

INFLUENCIA DE FACTORES EN EL RATIO DE PRODUCTIVIDAD DE PROYECTOS DE SOFTWARE DEL BANCO DE DATOS ISBSG

Marta Fernández-Diego

Sanae Elmouaden

Víctor Núñez-Monsálvez

María-Dolores Sanz-Berzosa

José-María Torralba-Martínez

Universitat Politècnica de València

(Departamento de Organización de Empresas)

Abstract

This paper proposes the study of factors influencing software projects productivity rate using Bayesian networks. There are many estimation methods, however Bayesian networks are being increasingly used in software engineering. Although expert knowledge can be included on the base of their estimates, this study seeks a preliminary estimate of productivity rate based only on the information contained in ISBSG (International Software Benchmarking Standards Group) dataset. This dataset provides a reliable and relevant sample of over 5000 software projects. The results showed a greater influence on software projects productivity rate of the team size and development platform type, compared to other factors such as the use of methodology, language type and development type.

Keywords: *productivity; software projects; bayesian networks; ISBSG dataset*

Resumen

Este artículo propone el estudio de factores influyentes en el ratio de productividad de proyectos de software utilizando redes bayesianas. Métodos de estimación hay muchos y las redes bayesianas están siendo cada vez más utilizadas dentro de la ingeniería del software. Aunque permiten incluir el conocimiento experto como base en sus estimaciones, este estudio pretende una estimación previa del ratio de productividad basándose únicamente en la información contenida en el banco de datos ISBSG (International Software Benchmarking Standards Group). Dicho banco de datos presenta una muestra fiable y relevante de más de 5000 proyectos de software. Como resultado destaca fundamentalmente la influencia del tamaño medio del equipo y del tipo de plataforma de desarrollo en el ratio de productividad de los proyectos, frente a otros factores como el uso de metodología, el tipo de lenguaje y el tipo de desarrollo.

Palabras clave: *productividad; proyectos de software; redes bayesianas; banco de datos ISBSG*

1. Introducción

En la estimación de la productividad del desarrollo de software se vienen utilizando diversidad de herramientas estadísticas, como la regresión lineal empleada en un trabajo previo (Fernández-Diego, Maciel, Marcelo-Llácer, & Torralba-Martínez, 2010), potenciándose en la última década la utilización de las redes bayesianas.

Frente a las técnicas actuales, existen otros trabajos previos realizados sobre este tema en España. En la tesis (Villanueva-Balsera, 2005) y posteriormente en el artículo (Villanueva-Balsera, Ortega-Fernandez, Rodriguez-Montequin, & Concepcion-Suarez, 2009) se analizan la viabilidad y las ventajas del desarrollo de un sistema basado en técnicas de inteligencia artificial, capaz de seleccionar las variables que afectan a la duración del proyecto y al esfuerzo necesario para realizarlo a partir de un conjunto de datos históricos.

En (Jiang et al., 2007) se realiza un estudio de la productividad del desarrollo software sobre el periodo 1995-2005, basado en la versión 10 de ISBSG, determinando como factores que influyen el tamaño medio del equipo de proyecto y el uso desequilibrado de lenguajes 3GL menos productivos.

En (Radliński, Fenton, Neil, & Marquez, 2007) se presenta un modelo de productividad utilizando redes bayesianas. Dicho modelo tiene como entradas estimaciones a priori del ratio de productividad y el ratio de defectos. En (Radliński, Fenton, & Marquez, 2008), que tomamos como referencia principal en esta comunicación, se construye precisamente un modelo basado en ISBSG versión 9 para proporcionar estos ratios a priori para aquellos casos en que no se dispone de información sobre proyectos anteriores.

El objetivo de esta comunicación es presentar un modelo de red bayesiana que permita una estimación previa del ratio de productividad basándose únicamente en la información contenida en el banco de datos ISBSG.

El plan de la comunicación es el siguiente: En el apartado 2 siguiente se realiza una descripción de la metodología utilizada; en el 3, la topología del modelo y los resultados obtenidos; y en el 4, para acabar, se presentan las conclusiones.

