

Tipo de artículo: Artículo original
Temática: Ingeniería y Gestión de software
Recibido: 11/03/2013 | Aceptado: 30/04/2013

Aplicando métricas de calidad a proyectos y procesos durante las pruebas exploratorias

Quality metrics applied to projects and processes during exploratory testing

Yeniset León Perdomo, Asnier Enrique Góngora Rodríguez, Ailyn Febles Estrada

Calisoft. Centro Nacional de Calidad de Software. Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños, km 2½, Torrens, Boyeros, La Habana, Cuba. CP.: 19370

{yleonp, agongora, ailyn}@uci.cu

Resumen

En muchas áreas de la gestión, se proponen las mediciones como una herramienta eficaz para ayudar en la obtención del éxito de proyectos de software y sistemas a partir de cuatro razones básicas conocidas: caracterizar, evaluar, predecir y mejorar. Dentro de los servicios que brinda el Centro Nacional de Calidad de Software se encuentran las Pruebas Exploratorias a productos de software desarrollados por terceros. En este artículo se describe la experiencia del Departamento de Pruebas de Software de incorporar durante las Pruebas Exploratorias un grupo de métricas internas y externas para evaluar el producto desde el propio proceso de prueba. En este artículo se describe el Departamento de Pruebas de Software, se presentan las diferentes ideas y conceptos de las Pruebas Exploratorias existentes en la literatura y la aplicación de un grupo de métricas de calidad dentro del proceso de exploración del producto.

Palabras clave: Calidad, métricas, pruebas exploratorias.

Abstract

In many areas of management measures are proposed as an effective tool to assist in securing the success of software projects and systems from four known basic reasons: to characterize, evaluate, predict and improve. Among the services offered by the National Centre for Software Quality (CALISOFT) are the Exploratory Testing (ET) to software products developed by third parties. This article describes the experience of the Department of Software Testing (DPSW) during ET by incorporating a group of internal and external metrics to evaluate the product from the testing process itself. It also describes the DPSW, and presents different ideas and concepts of ET in the literature and the application of a set of quality metrics within the exploration process of the product.

Keywords: *Exploratory testing, metrics, quality.*

Introducción

La medición es considerada como una eficaz herramienta en las pruebas a los software, es la base para: detectar las desviaciones del rendimiento aceptable en los procesos y producto de software, y las oportunidades de mejora, identificar y priorizar las principales preocupaciones, dar seguimiento a la solución y mejorar la calidad del producto. Las mediciones permiten además cuantificar tanto el proceso como el producto. Proporcionan la visión del desempeño del proceso permitiendo: desarrollar perfiles de los datos de los proyectos anteriores que se pueden utilizar para la planificación y mejora del proceso; analizar un proceso para determinar cómo mejorarlo; determinar la eficacia de modificaciones en el proceso, (Pomero y Cannon, 2003).

En el proceso de prueba las mediciones pueden ser usadas para, (Kaner y Pretichord, 2001).

- Monitorizar el proceso de prueba: Mostrar visibilidad sobre las actividades de pruebas. Esta información puede ser usada para medir el criterio de terminación de las pruebas y evaluar el progreso contra lo planificado.
- Reportar las pruebas: Métricas recolectadas al finalizar cada etapa de prueba para evaluar la adecuación de los objetivos de la etapa, la adecuación de la estrategia de pruebas tomada y la efectividad de las pruebas con respecto a sus objetivos.
- Controlar las pruebas: Acciones correctivas tomadas como el resultado de la información, las métricas tomadas y reportadas.

En el Laboratorio Industrial de Pruebas de Software (LIPS) del Departamento de Pruebas de Software de (Calisoft) no se evalúan los productos de software por las características de calidad debido a la carencia de mediciones realizadas a

los mismo, lo que influye negativamente en la opinión del cliente sobre la calidad del producto, de ahí la necesidad de insertar desde etapas tempranas de las pruebas un grupo de métricas de calidad.

Centro Nacional de Calidad de Software. Departamento de Pruebas de Software

Calisoft es una unidad presupuestada subordinada al Ministerio de la Informática y las Comunicaciones (MIC). Cada uno de los grupos que conforman el centro tienen definidos sus objetivos, misión y estrategia de trabajo.

