

Análisis de tiempos de lanzamiento de
proyectos comunitarios de software libre
mediante análisis de redes sociales: Caso
practico Debian GNU/Linux
Protocolo

Wences Arana

18 de noviembre de 2015

Índice general

Propósito del proyecto

Todo software desde su comienzo como una idea general hasta llegar a ser un sistema integro y funcional, conlleva un proceso ordenado con limites definidos en cuestiones de tiempo y cumplimiento de metas. Dicho cumplimiento de metas debe ser dentro de un marco tolerable de tiempo, esto con el fin de tener un proceso eficiente de desarrollo de software, un mejor manejo de costos de producción, calidad del mismo entre otras cualidades. Por esto mismo la medición de los factores que influyen en el atraso del cumplimiento de metas se hace de vital importancia para el estudio de la ingeniería de software, para lograr crear un proceso eficiente de desarrollo replicable y dentro de los recursos disponibles. Por esto mismo se plantea realizar dicho análisis utilizando el área de la teoría de análisis de redes sociales con el fin de por medio de un estudio de los datos disponibles crear un modelo replicable en proyectos de software con características similares, en nuestro caso se hará el estudio en base a sistemas operativos libres desarrollados de manera colaborativa por voluntarios y otros actores interesados, tomando en cuenta factores relacionados con la unidad mínima de software denominada paquete que conforma una red de estos que a su vez conforman el sistema operativo en su totalidad, además de las redes que existen en las relaciones sociales de los participantes en el desarrollo de dicho software. El estudio servirá a todos los interesados en obtener un modelo para poder encontrar posibles fallas en la gestión del proyecto relacionadas con los distintos componentes que forman parte de su proyecto de software tanto social como técnico, tangible e intangible.

Introducción

La teoría de análisis de redes tiene como objetivo encontrar relaciones en eventos de la naturaleza que no parecen tener incidencia unos con otros a primera vista, sin embargo al realizar un análisis profundo de sus relaciones, en algunos casos, pueden encontrarse desde relaciones de causa efecto entre eventos, actores, recursos, hasta la importancia de algunos de estos en el funcionamiento de todo el sistema en su conjunto, subgrupos relacionados, relaciones inesperadas, entre otras. Utilizando dicha teoría se pretende analizar el proyecto colaborativo de software libre Debian y su sistema operativo Debian GNU/Linux con el objetivo de encontrar los factores que inciden en el tiempo de lanzamiento de versiones estables de este. Para lograr este objetivo se manejarán varios enfoques para medir dicho impacto y podemos definirlos como: utilizar como principales nodos la unidad mínima funcional de software en el proyecto denominada paquete, la figura de desarrollador de software siendo en este caso tanto un desarrollador Debian como un desarrollador original del software que da origen al paquete, como relaciones entre nodos se tomarán en cuenta, dependencias de construcción entre paquetes, dependencias en tiempo de ejecución de los paquetes, bugs detectados en los paquetes que interfieran en la relación entre ellos, cantidad de bugs reportados de parte de un desarrollador Debian a un desarrollador upstream, cantidad de parches aplicados por parte de upstream enviados por un desarrollador Debian, cantidad de correos entre desarrolladores. Con esta información modelada se podrá analizar por medio de algoritmos existentes de teoría de análisis de redes sociales, en caso de poder ser aplicables, de lo contrario se creará el adecuado para poder realizar de manera apropiada los datos y de esta forma poder encontrar, patrones de centralidad, incidencia de clanes, nodos de mayor importancia, etcétera, y con esto podrá interpretarse los valores y hacer un análisis del impacto en el tiempo de desarrollo de una nueva versión estable del proyecto de software a otra.

Antecedentes

En el campo de estudio de la teoría de análisis de redes, se han realizado investigaciones relacionadas a proyectos de software pudiendo mencionar análisis de comunicación entre desarrolladores en un proyecto de software¹, las implicaciones de la coordinación de la arquitectura de software en proyectos de software², para encontrar patrones de colaboración en equipos de desarrollo de software³, entre otros casos de estudio en el área de ingeniería de software. De tal forma que el estudio se encargara en el área de tiempos de lanzamiento y de desarrollo que se basan en dichos estudios anteriores, aplicándolo a un proyecto colaborativo de software libre y definiendo un modelo en base a los datos, buscando encontrar los principales cuellos de botella y áreas a mejorar para poder eficientar el proceso de desarrollo de software de la distribución Debian GNU/Linux.

¹Crowston, Kevin, and James Howison. "The social structure of free and open source software development." *First Monday* 10.2 (2005).

²Ryan, Sharon, and Rory V. O'Connor. "Development of a team measure for tacit knowledge in software development teams." *Journal of Systems and Software* 82.2 (2009): 229-240.

³Ehrlich, Kate, Giuseppe Valetto, and Mary Helander. "Seeing inside: Using social network analysis to understand patterns of collaboration and coordination in global software teams." *Global Software Engineering, 2007. ICGSE 2007. Second IEEE International Conference on*. IEEE, 2007.

Justificación

La investigación tiene un fundamento esencial de existir debido a la necesidad de poder encontrar los problemas que se llevan a cabo en proyectos colaborativos de software de gran envergadura.

El motivo de la elección de un proyecto comunitario de software libre como Debian GNU/Linux se debe a factores sociales y técnicos, sociales debido a que la estructura organizacional de la comunidad de desarrolladores es pública y abierta lo cual es una ventaja para poder obtener información de relaciones entre toda la comunidad. Desde el punto de vista técnico, al ser un proyecto que desarrolla un sistema operativo de software libre permite poder tener acceso a toda la información referente a los enlaces y relaciones que existen entre las unidades de software que forman al sistema operativo en general, desde relaciones de dependencias directas hasta relaciones indirectas en el funcionamiento integral del sistema operativo como tal.

Debido a lo anteriormente expuesto tomando en cuenta otras teorías de análisis de datos existentes, la mas adecuada para analizar información y poder encontrar factores medibles que afecten a la liberación de nuevas versiones de software es la teoría de análisis de redes sociales esto debido a que la organización del proyecto tanto social como técnicamente pueden ser descritos como una red donde pueden encontrarse tendencias y poder aplicarse un modelo replicable en otras investigaciones, el proyecto ya provee un entramado interrelacionado lo cual facilita la utilización de análisis de redes sociales en la organización.

El proyecto es beneficioso por que brindará un modelo replicable para todo gestor de software que cuente con un proyecto de software con múltiples componentes interrelacionados y un equipo de colaboradores heterogéneos tanto internos(desarrolladores) como externos(terceros).

Las implicaciones que tendrá el trabajo irá orientado a la creación de un modelo replicable para determinar tiempos de lanzamientos.

El trabajo va dirigido a toda persona que necesite modelos para poder determinar que puntos modificar en su organización de manera técnica y social de tal forma que pueda mejorar tiempos de lanzamiento y obteniendo calidad en un tiempo razonable con una mayor anticipación.

MARCO TEÓRICO

- 1. Análisis de redes sociales
 - 1.2 Centralidad, cuellos de botella,
 - 1.3 Clicas, clusters y componentes,
- 2. Proyecto Debian,
 - 2.1 Estructura social
 - 2.2 Estructura técnica
 - 2.3 Medios de comunicación