2. Metodología

2.1 Descripción del banco de datos

Para llevar a cabo una investigación empírica dentro de la ingeniería del software, es importante contar con un banco de datos fiable y que esté formado por una muestra relevante para el estudio. Aquí, se utiliza una fuente disponible, el banco de datos ISBSG en su versión 11 (ISBSG, 2009) que se facilita a las Universidades, que incluye más de 5000 proyectos de software de multitud de países y de muy diversa índole. Un trabajo previo (Maciel, Fernández-Diego, Sanz-Berzosa, & Torralba-Martínez, 2010) presenta la evolución producida en el banco de datos ISBSG, basada en la nueva versión 11, referida a algunas de las variables relacionadas con la productividad.

2.2 Criterios de selección de los proyectos del banco de datos

Los criterios de selección utilizados vienen determinados por la necesidad de utilizar datos de calidad, y de coherencia a la hora de comparar el tamaño en puntos de función y el esfuerzo en horas/persona (Lokan & Mendes, 2009):

1. Calidad de los datos

El primer criterio impuesto es un criterio de calidad de los datos. En el banco de datos aparecen dos variables relativas a la calidad de los datos:

- La calidad de los datos en general del proyecto, que está categorizada en cuatro niveles decrecientes de calidad (A, B, C y D). Generalmente los artículos revisados que tienen por objeto este banco de datos dan por aceptables tanto los proyectos de calidad A como calidad B, desechando los proyectos de calidad C y D. Está codificada como V2;
- La calidad de los datos de medición del tamaño funcional del proyecto. La categorización de esta variable es la misma que lo planteado para la calidad en general. Sin embargo esta variable hace referencia únicamente a la calidad de la medición del tamaño funcional del proyecto. Está codificada como V3.

En un trabajo previo (Fernández-Diego, Martínez-Gómez, & Torralba-Martínez, 2010), se realiza una revisión de los criterios adoptados por los investigadores en cuanto a calidad se refiere, de la que se concluye que no hay un aparente consenso. Sin embargo, se demuestra que este criterio influye de manera estadísticamente significativa en los modelos de estimación de esfuerzo. Como conclusión, se trabajará en este estudio con calidad general de los datos A o B y además se restringirá también la calidad de la medición del tamaño funcional a A o B.

2. Ratio de productividad

La versión 11 cuenta ya con la variable "Normalised Level 1 Productivity Delivery Rate". Este ratio en horas por tamaño funcional está calculado como el cociente de la variable "Normalised Level 1 Work Effort" y el tamaño funcional (Unadjusted Function Points). Corresponde en definitiva a la inversa del ratio de productividad. El uso del esfuerzo normalizado para el equipo de desarrollo únicamente (Level 1) y los puntos de función no ajustados proporciona ratios más comparables, por lo que es la medida recomendada por ISBSG.

El criterio que se impone es que dicha variable (V12) no esté vacía.

3. Coherencia en la comparación de la medida funcional

Estamos trabajando con puntos de función y no con cualquier otro tipo de medida del tamaño, por lo que es necesario que el método de medida sea el mismo para poder hacer comparaciones. Como método de medida del tamaño en puntos de función se ha elegido IFPUG (International Function Point Users Group) por ser el más representativo. Por otro lado, dentro de IFPUG existen varias versiones. Parece que a partir de la versión 4 (1994) y posteriores es posible realizar una comparación de los tamaños medidos según las distintas versiones.

Tabla 1: Criterios de selección de proyectos

Criterio	Filtro	NP 1	NP 2	NP 3
Alta calidad de los datos en general	V2="A" V2="B"	4744	308	308
Alta calidad de los datos del tamaño funcional	V3="A" V3="B"	3184	1560	1868
"Normalised Level 1 Productivity Delivery Rate" calculada	MISSING (V12)~=1	2710	474	2342
Versión IFPUG 4.0 o posterior	CHAR.INDEX(V66,'IFPUG 4')~=0 (V66="IFPUG" & V4>"1994")	1612	1098	3440

La tabla 1 resume los criterios de selección empleados utilizando el lenguaje de SPSS versión 16. El significado de las columnas finales es el siguiente: NP 1: N° de proyectos que cumplen la condición de la fila y las anteriores; NP 2: N° de proyectos que excluye la condición de la fila; NP 3: N° de proyectos que excluye la condición de la fila y las anteriores. Finalmente contamos con un total de 1612 proyectos para el siguiente análisis de los 5052 proyectos con que cuenta la versión 11 del repositorio.

