

Emilio José Rodríguez Posada

Oficina de Software Libre y Conocimiento Abierto, Universidad de Cádiz; Ganador del III Concurso Universitario de Software Libre en la categoría "Mejor proyecto comunitario"

<emiliojose.rodriguez@uca.es>

1. Introducción

Wikipedia [5] es la enciclopedia libre que todos pueden editar. Cualquier persona con un navegador web puede modificar sus contenidos. Esto es posible en gran medida gracias al software que utiliza, MediaWiki [6]. La popularidad y el éxito de Wikipedia han alcanzado cotas inimaginables hace unos años, se encuentra entre los 10 sitios más visitados de Internet, suele aparecer entre los primeros resultados de la mayoría de las búsquedas en Google y ha recibido casi 1.000 millones de aportaciones [13] en sus primeros 9 años de vida. Incluso cuando Microsoft cerró la enciclopedia Encarta a finales de 2009, indicó que "la categoría de enciclopedias tradicionales y el material de referencia ha cambiado", en una clara alusión a Wikipedia.

Sin embargo, algunas personas hacen uso de la facilidad de edición de los artículos para introducir palabras malsonantes o insultos ("vandalismos") y eliminar párrafos de contenido válido (blanqueos). Todas estas modificaciones no deseadas pueden deshacerse fácilmente utilizando el historial de cada artículo, pero es una tarea tediosa para un humano y durante el tiempo transcurrido desde el vandalismo hasta su corrección la página de Wikipedia afectada está en estado inconsistente. Conforme la visibilidad del proyecto ha ido creciendo, el problema del vandalismo se ha agravado, haciendo patente la necesidad de encontrar nuevos mecanismos para combatirlo.

De este modo se facilita que los usuarios comprometidos con el proyecto empleen su tiempo en incorporar nuevos contenidos, mejorar los existentes, tutorizar a nuevos miembros de la comunidad u otras tareas más productivas y gratificantes, en vez de velar por la seguridad de los artículos.

AVBOT [1] [4] acrónimo de *Anti-Vandalism BOT* (en español: "BOT¹ Anti-Vandalismo") ofrece una solución automática a muchos de estos vandalismos. Las aportaciones no deseadas son, la mayoría de las veces, detectables fácilmente para un ordenador, ya que lo modificado resalta por contener expresiones malsonantes o texto sin sentido. Este programa analiza las ediciones que se realizan en Wikipedia en español, buscando contribuciones maliciosas, y las deshace. AVBOT no podrá sustituir nunca a una persona, pues la capacidad de comprensión humana a la hora

AVBOT: detección y corrección de vandalismos en Wikipedia

Resumen: Wikipedia es un proyecto cuyo objetivo es construir una enciclopedia libre para hacer llegar todo el conocimiento humano a cada persona en su idioma. A día de hoy, puede decirse que va camino de cumplirlo, pues recientemente ha superado la cifra de 15 millones de artículos enciclopédicos en un total de más de 270 idiomas. Además, teniendo en cuenta a sus proyectos hermanos (Wikcionario, Wikilibros, Wikisource...) ha recibido casi 1.000 millones de aportaciones en estos 9 años y sirve más de 4.000 millones de páginas cada mes. La creación de una enciclopedia de forma colaborativa ha sido posible gracias al software MediaWiki, que permite a los visitantes modificar el contenido del sitio web con facilidad. Pero esto plantea un problema, ya que no todas las ediciones son beneficiosas para el proyecto. AVBOT es un robot creado para proteger Wikipedia en español de ciertas modificaciones no deseadas, conocidas como "vandalismos". Aunque en un principio AVBOT fue desarrollado para Wikipedia, es posible utilizarlo en cualquier sitio web que emplee el motor MediaWiki. En sus dos años de funcionamiento ha reparado más de 200.000 vandalismos y esta cifra crece a una media de 500 al día.

Palabras clave: MediaWiki, monitorización, robot, software libre, wiki, Wikipedia.