Uno de ellos es el DPSW que está constituido por tres grupos: el Grupo de Ingeniería de Pruebas (GIPS), el Laboratorio Industrial de Pruebas de Software (LIPS) y el Grupo de Seguimiento y Control de la Calidad de los Servicios (SCCS). Este departamento tiene como misión principal brindar servicios de pruebas de software a la Industria Cubana del Software (InCusoft).

Los diferentes tipos de pruebas a ejecutar en este departamento son: Funcionalidad, Confiabilidad, Usabilidad, Eficiencia, Mantenibilidad, Potabilidad, Fiabilidad y Revisión Técnica a la Documentación.

Las pruebas se ejecutarán chequeando los atributos de McCall (Macías Gómez, 2010) y lo planteado en la norma ISO 9126 (ONN, 2005).

Aunque las PE no son un tipo de pruebas definido por la 9126 estas se aplican en el DPSW asociadas a los diferentes atributos de calidad, dada la necesidad de obtener retroalimentación rápida de cierto artefacto, producto o funcionalidad.

El funcionamiento del DPSW (Figura.1) comienza desde que es realizada la solicitud de pruebas del/los artefacto(s) hasta que se libera. Cada uno de ellos puede ser sometido a diferentes períodos de pruebas (iteraciones) antes de ser liberado.