2.3 Análisis estadístico

El tratamiento estadístico de los datos se ha realizado utilizando SPSS versión 16.

2.3.1 Variable dependiente

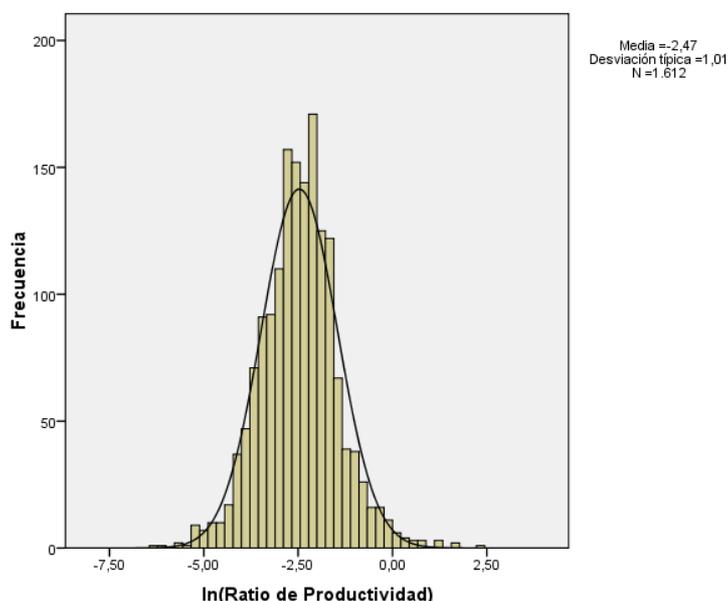
El primer paso a realizar consiste en calcular la inversa de la variable “Normalised Level 1 Productivity Delivery Rate” proporcionada en el banco de datos, que como se ha indicado correspondería entonces al ratio de productividad.

Tras una transformación logarítmica sobre la variable en cuestión, la figura 1 muestra una distribución aproximadamente normal:

$$\ln(\text{Ratio de Productividad}) \sim N(-2,47; 1,02) \quad (1)$$

Esta transformación podrá permitir más adelante una discretización más apropiada de la misma.

Figura 1: Distribución de la variable $\ln(\text{Ratio de Productividad})$



2.3.2 Posibles factores influyentes

La lista inicial de posibles factores se basa en la identificación presentada en (Jiang, Naudé, & Comstock, 2007). En dicho trabajo se realiza una investigación de la variación de la productividad en proyectos de software del banco de datos ISBSG. La tabla 2 presenta esa selección previa de variables de ISBSG que serán utilizadas posteriormente en el modelo si estadísticamente se verifican una serie de condiciones.

Tabla 2: Listado de los posibles factores influyentes

Identificador	Factor	Tipo	Estados
V101	Tamaño medio del equipo	Continua	[1..49]
V70	Tipo de lenguaje	Nominal	2GL, 3GL, 4GL, ApG
V32	Tipo de desarrollo	Nominal	new development, enhancement, re-development
V69	Plataforma de desarrollo	Nominal	PC, MidRange, MainFrame, Multi platform
V91	Uso de metodología	Nominal	Yes, No, Don't know
V90	Uso de herramienta CASE	Nominal	Yes, No, Don't know
V92	Adquisición de la metodología	Nominal	Combined Developed/Purchased, Developed In-house, Purchased, Traditional

Con respecto a los siete factores propuestos, se ha preferido no entrar en las técnicas de desarrollo sustituyendo esta variable por la que indica simplemente el uso o no de metodología, como hacen en (Radliński et al., 2008).