Autor

Emilio José Rodríguez Posada es Ingeniero Técnico en Informática de Sistemas, becario en la Oficina de Software Libre y Conocimiento Abierto de la Universidad de Cádiz y alumno colaborador del Departamento de Lenguajes y Sistemas Informáticos. Entusiasta del conocimiento libre y los wikis, es un usuario veterano de Wikipedia. Ha sido en este proyecto donde han tenido lugar la mayor parte de sus aportaciones al mundo del software libre, desarrollando *bots* que realizan tareas de mantenimiento de manera automática y rutinaria. El más destacado es sin duda AVBOT, un robot que repara vandalismos en Wikipedia en español, que resultó ganador en el III Concurso Universitario de Software Libre en la categoría "Mejor proyecto comunitario" y recibió el premio José Enrique Díaz Arozamena III Edición al mejor Proyecto Fin de Carrera Mención Informática en la Escuela Superior de Ingeniería. Ha impartido varias charlas y talleres sobre programación de bots, Wikipedia y edición en MediaWiki. A su vez, ha sido miembro del comité organizador de varias jornadas y encuentros sobre software libre y conocimiento abierto en la Universidad de Cádiz. En la actualidad está participando en el desarrollo de una herramienta para análisis estadístico de wikis MediaWiki, llamada StatMediaWiki.

de leer un texto sobrepasa con creces la de un ordenador, aunque sí es capaz de revertir gran cantidad de vandalismos, ahorrando mucho esfuerzo a la comunidad de Wikipedia. En sus dos años de funcionamiento ha reparado más de 200.000 vandalismos [3] y esta cifra crece a una media de 500 al día.

2. Desarrollo

AVBOT está programado en Python [7] y utiliza los paquetes *pywikipediabot* [8] y *python-irclib* [9]. Funciona en tres fases principales: captura de los cambios en artículos, análisis de las modificaciones y toma de decisiones.

En el primer paso obtiene una lista de los últimos cambios realizados en los artículos. Este "log" se publica en tiempo real a través de un canal de IRC [10] al que AVBOT permanece conectado 24 horas al día. A continuación, se descartan las modificaciones realizadas por bibliotecarios (administradores de Wikipedia), *bots* de mantenimiento y usuarios veteranos, ya que se presupone no dañan

rán los artículos. Las restantes ediciones se analizan para detectar posibles vandalismos.

Para cada modificación, el robot consulta el texto anterior y posterior al cambio, comparándolos. Entonces, un módulo de análisis comprueba la aparición de palabras malsonantes mediante una lista de expresiones regulares configurable. Cada expresión cuenta con un peso asociado, de modo que si la suma de los pesos provocados por una edición supera un determinado umbral, ésta se considera vandalismo y se revierte el artículo de Wikipedia a su estado anterior. Tanto las expresiones como los pesos han sido mejorados con la experiencia acumulada en los dos años de ejecución de AVBOT en Wikipedia en español.

Otro tipo de edición considerada maliciosa son los blanqueos, en los cuales un usuario elimina parcial o totalmente el contenido de un artículo. Esto es detectable calculando el porcentaje de texto retirado y revirtiendo si se supera cierta cifra. Este porcentaje de activa-

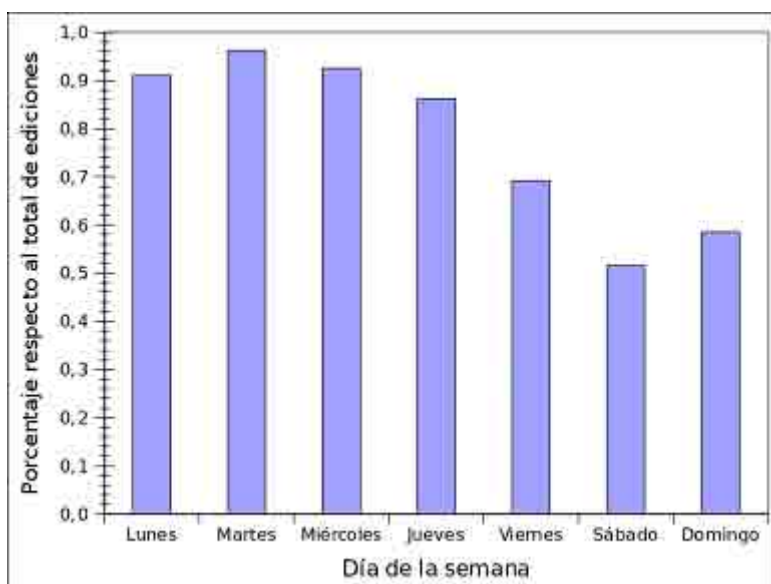


Figura 1. Porcentaje de intervenciones de AVBOT a lo largo de la semana, respecto al total del resto de usuarios.

