

MODELO DINÁMICO PARA LA ESTIMACIÓN TEMPRANA DE ESFUERZO EN PROYECTOS DE DESARROLLO DE SOFTWARE

Ana Lucía Pérez*, Liliana González**, Astrid Duque***, Felipe Millane**** y Germán Ospina*****

Recibido: 23/05/2006
Aceptado: 25/08/2006

RESUMEN

Existen modelos para estimación del esfuerzo requerido en proyectos software que presentan limitaciones cuando se utilizan en etapas tempranas del ciclo de vida de desarrollo. En este artículo se presenta una revisión de los modelos existentes y se describe un nuevo modelo capaz de estimar esfuerzo en las primeras etapas del desarrollo, cuyas entradas dependen de históricos de proyectos realizados y de la experiencia de proyectos similares.

El modelo expuesto en este artículo fue resultado de un proyecto de investigación aplicada realizado entre la Universidad de Antioquia y Orbitel S. A., con el objetivo de estimar el esfuerzo requerido por los analistas de la Gerencia de Informática para la creación y operación de soluciones. Considerando los históricos disponibles en Orbitel, nuestro modelo entrega estimaciones precisas desde el punto de vista estadístico. Adicionalmente, el modelo propuesto puede ser simulado con una herramienta orientada a la Web.

Palabras clave: estimación temprana, modelo de estimación de esfuerzo en proyectos de software.

ABSTRACT

There are models for estimation of the effort required in software projects that present limitations when they are used in early stages of the service life of development. In this article a revision of the existing models is presented and a new model able of estimating effort in the first stages of the development is described, whose entrances depend on both historical of projects made and on the experience of similar projects.

The model exposed in this article was the result of an applied research project done between Universidad de Antioquia and Orbitel S.A., with the objective of considering the effort required by the analysts of the Management of Computer science for the creation and operation of solutions. Considering the historical available in Orbitel, our model gives precise estimations from the statistical point of view. Additionally, the proposed model can be simulated with a tool oriented to the Web.

Key words: Early estimation, model of effort estimation in software project.

* MSc. Profesora de tiempo completo. Universidad de Antioquia. Teléfono: 234 88 43. e-mail: alperez@udea.edu.co

** Ingeniera de Sistemas. Estudiante de la Maestría en Ingeniería con énfasis en informática de la Universidad de Antioquia. Teléfono: 2105508. e-mail: lilianagonpa@yahoo.com

*** Ingeniera de Sistemas Universidad de Antioquia. Teléfono: 2105532. e-mail: asdur02@gmail.com

**** Ingeniero de Sistemas Universidad de Antioquia. Teléfono: 2105532. e-mail: jfelipe885@hotmail.com

***** Estudiante de Ingeniería de Sistemas Universidad de Antioquia. Teléfono: 2105532. e-mail: germanospina@hotmail.com

INTRODUCCIÓN

La inversión total en desarrollo y mantenimiento de software se ha incrementado rápidamente en los últimos años y se estima que tiene un costo de más de US\$200 millones por año (Arora et al., 2005). Claramente, una parte importante para el control de este costo es la capacidad que tienen las firmas de anticipar de manera precisa el esfuerzo requerido para el proceso de desarrollo de software. El objetivo de este artículo es describir un modelo que, utilizando históricos y experiencia disponible de proyectos similares, genera estimaciones tempranas del esfuerzo necesario para crear soluciones de tecnología de información (TI).

Estimar el esfuerzo de las soluciones que involucran TI implica un alto grado de incertidumbre debido a que su comportamiento en el tiempo es altamente sensible, entre otras, a variables como la habilidad y cantidad del recurso humano disponible en la firma, al cambio tecnológico y a las decisiones tomadas en niveles estratégicos de la organización para atender las condiciones de la dinámica misma del mercado. Estas características particulares y complejas de industrias dedicadas o dependientes de las TI, dificultan la precisión de la labor de planificación inicial, realizada por los gestores de proyectos.