El primer paso consiste en identificar las relaciones entre los posibles factores y la variable dependiente. Para ello, puesto que la variable dependiente es una variable numérica continua, si los factores son también variables numéricas continuas se utiliza el coeficiente de Spearman. En (Maxwell, 2002) se explica porqué se prefiere al de Pearson. En el caso de que los factores sean nominales, se recurre a la prueba de Kruskal-Wallis. Frente a la prueba de Mann-Whitney, la de Kruskal-Wallis se utiliza para un único factor de varios niveles. Según este primer análisis, se puede confirmar que todos los factores considerados en la tabla 2 influyen en la variable dependiente.

Tabla 3: Relación entre el factor "Tamaño medio del equipo" y los factores nominales

Factores nominales	Factor continuo					
	Tipo de lenguaje	Tipo de desarrollo	Plataforma de desarrollo	Uso de metodología	Uso de herramienta CASE	Adquisición de la metodología
Tamaño medio del equipo	-	+	+	+	+	-

Note: Factor relacionado con "Tamaño medio del equipo": +; Factor no relacionado con "Tamaño medio del equipo": -.

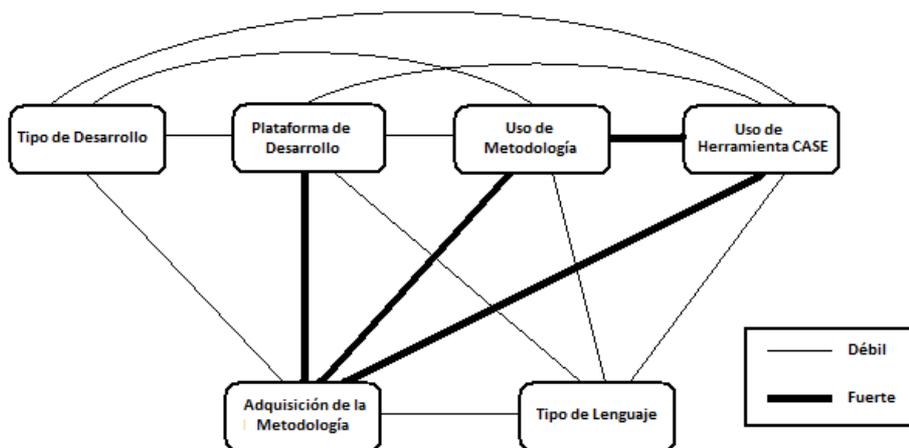
En segundo lugar se identifica el grado de correlación o asociación entre los distintos factores. Aunque no es necesario, esto permite simplificar en gran medida el modelo de red bayesiana convirtiéndola en un Clasificador Naïve Bayes (NBC) (Radliński et al., 2008). Para

el factor continuo con el resto de factores nominales, se recurrirá de nuevo a la prueba de Kruskal-Wallis.

En cuanto a la relación entre el factor continuo “Tamaño medio del equipo” y el resto de factores nominales, la tabla 3 presenta un resumen en que queda manifiesta la independencia con respecto a los factores “Tipo de lenguaje” y “Adquisición de la metodología”.

Para identificar la posible asociación entre dos factores nominales se emplea la prueba de Cramer. Los resultados de la relación entre los factores nominales se presentan en la figura 1.

Figura 1: Relación entre los factores nominales



En la figura 1 se aprecia no sólo la existencia o no de asociación, sino también la fortaleza o no de dicha asociación. A la vista de los resultados, la primera medida consiste en despreciar las asociaciones débiles, esto es con un valor de medida por debajo de 0.3. A continuación, se va a proceder a romper el triángulo de dependencias fuertes entre los factores “Uso de metodología”, “Uso de herramienta CASE” y “Adquisición de la metodología”. Dado que el empleo de herramientas CASE, a pesar de influir en el ratio de productividad presentaba un valor de significación límite (0.05) para el rechazo de la hipótesis nula, queda descartado de aquí en adelante. Por otro lado, entre los otros dos factores es preferible descartar el factor “Adquisición de la metodología” que tiene a su vez una dependencia fuerte con el factor “Plataforma de desarrollo”.