ción varía en función del tamaño del artículo, y se ha estimado analizando los blanqueos realizados a lo largo de los años.

Por cuestiones de eficiencia, dada la enorme cantidad de ediciones que recibe Wikipedia en español (aproximadamente 30.000 al día), ha sido necesario realizar distintas optimizaciones al sistema, como el uso de hilos para paralelizar el trabajo y la precompilación de la lista de expresiones regulares.

Además, cada vez que AVBOT revierte un texto añadido por un usuario, el robot le envía un mensaje indicando que su edición ha sido deshecha y le solicita que se detenga. Si el usuario prosigue con su conducta disruptiva y comete tres infracciones seguidas en pocas horas, este hecho se notifica en un tablón conocido como "Wikipedia:Vandalismo en curso" [11]. Este tablón es vigilado con frecuencia por los administradores, que se encargarán de analizar la situación y actuar en consecuencia (generalmente bloqueando la cuenta de usuario reincidente o la dirección IP si son ediciones anónimas).

Por otro lado, dado que existe la posibilidad de error por parte del robot, que se ha estimado inferior al 0,5% (un error cada 200 acciones satisfactorias), el software detecta los falsos positivos vigilando si sus reversiones perduran. En el caso de que alguien haya eliminado una reparación del robot, éste emitirá un informe a una página vigilada por el controlador del bot, para su posterior diagnóstico. Esto permite ir afinando la capacidad de detección del programa.

Respecto a los errores de AVBOT, la mayoría están relacionados con palabras polisémicas. Un claro ejemplo lo constituye la expresión

"Patito feo", que contiene la palabra "feo", la cual es utilizada como ataque muy frecuentemente en biografías, aunque es correcta en este contexto concreto. Esto se soluciona insertando una expresión regular con contrapeso, que anula el peso negativo en las situaciones donde la expresión de que se trate es adecuada.

Por último, el programa incorpora otras funcionalidades, como una lista de exclusiones compuesta por páginas que el robot no vigilará. Por ejemplo, páginas de discusión o de mantenimiento interno de Wikipedia, debido a que en ellas está tolerado el uso de lenguaje coloquial, que difiere del lenguaje enciclopédico de los artículos. Además, el programa revisa la creación de páginas (en Wikipedia en español se crean unos 400 artículos nuevos cada día), marcando para su eliminación aquellas que contienen ediciones de prueba de anónimos, careciendo completamente de contenido útil que amplíe la enciclopedia.

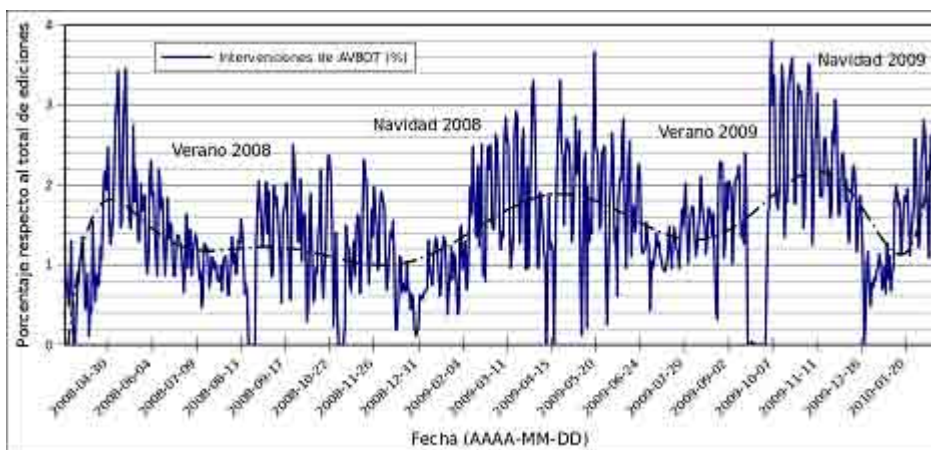


Figura 2. Porcentaje de intervenciones de AVBOT desde que entró en funcionamiento, respecto al total del resto de usuarios.