En particular, la habilidad y la cantidad de recurso humano, expresada en términos de esfuerzo y con mayor precisión como la fuerza de trabajo requerida para el desarrollo, medida en meses-hombre, días-hombre, y en general, unidad de tiempo-hombre (Zhenyou 2004) requiere especial atención durante la planificación del proyecto, motivo por el cual, debe estimarse lo más temprano posible, idealmente en cuanto se tiene la especificación de requisitos.

El problema que motiva la investigación realizada con la Gerencia Informática de Orbitel es que los analistas requieren estimar el esfuerzo para la gestión de cada uno de los proyectos a realizar antes de tener una completa especificación de

requisitos, y la utilización de los modelos empíricos anteriormente utilizados generaba imprecisiones en la planificación de la capacidad para atención de fallas, creación de soluciones, operación de los procesos del negocio.

TÉCNICAS PARA EL CÁLCULO DEL ESFUERZO DE ACUERDO CON EL TAMAÑO DEL PRODUCTO

El principal factor que influye en el cálculo del esfuerzo es el tamaño del producto a desarrollar; es por esto que se han propuesto métodos diferentes del juicio experto que buscan disminuir la incertidumbre en la estimación del tamaño (Jorgensen et al., 2001). Entre las técnicas más estructuradas se encuentran: puntos función, puntos característica, puntos de casos de uso, entre otras. Cada una de estas técnicas tiene fórmulas para calcular el esfuerzo de acuerdo con el tamaño del producto a construir.

La técnica de Puntos Función o Function Points (Albrecht, 1979; Albrecht et al., 1983) proporciona una unidad de medida para la funcionalidad de los sistemas software determinando sus componentes principales: entradas, salidas, consultas o peticiones interactivas (cuando el usuario hace una petición al sistema y éste devuelve una respuesta), archivos lógicos internos (archivos maestros) y archivos lógicos externos (interfaces con otras aplicaciones), y luego asociando estos componentes a características generales de un sistema (eficiencia, reusabilidad, facilidad de operación y mantenimiento entre otras) (Sánchez, 1999).

La técnica ya descrita fue pensada para medir el tamaño funcional de sistemas software orientados a la gestión, pero, era necesario contar con una técnica útil para medir el tamaño funcional de otras aplicaciones. Con esta intención, se desarrolló una técnica experimental, denominada Puntos Característica (Feature Points), para adaptar la técnica de puntos función a sistemas software científicos y de ingeniería (Jones, 1996).

Los puntos característica se han venido utilizando con gran éxito en la medición de diversos sistemas software: sistemas en tiempo real, sistemas embebidos, software para inteligencia artificial, los cuales se caracterizan por la complejidad algorítmica que implementan y el escaso número de entradas y salidas que tienen.

La técnica de Puntos de Casos de Uso (Peralta, 2004, Ribu, 2001) permite determinar el tamaño de una aplicación de acuerdo con el número de actores y casos de uso involucrados. Luego, permite refinar este cálculo teniendo en cuenta el factor de complejidad técnica y el factor de ambiente en el cual se incluyen factores como las habilidades y el entrenamiento del grupo involucrado en el desarrollo.

Las técnicas presentadas anteriormente tienen como ventaja la confiabilidad en los resultados si se tiene toda la información requerida, pero, no son aplicables a proyectos de TI en general, y sólo se limitan a proyectos de desarrollo software, además, necesitan gran cantidad de datos que aún no se tienen en fases tempranas del desarrollo y trabajan sobre la base de una especificación de requisitos buena, es decir, no ambigua y bastante completa (Varas, 2002).

Luego de una rigurosa revisión de la literatura para el estudio de la problemática de la Gerencia Informática de Orbitel, se llegó a la conclusión de que las técnicas existentes no permiten estimar el esfuerzo de los proyectos a realizar antes de tener una completa especificación de requisitos, por lo tanto, se propone a continuación un nuevo método para estimar el esfuerzo en etapas tempranas.