Finalmente queda pendiente la relación del factor “Tamaño medio del equipo” con los factores nominales “Tipo de desarrollo”, “Plataforma de desarrollo” y “Uso de metodología”. Tras la categorización de dicha variable atendiendo al criterio de repartición homogénea de casos, dando lugar a 4 intervalos (entre 1 y 2 personas, entre 3 y 5 personas, entre 6 y 9 personas, y 10 o más personas), queda manifiesta una asociación débil con todos ellos.

Por consiguiente, una vez finalizado este análisis estadístico puede asumirse la hipótesis de independencia entre los factores.

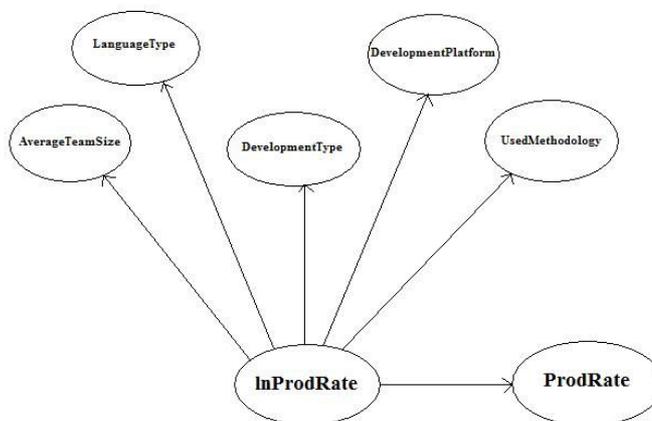
3. El modelo

La herramienta empleada para el modelado es AgenaRisk (AgenaRisk, 2008). AgenaRisk es una herramienta potente pero intuitiva para modelar, analizar y predecir en ambientes de incertidumbre. Soporta el razonamiento tanto diagnóstico como predictivo gracias al uso de redes bayesianas.

3.1 Estructura del modelo

El modelo propuesto en la figura 2 sigue la topología de un Clasificador Naïve Bayes (NBC). El fundamento principal es la suposición de que todos los factores son independientes conocido el valor de la variable clase, es decir la variable dependiente (Hernández Orallo, José, Ramírez Quintana, María José, & Ferri Ramírez, César, 2004). Esta hipótesis da lugar a un grafo en el que existe un único nodo raíz, la clase, y en la que todos los factores son nodos hoja que tienen como único padre a la variable clase.

Figura 2: Topología del modelo



Al ser la variable dependiente una variable continua ha sido discretizada (Tuya, Román, & Cosín, 2007), aunque AgenaRisk permite el uso de modelos paramétricos para abordar el problema de la especificación de las tablas de probabilidad condicional. La variable transformada del ratio de productividad ha sido redondeada al entero más cercano. Como el valor mínimo que toma la variable es -6,43 y el valor máximo es 2,30, han resultado 9 intervalos utilizando un mismo rango de 1.

3.2 Resultados

Al introducir evidencias u observaciones en el modelo, se actualizan las probabilidades del resto del sistema. A continuación se muestra cómo afecta en la mediana del ratio de productividad la instanciación de cada uno de los factores por separado (Tablas 4 a 8).

Tabla 4: p(Ratio de Productividad | Tamaño medio del equipo)

	Tamaño medio del equipo			
	1-2	3-5	6-9	10+
Mediana del ratio de productividad	0,171	0,106	0,075	0,05

Los resultados obtenidos sobre el factor “Tamaño medio del equipo” confirman lo ya sugerido en la literatura, es decir una asociación negativa entre ambos parámetros (Maxwell,

Van Wassenhove, & Dutta, 1996). En (Pendharkar & Rodger, 2007) se realiza un estudio sobre el impacto del tamaño del equipo de desarrollo en el esfuerzo de desarrollo.

Tabla5: p(Ratio de Productividad | Tipo de lenguaje)

	Tipo de lenguaje			
	2GL	3GL	4GL	ApG
Mediana del ratio de productividad	0,068	0,074	0,13	0,098

En relación al tipo de lenguaje los resultados obtenidos tienen también sentido. En principio los lenguajes de cuarta generación están diseñados para reducir el esfuerzo del equipo de desarrollo, por lo que serían más productivos que los de tercera generación, que a su vez serían más productivos que los de segunda generación. El dato curioso se observa en cuanto a los generadores de aplicaciones, supuestamente los más productivos de todos ellos y confirmado en (Radliński et al., 2008).