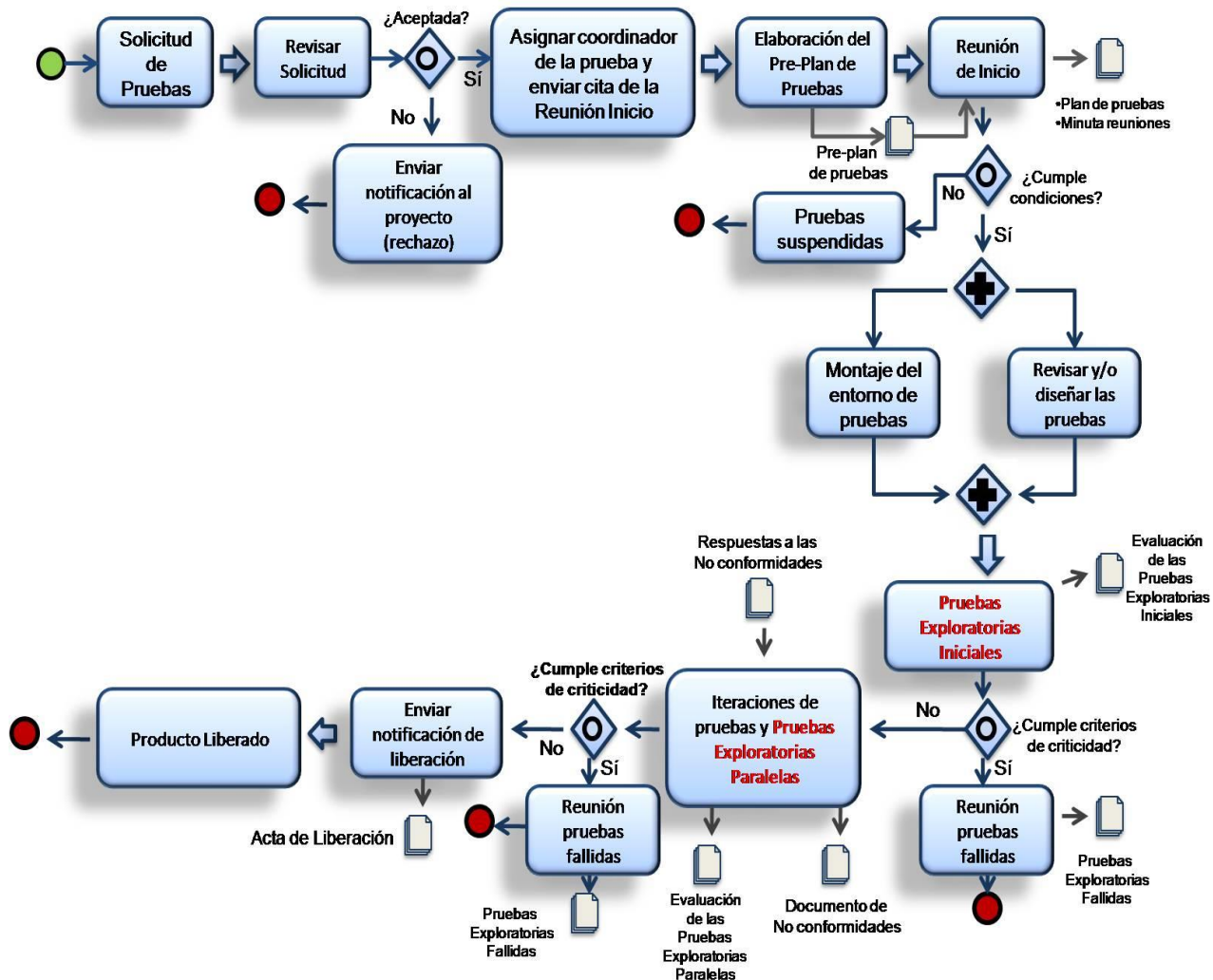


Figura. 1. Inserción de las PEI y PEP en el proceso del Laboratorio Industrial de Pruebas de Software.

Primeramente se confecciona el Pre-Plan de Pruebas por especialistas del departamento y se convoca a la reunión de inicio donde queda conformado todo el proceso de pruebas. Una vez finalizada la reunión se ejecuta de forma paralela la revisión y refinamiento de los diseños de prueba y el montaje del entorno.

Posteriormente se aplican las PEI, para ello se define la muestra por artefacto y se utilizan como herramientas las listas de chequeo y/o los casos de prueba.

Finalizadas en estas pruebas el entregable puede o no ser abortado, o sea, declararse la prueba detenida o abortada. Si la misma aborta se convoca a una reunión con los clientes donde se le exponen las causas por las que no se continúa con el proceso de prueba, mostrándole los documentos: Prueba Detenida o Prueba Abortada.

En caso contrario se ejecutan las Pruebas Funcionales para dicho entregable, siguiendo cierta cantidad de iteraciones (≤ 3) hasta que el mismo sea liberado o abortado.

Para declarar que un artefacto se encuentra en estado crítico de terminación, dígase, abortado o detenido, se aplican los Criterios de Criticidad (León, 2012).

Los Criterios de Criticidad establecen los parámetros para declarar un producto de desarrollo de software en estado crítico de terminación.

Son aplicables en 3 momentos fundamentales:

- Cuando se hace la reunión de inicio con el equipo de proyecto
- Durante el proceso de PEI
- Durante las iteraciones de prueba

Pruebas exploratorias y Pruebas exploratorias iniciales en el DPSW

El término “Pruebas Exploratorias” fue introducido por Cem Kaner, se refiere a ejecutar las pruebas a medida que se piensa en ellas, sin gastar demasiado tiempo en preparar o explicar las pruebas, confiando en los instintos. (Bach, 2002).

James Bach define como PE al proceso simultáneo de exploración del producto (aprendizaje), diseño y ejecución de pruebas (Bach, 2002).

De acuerdo a las características del LIPS se ha tenido que adaptar las diferentes ideas y conceptos de PE existentes en la literatura, para dar lugar a un nuevo concepto. Esto no excluye la utilización de las PE tal y como las define James Bach. El propio proceso de pruebas del laboratorio requiere que se utilicen ambas técnicas de pruebas: PE y PEI. Para mejor entendimiento y explicación del momento en que se utilizarán las PE en esta investigación, las mismas se dividen en: PEI y Pruebas Exploratorias Paralelas (PEP).

“Las Pruebas Exploratorias Iniciales se definen como el proceso inicial de exploración del producto, diseño y ejecución de pruebas, que valida la calidad de este, permitiendo o no el paso a nuevas etapas”. (León, 2012).

Estas pruebas como su nombre lo indica serán las primeras que se realizarán en el DPSW, antes de la primera iteración de las pruebas funcionales. A su vez las PEP serán aplicadas paralelas fundamentalmente a la segunda iteración de las pruebas funcionales o a cualquier otra iteración en que el especialista de prueba las solicite. En la Figura 1 se evidencia la aplicación de las PEI y PEP dentro del proceso de pruebas de liberación del DPSW.