UN NUEVO MÉTODO PARA CALCULAR EL ESFUERZO EN ETAPAS TEMPRANAS

Esta investigación tiene lugar en la gerencia informática de Orbitel S.A, cuyo objetivo es crear y operar soluciones. La gerencia informática, cuenta

actualmente con una estructura conformada por dos direcciones y 11 equipos de trabajo. La Dirección de Arquitectura y Soluciones Informáticas está integrada por los equipos: Arquitectura e Integración (AEI), Business Intelligence (BI), Ventas y Servicio al Cliente (VSC), Sistemas Administrativos y Financieros (SAF), Producto, Facturación, Plataformas Afines a Red (PAR). La Dirección de Infraestructura Informática y Atención a Usuario Final está integrada por los equipos: Planeación y Consecución de Recursos (PCR), Prestación de Servicios (PS), Gestión de la Relación con el Cliente (GRC), el cual no está ligado a ninguna de las direcciones ya mencionadas.

Cada uno de los equipos sigue un mapa de procesos bien establecido y trabaja bajo un esquema de autogestión que se rige por una política de hechos y datos en la cual cada procedimiento debe estar soportado y documentado. Por esto, el modelo para la estimación temprana de esfuerzo en proyectos de desarrollo de software cobra gran importancia en la definición de indicadores y métricas que facilitan la planeación y control de las actividades que se realizan.

Esta investigación toma como punto de partida la información de proyectos terminados, de los cuales existen datos históricos de esfuerzo. En conjunto con los analistas y directores de la gerencia informática, y haciendo uso tanto de la experiencia como de técnicas estadísticas, se logra tipificar y caracterizar los proyectos, con el fin de lograr, primero, un modelo capaz de sugerir intervalos de esfuerzo en etapas tempranas para un nuevo proyecto y, segundo, una herramienta de simulación que permita predecir, dado el esfuerzo estimado para cada nuevo proyecto, el comportamiento de la capacidad y los niveles de sobreesfuerzo de los analistas, las direcciones y en general de la gerencia informática.

La figura 1 presenta el diagrama de flujo que describe de forma global el modelo para calcular el esfuerzo en etapas tempranas:

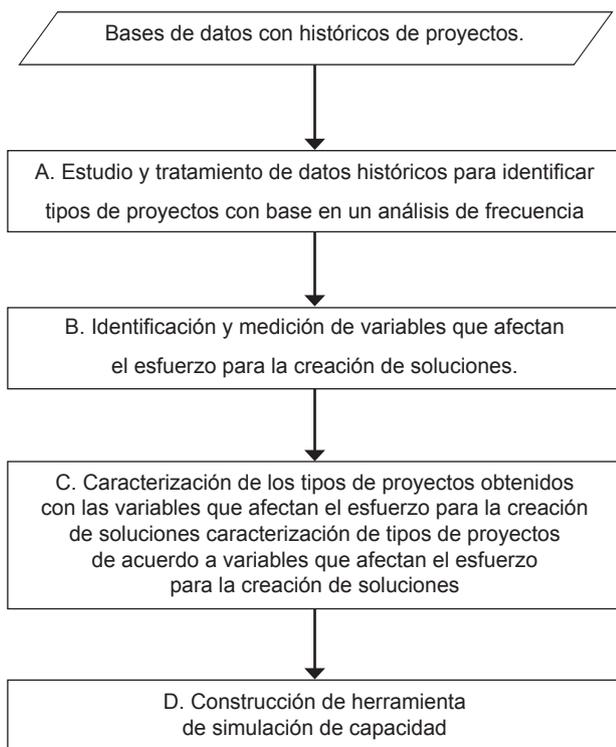


Figura 1. Diagrama de flujo con el procedimiento para la creación del modelo dinámico de estimación temprana

Para la creación de este modelo se desarrollaron las siguientes actividades: estudio y tratamiento de datos históricos, identificación de variables que afectan el esfuerzo para la creación de soluciones, caracterización de tipos de proyectos de acuerdo con las variables asociadas y, finalmente, construcción de una herramienta de simulación y pruebas.

ESTUDIO Y TRATAMIENTO DE DATOS HISTÓRICOS

En esta actividad se hizo un estudio estadístico para determinar de qué tipo de distribución provenían los datos relacionados con el esfuerzo (horas-hombre) requerido en el desarrollo de pro-

yectos, para luego dar una posible clasificación de proyectos.

El estudio se orientó a buscar la normalidad de los datos, o la forma de obtenerla, teniendo en cuenta que existe más información acerca del tratamiento de datos que siguen una distribución normal.