Tabla 6: p(Ratio de Productividad | Tipo de desarrollo)

	Tipo de desarrollo		
	Enhancement	New Development	Re-development
Mediana del ratio de productividad	0,078	0,122	0,108

Pese a lo que puede pensarse de manera más intuitiva, el desarrollo de mantenimiento no es más fácil y por consiguiente no mejora necesariamente la productividad, puesto que consume muchos recursos del equipo (Lientz, Swanson, & Tompkins, 1978). Esto lo confirman los resultados de la tabla 6 y a su vez en (Radliński et al., 2008). Sin embargo en (Kitchenham, Lawrence Pfleeger, McColl, & Eagan, 2002) y confirmado en (Premraj, Kitchenham, Shepperd, & Forselius, 2005), se concluye que no existe una diferencia significativa entre proyectos de nuevo desarrollo y de mantenimiento. Con respecto a los proyectos de re-desarrollo no se puede concluir gran cosa, al ser poco significativos sobre el conjunto de proyectos.

Tabla 7: p(Ratio de Productividad | Plataforma de desarrollo)

	Plataforma de desarrollo			
	MF	MR	Multi	PC
Mediana del ratio de productividad	0,069	0,086	0,108	0,156

La tabla 7 presenta la influencia del tipo de plataforma de desarrollo en el ratio de productividad. El impacto de cada una de las etiquetas es acorde al resultado obtenido en (Radliński et al., 2008) (Jiang et al., 2007).

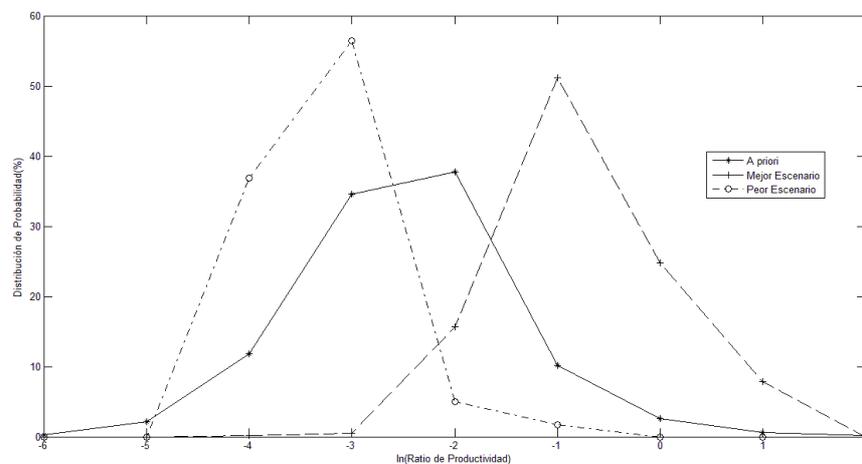
Por último el uso de metodología o no en el proyecto tiene un efecto aparentemente contradictorio. El empleo de una metodología concreta no mejora la productividad, sino lo contrario, puesto que se transfiere parte del esfuerzo dedicado al producto a actividades asociadas con el proceso, como la elaboración de documentación. Al mismo resultado se llegó en (Radliński et al., 2008).