Las PE tiene cinco elementos externos básicamente: tiempo, probador, producto, misión e informes. Esta aparente simplicidad permite desplegar una amplia gama de posibilidades para la aplicación de las PE. Una prueba exploratoria

se lleva a cabo en un tiempo delimitado, sobre un producto específico, por medio de un/unos probadores en particular, tratando de cumplir una misión, el cual puede reportar el estado y resultados de la misma en cualquier momento.

Métricas de calidad en las PE

Pressman clasifica el campo de las métricas en 6 categorías o grupos de métricas distintos, dentro de ellas se encuentran las: Métricas de calidad. Éstas proporcionan una indicación de cómo se ajusta el software a los requisitos implícitos y explícitos del cliente. Es decir cómo se va a medir para que el sistema se adapte a los requisitos que pide el cliente. (Pressman, 2009).

Métricas internas

Existe un grupo de métricas internas que aunque son aplicadas fundamentalmente a los productos intermedios que se desarrollan a lo largo del ciclo de vida de desarrollo, los autores consideran que las mismas pueden ser adaptadas para reportar el avance en la ejecución de las pruebas en evaluaciones realizadas al final del proceso o ciclos de pruebas (PEI, PEP, iteraciones), por eso es que se toma la decisión de incluirlas como parte de la estrategia desarrollada en esta investigación.

A continuación se listan las mediciones al proceso que se tienen en cuenta en esta investigación. Dichas mediciones fueron recolectadas de: (ISTQB, 2005; Kaner y Pretichord, 2001; Kit, 1995).

Tabla 1. Métricas sobre el avance del proceso en la ejecución de las pruebas.

Elemento	Métrica	Propone medir
Funcionalidad	Cubrimiento de funcionalidades	Cantidad de casos de prueba ÷ Total de funcionalidades a probar.
	Tiempo por funcionalidad	Total de tiempo en ejecutar las pruebas ÷ Total de funcionalidades probadas.
	Densidad de defectos por funcionalidad	Total de No conformidades (NC) encontrados ÷ Total de funcionalidades probadas.
PE	Esfuerzo de las PE	Contar el tiempo dedicado a las PE en el ciclo de pruebas.
	NC encontradas en las PE	Contar las NC encontradas durante las PE.
	Funcionalidades exploradas	Contar las funcionalidades exploradas durante las PE del ciclo de prueba.

Esfuerzo	Esfuerzo planificado (5)	Sumar el tiempo planificado para todas las actividades del proyecto de prueba.
	Esfuerzo realizado en el ciclo de prueba	Sumar el tiempo invertido en todas las actividades del ciclo de prueba.
	Esfuerzo realizado hasta el momento (3)	Sumar el tiempo realizado desde que el proyecto comenzó hasta el último ciclo realizado.
	Esfuerzo por realizar (4)	Sumar el tiempo que se planifica para cada actividad hasta que el proyecto de prueba termine.
	Esfuerzo total (6)	Sumar (3) y (4).
	Desviación en el esfuerzo planificado	Restar (6) menos (5).
	Esfuerzo por persona	Contar el esfuerzo por persona del equipo de prueba.

Métricas externas

La ISO 9126 parte 2 (ONN, 2005) contiene un conjunto de métricas externas que permiten evaluar calidad del producto de software durante la prueba, sin asignar rangos de valores a las métricas que propone ya que son específicas para cada producto, en dependencia de su categoría, nivel de integridad y necesidad del usuario final. Las métricas están distribuidas por las seis características que recoge la parte 1 Modelo de Calidad, (ONN, 2005) ellas son: Funcionalidad, Confiabilidad, Usabilidad, Eficiencia, Mantenibilidad y Portabilidad.

Para esta investigación solamente se tendrán en cuenta las métricas asociadas a las características de calidad de funcionalidad midiendo la idoneidad y funcionalidad, confiabilidad teniendo en cuenta la madurez y usabilidad midiendo comprensibilidad e instructibilidad. (León, 2012; Góngora, 2011).