En la Gerencia Informática, los datos del esfuerzo se mueven en un rango bastante amplio de 0,25 a 15620 horas-hombre. Por medio de un histograma de frecuencias (figura 2) y algunas pruebas de normalidad (Prueba de Shapiro-Wilk y prueba de Anderson-Darling), se pudo determinar que los datos no seguían una distribución normal.

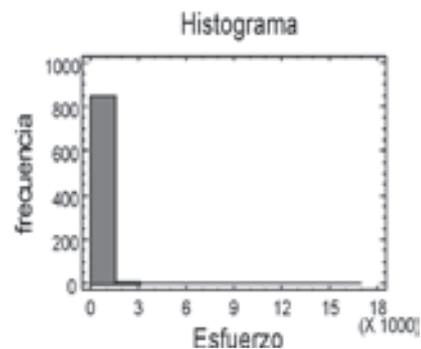


Figura 2. Histograma de frecuencia del esfuerzo (horas-hombre) requerido para el desarrollo de proyectos software

Para ajustar los datos a una distribución normal, se aplicó logaritmo natural y se eliminaron los proyectos que tienen un esfuerzo menor de 5 horas-hombre y mayor de 994,858 horas-hombre. Posteriormente, usando test de bondad de ajuste, reincorporando los datos que habían sido excluidos, y haciendo otros procedimientos estadísticamente válidos, se obtuvieron 13 intervalos tal como se muestra en la tabla 1, la tabla 1 muestra la clasificación de proyectos asociados a la creación de soluciones en la Gerencia Informática de Orbitel de acuerdo con el esfuerzo requerido para su desarrollo.

Tabla 1. Clasificación de proyectos de acuerdo con el esfuerzo asociado

Intervalo	Límite inferior	Límite superior
T0	0,25	4,9998
T1	4,9998	7,2087
T2	7,2087	13,0771
T3	13,0771	19,9846
T4	19,9846	28,5604
T5	28,5604	39,6226
T6	39,6226	54,4967
T7	54,4967	75,6048
T8	75,6048	108,0485
T9	108,0485	165,1229
T10	165,1229	299,5406
T11	299,5406	994,8580
Tx	994,858	15620

El intervalo T0 corresponde a los proyectos que están entre 0,25 y 5 horas-esfuerzo. Los intervalos T1 a T11 fueron obtenidos con el test de bondad de ajuste, y el intervalo TX corresponde a los proyectos que superan las 994,858 horas-hombre. Si bien los extremos de los intervalos obtenidos con los históricos disponibles son bastante distantes, el modelo creado y la herramienta de simulación permiten que los intervalos se ajusten de manera dinámica a medida que se ingresan nuevos proyectos.

En apartados posteriores, los tipos de proyectos obtenidos serán caracterizados con variables obtenidas para tal fin.

IDENTIFICACIÓN Y MEDICIÓN DE VARIABLES QUE AFECTAN EL ESFUERZO.

En esta actividad se estableció un método para descubrir las variables, eventos y retardos que aumentan el esfuerzo en el proceso de creación

de soluciones, incluyendo para cada una de ellas: nombre, descripción, unidades de medida, clasificación de valores y comentarios. Dichas variables se clasificaron en “Variables generales” (tabla 2) y “Variables por equipo” (tabla 3). Las variables generales son aquellas que después del análisis realizado, aplican para todos los equipos de la Gerencia, y las variables por equipo son variables que aplican a equipos particulares. A continuación se presentan algunas variables y sus atributos.

Tabla 2. Algunas variables generales que afectan el esfuerzo para la creación de soluciones en la Gerencia Informática de Orbitel

Nombre	Descripción	Unidades de medida
Sistemas de información y/o plataformas intervenidos	Cantidad y nombre de sistemas de información y/o plataformas que serán modificados o actualizados durante la creación de la solución.	Número y nombre de los sistemas de información y/o plataformas
Áreas usuarias involucradas en la especificación de requisitos	Áreas usuarias que deben participar en la especificación de requisitos de la solución a crear.	Número de áreas usuarias
Conocimiento del negocio	Habilidades y conocimientos adquiridos por el analista con soluciones similares a la que se va a gestionar.	Nivel de conocimiento del negocio

Para la evaluación y medición de variables como “conocimiento del negocio”, “conocimiento de nuevas tecnologías” y “conocimiento técnico”, las cuales hacen parte del conjunto de variables generales, se construyó un método, el cual no es ampliado en este trabajo.