Tabla 8: $p(\text{Ratio de Productividad} \mid \text{Uso de metodología})$

	Uso de metodología	
	No	Yes
Mediana del ratio de productividad	0,158	0,082

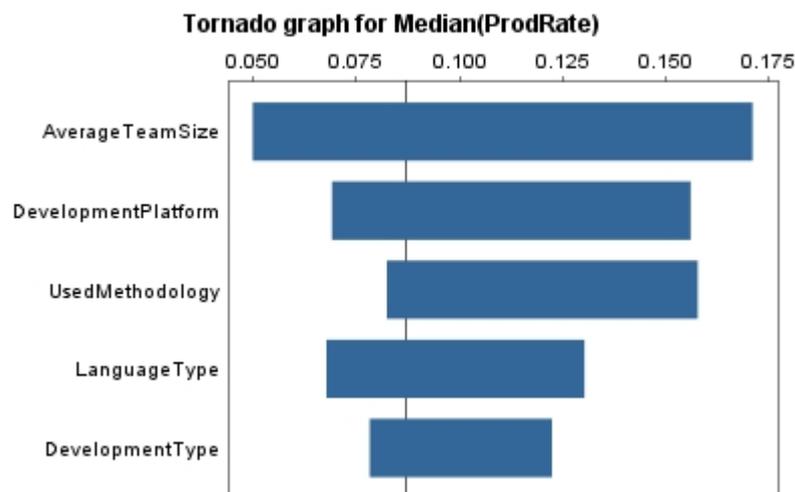
A continuación se presenta el escenario más y menos favorable al ratio de productividad combinando las evidencias más y menos favorables de todos los factores. El resultado aparece en la figura 3.

Figura 3: Escenarios más y menos favorable



Por último, para determinar en qué medida cada uno de los factores tiene influencia en el ratio de productividad, la herramienta permite realizar un análisis de sensibilidad. Los resultados se muestran en la figura 4.

Figura 4: Análisis de sensibilidad



Como puede observarse, el factor más influyente en el ratio de productividad es el tamaño medio del equipo, seguido del tipo de plataforma de desarrollo y del uso o no de

metodología. Los factores menos influyentes son el tipo de lenguaje y el tipo de desarrollo. Sin embargo, en (Jiang et al., 2007), sobre los mismos factores estudiados concluyen que el tamaño medio del equipo y el tipo de lenguaje son los más influyentes en la productividad. La posición segunda del tipo de plataforma de desarrollo respalda el trabajo que pone en evidencia su efecto significativo en el esfuerzo de desarrollo software (Subramanian, Pendharkar, & Wallace, 2006).

Por otro lado y comparando los parámetros en común, la sensibilidad de éstos en el ratio de productividad resulta menor en nuestro caso que en (Radliński et al., 2008).

4. Conclusiones

Partiendo de un conjunto de factores reconocidos como influyentes en la productividad de los proyectos software, un análisis estadístico ha permitido confirmar dicha influencia. A partir de ellos se ha estudiado la posibilidad de su inserción en un Clasificador Naïve Bayes (NBC) asegurando la independencia entre ellos.

Como resultado destaca fundamentalmente la influencia del tamaño medio del equipo y del tipo de plataforma de desarrollo en el ratio de productividad de los proyectos, frente a otros factores como el uso de metodología, el tipo de lenguaje y el tipo de desarrollo.

Una de las limitaciones de este trabajo es que hay otros muchos factores que influyen en la productividad y que no han sido contemplados simplemente por no venir recogidos en ISBSG, como pueden ser las cuestiones más relacionadas con el equipo de desarrollo, su formación, habilidades, grado de motivación, ...

Por otro lado, una línea de futuro sería la incorporación de otras variables, como el número de defectos encontrados en el software, en el modelo de productividad.

5. Agradecimientos

Los autores agradecen a los revisores las oportunas recomendaciones que realizaron sobre la propuesta de comunicación.

6. Referencias

AgenaRisk. (2008). AgenaRisk software v5.0. *Bayesian Network and Simulation*

Software for Risk Analysis and Decision Support. Retrieved from

<http://www.agenarisk.com/>

Fernández-Diego, M., Maciel, J., Marcelo-Llácer, D., & Torralba-Martínez, J. (2010).

Software projects size and economies of scale in the ISBSG dataset. In *XIV*

International Congress on Project Engineering. Presented at the XIV International Congress on Project Engineering, Madrid, Spain.

Fernández-Diego, M., Martínez-Gómez, M., & Torralba-Martínez, J. (2010).

Sensitivity of results to different data quality meta-data criteria in the sample selection of projects from the ISBSG dataset. In *Proceedings of the 6th International Conference on Predictive Models in Software Engineering, PROMISE '10* (pp. 13:1–13:9). New York, NY, USA: ACM.

Hernández Orallo, José, Ramírez Quintana, María José, & Ferri Ramírez, César.