Característica de funcionalidad

Las métricas externas de funcionalidad deben ser capaces de medir de un atributo como es el comportamiento funcional del sistema en el cual el software está presente. En esta investigación solamente se tiene en cuenta la métrica de idoneidad.

- Métricas de idoneidad

Las métricas externas de idoneidad deben ser capaces de medir de un atributo como es la ocurrencia de un funcionamiento insatisfactorio o la ocurrencia de una operación insatisfactoria.

Un funcionamiento u operación insatisfactoria puede ser:

- Funcionamiento u operación que no se desempeña de la forma especificada en el manual de usuario o la especificación de requisitos.
- Funcionamiento u operación que no provee una salida aceptable o razonable al tomar en consideración un objetivo específico de las tareas del usuario.
- Métricas de funcionalidad

Una métrica de funcionalidad externa debe ser capaz de medir el cumplimiento de un atributo con el número de funciones, o por otros acontecimientos de los problemas de cumplimiento como que el producto de software no cumpla con las normas, convenios, contratos u otros requisitos reglamentarios.

Tabla 2. Métricas externas de calidad. Características de funcionalidad..

Funcionalidad		
Elemento	Métrica	Propone medir
Idoneidad	Estabilidad en las especificaciones funcionales	¿Cuán estable son las especificaciones funcionales después de entrar a funcionar?
Funcionalidad	Cumplimiento de estándar de interfaz	¿Cómo se cumplen las interfaces con los reglamentos aplicables, las normas y regulaciones?

Característica de confiabilidad

Las métricas externas de confiabilidad deben ser capaces de medir atributos relacionados con el comportamiento del sistema del cual el software forma parte durante la ejecución de las pruebas para indicar la magnitud de la confiabilidad, o sea, seguridad de funcionamiento del software durante la operación del sistema, con las que en la mayor parte de los casos no se distingue entre el software y el sistema. En esta investigación solamente se tiene en cuenta la métrica de madurez.

- Métricas de madurez

Las métricas externas de madurez deben ser capaces de medir de un atributo como la exención de fallas en el software, causados por la ocurrencia de fallos existentes en el propio software.

Tabla 3. Métricas externas de calidad. Características de confiabilidad.

Confiabilidad		
Elemento	Métrica	Propone medir
Madurez	Grado de solución ante fallos totales	¿Cuántas condiciones de fallo total están resueltas?
	Erradicación de fallos	¿Cuántos fallos han sido corregidos?
	Densidad de fallos totales contra casos de prueba	¿Cuántos fallos totales fueron detectados durante un período de pruebas definido?
	Madurez de las pruebas	¿Está bien probado el producto?

Característica de usabilidad

Las métricas externas de usabilidad miden la dimensión con que el software puede ser comprendido, estudiado, operado, atractivo y acorde con las regulaciones y guías relativas a la usabilidad.

Resulta recomendable que la evaluación de estas métricas se haga por un grupo (7 u 8 aunque menores pueden obtener información de utilidad) de usuarios o evaluadores, usuarios simulados o clonados (pero representativos de un rango de usuarios) sin que reciban asistencia externa alguna.

- Métricas de comprensibilidad

Las métricas externas de comprensibilidad deben ser capaces de valorar cómo un nuevo usuario podría comprender:

- Si el software es idóneo para la aplicación a la cual lo destina.
- Cómo el software puede ser usado para una tarea en particular.

- Métricas de instructibilidad

Una métrica externa de instructibilidad debe ser capaz de evaluar cómo los usuarios a larga distancia le llevará aprender a utilizar funciones específicas y la eficiencia se los sistemas de ayuda y documentación.

La instructibilidad está muy relacionada con claridad, comprensibilidad y medidas pueden ser indicadores de la potencial capacidad de aprendizaje del software.

Tabla 4. Métricas externas de calidad. Características de usabilidad.

Usabilidad		
Elemento	Métrica	Propone medir
Compresibilidad	Accesibilidad a demos	¿A qué proporción de demos/tutoriales pueden acceder los usuarios?
Instructibilidad	Eficiencia de la documentación del usuario y / o sistemas de ayuda en el uso.	¿Qué proporción de las funciones pueden ser utilizados correctamente después de leer la documentación o el uso de sistemas de ayuda?