Tabla 3. Algunas variables que afectan el esfuerzo para la creación de soluciones en la Gerencia Informática de Orbitel discriminadas por equipo

Equipo	Nombre	Descripción
BI	Impacto en la bodega de datos	Efecto producido por los cambios estructurales que se deben hacer en la bodega de datos para crear una solución.
	Tamaño de la ETL	Tamaño en la ETL, medido de acuerdo con la complejidad del proceso de extracción, transformación y carga de datos debido a la fuente y a la transformación de los datos.
PCR	Complejidad en la adquisición de hardware	Dificultad en la adquisición de hardware, medida de acuerdo con el número de proveedores existentes, con la disponibilidad del hardware en el mercado y con el proceso de selección y/o compra que se debe seguir.
FAC	Planes y/o servicios	Número y tipo de planes y servicios que es necesario crear o modificar para construir la solución.

El resultado de esta actividad fue utilizado para caracterizar los tipos de proyectos. La combinación de las variables asociadas a una nueva solución permite determinar la complejidad de una solución en términos del esfuerzo requerido.

CARACTERIZACIÓN DE TIPOS DE PROYECTOS

Luego de tener los resultados de las actividades de identificación de tipos e identificación y medición de variables, se realizó una caracterización de los tipos de proyectos con la intervención de analistas y directores de la Gerencia de Informática

de Orbitel. Para el estudio se tomó una muestra aleatoria de proyectos incluidos en cada uno de los intervalos de esfuerzo (tipos de proyectos) obtenidos con el análisis estadístico.

En la plantilla mostrada en la figura 3, los PROYECTOS representan la muestra de proyectos seleccionada tomando como base el total de proyectos que cada equipo realizó desde enero del año 2003 hasta marzo del año 2005. Las VARIABLES obtenidas en la segunda actividad, fueron asociadas a cada uno de los tipos de proyectos con base en la experiencia de los analistas de la Gerencia Informática y en la información recopilada de cada proyecto.



Figura 3. Plantilla de asociación de variables a tipos de proyectos

Los resultados obtenidos en esta actividad permiten que los analistas sólo necesiten indicar las variables que intervendrán en la nueva solución, y el modelo las combinará hasta conseguir el tipo que más se ajusta al nuevo proyecto. Es importante anotar que el modelo permite la inclusión de nuevas variables y tiene la capacidad de refinarse dinámicamente con cada nuevo proyecto incluido en la base de datos.

IMPLEMENTACIÓN DE LA APLICACIÓN

Luego de hacer el tratamiento estadístico de los datos y tener una clasificación de proyectos de acuerdo con las variables asociadas a cada tipo, fue necesario diseñar e implementar un algoritmo que le permitiera al analista conocer el tipo de proyecto a gestionar, luego de seleccionar las variables involucradas.

También fue necesario construir una base de datos para almacenar la clasificación de proyectos, las variables relacionadas, y otra información relevante.

Todo lo anterior se integró para obtener una aplicación orientada a la Web que permite a los analistas monitorear su capacidad, con sólo ingresar los proyectos que desean gestionar.

A continuación se muestran algunas pantallas generadas por la aplicación y sus respectivos comentarios (figuras 4 a 7):



Figura 4. Barra de menús de la aplicación

Los dos menús más relevantes de la aplicación son el menú “soluciones” y el menú “simular”. El menú SOLUCIONES permite ingresar una nueva solución, asignar participación de un analista y modificar una solución existente.

En el menú SIMULAR se pueden obtener series de tiempo sobre la capacidad empleada de un analista, un equipo, una dirección, y la Gerencia Informática. Los resultados obtenidos se mostrarán en la sección de resultados.

Para ingresar una nueva solución o proyecto en el cual intervendrá un analista se utiliza la interfaz mostrada en la figura 5.