(2004). *Introducción a la minería de datos*. Pearson. Retrieved March 30, 2011, from <http://www.agapea.com/libros/Introduccion-a-la-mineria-de-datos-isbn-8420540919-i.htm>

ISBSG. (2009). ISBSG dataset Release 11. *International Software Benchmarking Standards Group*. Retrieved from <http://www.isbsg.org/>

Jiang, Z., Naudé, P., & Comstock, C. (2007). An investigation on the variation of software development productivity. *International Journal of Computer and Information Science and Engineering*, 1(2), 72-81.

Kitchenham, B., Lawrence Pfleeger, S., McColl, B., & Eagan, S. (2002). An empirical study of maintenance and development estimation accuracy. *Journal of Systems and Software*, 64(1), 57-77.

Lientz, B. P., Swanson, E. B., & Tompkins, G. E. (1978, June). Characteristics of application software maintenance. *Communications of the ACM*, 21, 466–471.

Lokan, C., & Mendes, E. (2009). Investigating the use of chronological split for software effort estimation. *IET Software*, 3(5), 422-434.

Maciel, J., Fernández-Diego, M., Sanz-Berzosa, M., & Torralba-Martínez, J. (2010).

- The recent evolution of the ISBSG (International Software Benchmarking Standards Group) software projects dataset. In *XIV International Congress on Project Engineering*. Presented at the XIV International Congress on Project Engineering, Madrid, Spain.
- Maxwell, K. D., Van Wassenhove, L., & Dutta, S. (1996). Software Development Productivity of European Space, Military, and Industrial Applications. *IEEE Transactions on Software Engineering*, 22, 706–718.
- Maxwell, K. D. (2002). *Applied Statistics for Software Managers*. Prentice Hall.
- Pendharkar, P. C., & Rodger, J. A. (2007). An empirical study of the impact of team size on software development effort. *Information Technology and Management*, 8, 253–262.
- Premraj, R., Kitchenham, B., Shepperd, M., & Forselius, P. (2005). An empirical analysis of software productivity over time. *11th IEEE International Symposium On Software Metrics (METRICS 2005)*, IEEE Computer Society, 37.
- Radliński, Ł., Fenton, N., & Marquez, D. (2008). Estimating Productivity and Defect Rates Based on Environmental Factors. In *Information Systems Architecture and Technology: Models of the Organisation's Risk Management* (pp. 103--113). Wrocław, Poland: Oficyna Wydawnicza Politechniki Wrocławskiej.
- Radliński, Ł., Fenton, N., Neil, M., & Marquez, D. (2007). Improved decision-making for software managers using Bayesian networks (pp. 13–19). Anaheim, CA, USA: ACTA Press. Retrieved March 12, 2011, from <http://portal.acm.org/citation.cfm?id=1647636.1647640>
- Subramanian, G. H., Pendharkar, P. C., & Wallace, M. (2006). An empirical study of

the effect of complexity, platform, and program type on software development effort of business applications. *Empirical Software Engineering*, 11(4), 541-553.

Tuya, J., Román, I. R., & Cosín, J. D. (2007). *Técnicas cuantitativas para la gestión en la ingeniería del software*. Netbiblo.

Villanueva-Balsera, J. (2005, October 7). *Estimación de costes y plazos en proyectos de sistemas de información* (<http://purl.org/dc/dcmitype/Text>). Universidad de Oviedo. Retrieved May 1, 2011, from <http://dialnet.unirioja.es/servlet/tesis?codigo=19735>

Villanueva-Balsera, J., Ortega-Fernandez, F., Rodriguez-Montequin, V., & Concepcion-Suarez, R. (2009). Effort Estimation in Information Systems Projects using Data Mining Techniques. In *Proceedings of the 13th WSEAS International Conference on COMPUTERS* (pp. 652-657). Rhodes Island, Greece: WORLD SCIENTIFIC AND ENGINEERING ACAD AND SOC.

Correspondencia (Para más información contacte con):

Marta Fernández-Diego
Universitat Politècnica de València
Departamento de Organización de Empresas
Camino de Vera s/n, 46022 Valencia, Spain
Phone : +34 96 387 76 85
E-mail : marferdi@omp.upv.es