Aplicación de las métricas propuestas

Se presentan los resultados de aplicar las métricas propuestas en la estrategia a través de verificaciones realizadas durante las pruebas a proyectos seleccionados, logrando medidas de la efectividad del proceso de prueba a partir de la evaluación de los productos liberados, teniendo en cuenta el cumplimiento de las características de calidad de la ISO 9126 Parte 1: Modelo de calidad. Los datos para definir la media fueron tomados de (Góngora, 2011), la cual se rige por la metodología propuesta por la IEEE *Standard for a Software Quality Metrics Methodology*.

Resultados de las métricas

En la implementación de las métricas de calidad se utilizaron 42 proyectos, algunos se encuentran en pruebas en LIPS y otros ya fueron liberados por especialista del mismo departamento. Estos proyectos fueron seleccionados ya que tienen toda la documentación y los datos necesarios para obtener las métricas de calidad.

Los datos son recolectados mediante la herramienta para la gestión de las métricas (Góngora, 2011).

El cálculo de las métricas en esta investigación proporcionará una evaluación preliminar del producto antes de entrar a la 1ra iteración de las pruebas de liberación. A continuación se muestran los valores de la media por característica de calidad que se utilizará para evaluar el proceso y el producto.

Tabla 5. Resultado de la métrica de funcionalidad.

Subcaracterística de calidad	Nombre de la Métrica	% de cumplimiento de la característica de calidad
Idoneidad	Adecuación funcional	62,2 %
Idoneidad	Estabilidad en la especificaciones funcionales	64,1%

Funcionalidad	Cumplimiento de estándar de interfaz	67,0 %
---------------	--------------------------------------	--------

Tabla 6. Resultado de la métrica de confiabilidad.

Subcaracterística de calidad	Nombre de la Métrica	% de cumplimiento de la característica de calidad
Madurez	Densidad de fallos totales contra casos de prueba	6,8 %
Madurez	Grado de solución ante fallos totales	95,3 %
Madurez	Cobertura de las pruebas	99,0 %
Madurez	Madurez de las pruebas	45,4 %

Tabla 7. Resultado de la métrica de usabilidad.

Subcaracterística de calidad	Nombre de la Métrica	% de cumplimiento de la característica de calidad
Comprensibilidad	Accesibilidad a demos	50,0 %
Instructibilidad	Eficiencia de la documentación del usuario y / o sistemas de ayuda en el uso	83,7 %

Análisis de los resultados de las métricas

Una vez obtenidas las métricas de calidad se hace necesario comparar los resultados de las métricas calculadas en cada proyecto y la métrica general de cada sub-características de calidad, para así ver el cumplimiento de las sub-características en el producto.

Para demostrar cómo se realiza el cálculo de las métricas en las PE, se tomaron como muestra dos proyectos aleatorios, una vez culminado el ciclo de PE se emitió de cada uno el Informe de Evaluación de las Pruebas, donde dio como resultado que el proyecto A su prueba fue satisfactoria y el proyecto B se evaluación fue abortada. A continuación se presenta la Figura 2 donde se muestra la evaluación de los dos proyectos (A y B) con respecto a la media calculada anteriormente.

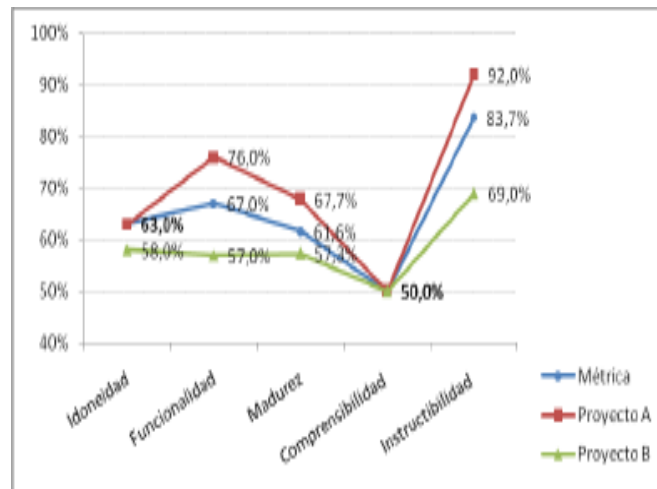


Figura 2. Evaluación del comportamiento de las características de calidad en los proyectos A y B.

