieringa, R.J.: Guest editors’ introduction: Stakeholders in requirements engineering. *IEEE software* **24**(2), 18–20 (2007)
11. Hartigan, J., Wong, M.: Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society, Series C (Applied Statistics)* **28**(1), 100 – 108 (1979)
12. Hujainah, F., Bakar, R.B.A., Abdulgaber, M.A.: Stakeqp: A semi-automated stakeholder quantification and prioritisation technique for requirement selection in software system projects. *Decision Support Systems* **121**, 94–108 (2019). <https://doi.org/https://doi.org/10.1016/j.dss.2019.04.009>
13. Hujainah, F., Bakar, R.B.A., Al-Haimi, B., Abdulgaber, M.A.: Stakeholder quantification and prioritisation research: A systematic literature review. *Information and Software Technology* **102**, 85–99 (2018)
14. Hujainah, F., Binti Abu Bakar, R., Nasser, A.B., Al-haimi, B., Zamli, K.Z.: Srptackle: A semi-automated requirements prioritisation technique for scalable requirements of software system projects. *Information and Software Technology* **131**, 106501 (2021)
15. Jung, H.W.: Optimizing value and cost in requirements analysis. *IEEE software* **15**(4), 74–78 (1998)
16. Lim, S.L., Finkelstein, A.: Stakerare: using social networks and collaborative filtering for large-scale requirements elicitation. *IEEE transactions on software engineering* **38**(3), 707–735 (2011)
17. Lim, S.L., Harman, M., Susi, A.: Using genetic algorithms to search for key stakeholders in large-scale software projects. In: *Aligning Enterprise, System, and Software Architectures*, pp. 118–134. IGI Global (2013)
18. Mitchell, R.K., Agle, B.R., Wood, D.J.: Toward a theory of stakeholder identification and salience: Defining the principle of who and what really counts. *The Academy of Management Review* **22**(4), 853–886 (1997)
19. Nayebi, M., Ruhe, G.: Asymmetric release planning: compromising satisfaction against dissatisfaction. *IEEE Transactions on Software Engineering* **45**(9), 839–857 (2018)

20. Pacheco, C., Garcia, I.: A systematic literature review of stakeholder identification methods in requirements elicitation. *Journal of Systems and Software* **85**(9), 2171–2181 (2012)
21. Ramirez, A., Romero, J.R., Ventura, S.: A survey of many-objective optimisation in search-based software engineering. *Journal of Systems and Software* **149**, 382–395 (2019)
22. Ruhe, G.: *Product release planning: methods, tools and applications*. CRC Press (2010)
23. del Sagrado, J., del Águila, I.M.: Assisted requirements selection by clustering. *Requirements Engineering* **26**(2), 167–184 (2021)
24. Sarmah, S., Datta, D.: Product priority problem: A multi-objective optimization approach for product development based on customers’ priority. *Modeling, Simulation and Optimization: Proceedings of CoMSO 2020* p. 101 (2020)
25. Zhang, Y., Harman, M., Finkelstein, A., Mansouri, S.A.: Comparing the performance of metaheuristics for the analysis of multi-stakeholder tradeoffs in requirements optimisation. *Information and software technology* **53**(7), 761–773 (2011)
26. Zhang, Y., Harman, M., Mansouri, S.A.: The multi-objective next release problem. In: *Proceedings of the 9th annual conference on Genetic and evolutionary computation*. pp. 1129–1137 (2007)