Figura 5. Interfaz para ingreso de nuevas soluciones

En este formulario, el analista debe ingresar datos como: nombre del proyecto o solución, fecha de inicio, duración, porcentaje de participación y tipo de solución (intervalo de esfuerzo) en caso de conocerlo. Si se trata de una estimación temprana, la herramienta de simulación permite calcular el tipo de proyecto de acuerdo con las variables involucradas (figura 6)

Figura 6. Herramienta para calcular el tipo de solución

El formulario mostrado en la figura anterior permite señalar cuáles variables (con sus respectivos niveles) están presentes en la solución que gestio-

nará el analista, luego, la herramienta “simula” se encarga de indicar en qué nivel de la clasificación se encuentra una solución en la que intervienen las variables ya señaladas por el analista

RESULTADOS DEL MODELO DE CAPACIDAD

En la actualidad, el resultado de este proyecto de investigación aplicada es un modelo que le permite a los analistas de la Gerencia Informática no solamente hacer estimaciones tempranas con base en históricos y experiencias de proyectos similares, sino que les permite monitorear su capacidad y determinar cuándo están en situación de sobre-esfuerzo.

La interfaz del modelo está diseñada de manera que si la capacidad sobrepasa el 100%, al simular

algún mes, se presentará una señal de alerta, buscando que el analista ajuste su capacidad y no entre en condiciones de sobre-esfuerzo (figura 7). La alerta seguirá mostrándose hasta que el analista haga los ajustes necesarios.

Además, en la simulación es posible observar el comportamiento de la capacidad al agregar una solución de un tipo determinado, sin necesidad de almacenar esta solución en la base de datos. Como lo muestra la figura 7, en la parte inferior aparecerá un listado de las soluciones asociadas al analista.

El software permite también mostrar gráficos del comportamiento de la capacidad por equipo, por Dirección, y de toda la Gerencia Informática, informando siempre si se ha entrado en condición de sobre-esfuerzo.

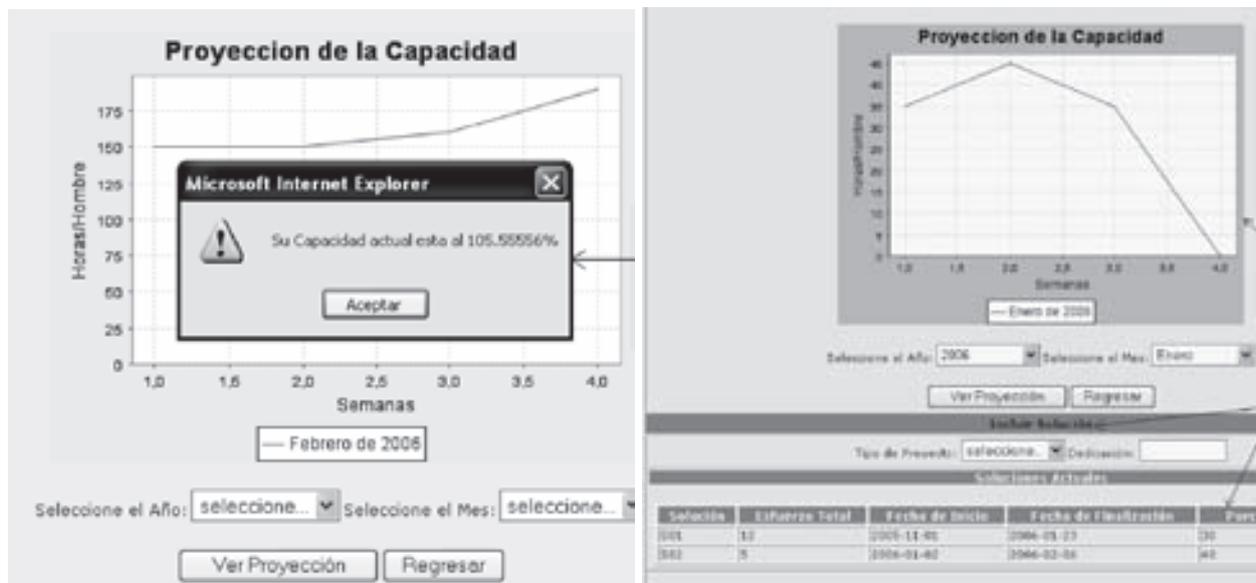


Figura 7. Interfaz gráfica del modelo de capacidad

CONCLUSIONES

La simulación del modelo dinámico para estimación temprana de esfuerzo en proyectos de software representa un avance importante que ayuda a entender la evolución de las divisiones de informática en términos de acumulación de conocimiento y explotación históricos de proyectos terminados.