Del resultado mostrado en la figura anterior el proyecto A representado con la línea roja obtiene mayores por ciento de cumplimiento de las sub-características evaluadas anteriormente que el proyecto B. Esto reafirma que el proyecto A puede pasar a la 1ra iteración de las pruebas de liberación, no siendo así el proyecto B representado con la línea verde, puesto que este último la evaluación emitida de las PE fue abortado, pues se le encontraron mayor cantidad de NC que la media establecida documentadas en los criterios de criticidad. Además el tiempo de ejecución de las pruebas fue mayor que el estipulado para este tipo de proyecto y la documentación presentada no cumple con los criterios establecidos.

Conclusiones

Se aplicaron métricas propuestas en la investigación a través de verificaciones realizadas durante las pruebas a proyectos seleccionados, logrando medidas de la efectividad del proceso de prueba a partir de la evaluación de los productos.

La definición y selección de las métricas de calidad teniendo en cuenta las características de calidad según la ISO 9126-1 Parte 1: Modelo de Calidad, permite tener una referencia de cuál es el por ciento de cumplimiento que tienen los software de las características de calidad, ayudando a los directivos a la toma de decisiones y a mejorar la calidad del proceso de pruebas y del producto final. Además estas métricas se incluyen en esta estrategia de PE con la idea de brindarle al equipo de desarrollo una evaluación completa de su producto.

Referencias

- BACH, J. “Exploratory Testing Explained”. The Test Practitioner, 2002. Disponible en: [<http://www.satisfice.com/articles/et-article.pdf>].
- GÓNGORA, A. "Catálogo automatizado de métricas de calidad para evaluar los productos en las pruebas". Tesis de Maestría en Calidad de Software, Universidad de las Ciencias Informáticas, Habana, Cuba, 2011.
- ISTQB. "International Software Testing Qualifications Board, Certified Tester Foundation Level Syllabus". 2005. Disponible en: [<http://www.istqb.org/fileadmin/media/SyllabusFoundation.pdf>].
- KANER, C. y B. PRETICHORD. “Lessons Learned in Software Testing”. ISBN 0471081124, 2001. Disponible en: [<http://www.istqb.org/fileadmin/media/SyllabusFoundation.pdf>].
- KIT, E. “Software Testing In The Real World : Improving The Process”. Addison Wesley. ISBN 0201877562, 1995.
- LEÓN, Y. “Estrategia de Pruebas Exploratorias para mejorar el rendimiento del Laboratorio Industrial de Pruebas de Software de Calisoft”. Tesis de Maestría en Calidad de Software, Universidad de las Ciencias Informáticas, Habana, Cuba, 2012.
- MACÍAS, J. E. y J. M. GÓMEZ REYNOSO. "Utilizando el Modelo de Calidad de McCall y el Estándar ISO-9126 para la Evaluación de la Calidad de Sistemas de Información por los Usuarios". AMCIS 2010 Proceedings. Paper 89, 2010. Disponible en: [<http://aisel.aisnet.org/amcis2010/89>].
- NC-ISO-IEC 9126-1. Parte 1. “Modelo de Calidad”. Oficina Nacional de Normalización, ONN, 2005. Disponible en: [<http://www.iso.org/iso/home.html>].
- NC-ISO-IEC 9126-2. Parte 2. “Métricas externas”. Oficina Nacional de Normalización, ONN, 2005. Disponible en: [<http://www.iso.org/iso/home.html>].
- POMEROY-HUFF, M.; CANNON, R. y S. SEBERN. “The Personal Software Process (PSP) Body of Knowledge”. Versión 1.0, CMU/SEI-2005-SR-2003. 2005. Disponible en: [<http://www.sei.cmu.edu/publications/documents/05.reports/05sr003.html>].
- PRESSMAN, R. S: “Ingeniería de Software. Un Enfoque práctico”. 6^{ta} Edición, 2009.