A pesar de la incertidumbre implícita en proyectos de desarrollo de software, la existencia de bases de datos con acumulación dinámica de históricos de proyectos permitió establecer de manera muy aproximada el conjunto de variables asociadas a los procesos de creación de soluciones de software.

El modelo dinámico usa un número de parámetros, variables y funciones que caracterizan los proyectos y el entorno de una Gerencia Informática. Para realizar una estimación temprana de un nuevo proyecto, es necesario seleccionar las variables que caracterizan la nueva solución; una vez se cuenta con el grupo de variables que caracterizan la nueva solución, el modelo puede, además de realizar la estimación temprana de esfuerzo, planear y simular la capacidad del recurso humano.

Considerando que las variables identificadas en esta investigación son frecuentes para cualquier

gerencia informática, debido a que se establecieron para los procesos sugeridos por la metodología ITIL y que la base de datos con históricos de proyectos puede crecer de manera dinámica, el modelo creado es una herramienta que puede ser usada para hacer estimaciones y experimentaciones considerando las siguientes condiciones de desarrollo de proyectos de software y control de la capacidad:

- a. Cuando se requiere hacer una estimación con una especificación de requisitos insuficiente.
- b. Cuando la información disponible en bases de datos con históricos de proyectos sea incompleta.

El modelo creado y la herramienta de simulación pueden ser utilizados para la toma de decisiones o para la experimentación de políticas de contratación de recurso humano.

AGRADECIMIENTOS

Los autores quieren agradecer al equipo ingenieros de la Gerencia Informática de Orbitel S. A., quienes participaron activamente en la interpretación y análisis de los resultados del modelamiento y la simulación. También, a la Universidad de Antioquia y a Orbitel S.A por patrocinar esta investigación.

BIBLIOGRAFÍA

- ALBRECHT, A. J. 1979. Measuring application development productivity. Proceedings SHARE/GUIDE IBM Applications Development Symposium, Monterrey, CA., Oct 14-17.
- ALBRECHT, A. J. & GAFFNEY, J. E. 1983. Software function, source lines of code, and development effort prediction: A software science validation. IEEE Transactions on Software Engineering, Vol. SE-9, no. 6, pp. 639-648
- ARORA, A. & GAMBARDILLA, A. 2005. From underdogs to tigers the rise and growth of the software industry in Brazil, China, India, Ireland, and Israel. Disponible en: <http://www.oup.co.uk/pdf/0-19-927560-2.pdf>
- JONES, C. 1996. Activity-based software costing. Computer, May 1996, p. 103-104.
- JORGENSEN, M. & SJOBERG, D. 2001. Impact of effort estimates on software project work. Information and software technology. 43: 939-948.

- PERALTA, M. 2004. Estimación del esfuerzo basada en casos de uso. Reportes Técnicos en Ingeniería de Software. Buenos Aires- Argentina. 6 (1): 1-16.
- RIBU, K. 2001. Estimating object-oriented software projects with use cases. Tesis de master. Universidad de Oslo. Disponible en: <http://heim.ifi.uio.no/~kribu/oppgave.pdf>. Acceso: Octubre 21 de 2005
- SÁNCHEZ, F. 1999. Medida del tamaño funcional de aplicaciones software. Universidad de Castilla- La Mancha.
- TORO, M., RUIZ, M. & RAMOS, I. 2002. Marco dinámico integrado para la mejora de los procesos software.
- VARAS, M. 2002. Una Experiencia con la estimación del tamaño del software disponible en: <http://www.inf.udec.cl/revista/edicion1/mvaras.htm>.
- ZHENYOU. 2004 Reduced models of software development effort estimation. Department of electrical and computer engineering. University of Alberta. Edmonton, Alberta. 2004
- 1987. Improving software productivity. IEEE computer. Vol 20 . Pag 43-57.