3. Análisis de resultados

Los resultados de AVBOT son muy satisfactorios, produciéndose un falso positivo cada 200 reversiones correctas o más. La precompilación de la lista de expresiones regulares permite que el tiempo total consumido (suma de realizar el análisis y tomar la decisión sobre si revertir o no) esté por debajo del milisegundo, por lo que el único retraso perceptible que existe es la comunicación a través de Internet, entre servidores.

Por ello, el tiempo transcurrido entre el acto vandálico y la restauración del artículo ronda los 5 segundos, haciendo que el vandalismo sea prácticamente invisible (a menos que algún lector consulte la página justo en ese periodo).

Un análisis detallado de la distribución de los vandalismos en el tiempo permite ver que éstos se concentran de lunes a jueves (ver figura 1), y durante la tarde europea, pues en esa franja horaria coinciden los vandalismos provenientes de España y América del Sur. Por otro lado, se observan caídas en periodos vacacionales, como el verano, Navidad o fin de año, no solo en términos totales sino también porcentuales (ver figura 2).

4. Comunidad

Para obtener permiso de la comunidad para ejecutar AVBOT sobre Wikipedia en español, se hizo una votación en la que hubo 19 votos a favor y 0 en contra [12].

Desde los comienzos de AVBOT, los usuarios de Wikipedia en español han sugerido mejoras y solicitado que fueran incorporadas nuevas funcionalidades al programa. Quizás, el aspecto en el que más directamente han colaborado algunos usuarios ha sido en la ampliación de la lista de patrones que detectan los vandalismos.

Además, dado que es software libre y su código se encuentra disponible en una forja

de acceso público [2], algunos usuarios han decidido descargarse y ejecutar una copia de respaldo de AVBOT, por si el servidor donde se ejecuta la instancia principal deja de funcionar momentáneamente. Esto permite que Wikipedia en español siempre esté protegida de manera redundante.

La evolución del proyecto ha sido narrada en gran medida en el blog oficial [1], que también sirvió como histórico de cambios para la participación del proyecto en la III edición del Concurso Universitario de Software Libre. A lo largo de la evolución del programa, se ha hecho especial énfasis en dotar de diversas opciones de configuración al robot, para que pueda ser empleado sin muchos conocimientos de programación, y utilizado en las plataformas MediaWiki de otras comunidades. En este sentido, recientemente desde Wikipedia en portugués se han interesado en ejecutar AVBOT sobre sus contenidos y se planea su expansión a más idiomas de Wikipedia.

De cara al futuro, las líneas de colaboración más inmediatas con el proyecto que proponemos a los lectores de este artículo serían ayudar en la traducción y adaptación del sistema a otros idiomas, crear listas de expresiones y pesos adecuadas a otras Wikipedias y traducir la documentación existente. A este respecto, el proyecto está abierto a la participación de todo aquel que lo desee.

5. Conclusiones y trabajos futuros

Conforme la popularidad de Wikipedia ha crecido, el problema del vandalismo se ha hecho más grave. AVBOT lleva funcionando 2 años en la versión en español de Wikipedia,

y en todo este tiempo ha reparado más de 200.000 vandalismos [3], evitando una gran cantidad de horas de trabajo a la comunidad. El mundo de los wikis evoluciona con rapidez, por lo que será necesario ir adaptándose a los cambios, añadiendo nuevas funcionalidades al sistema y mejorando las existentes. Un campo por explorar sería el del aprendizaje automático del robot, entrenándolo con los vandalismos revertidos por humanos. Por otro lado, dado que existe interés por utilizarlo en otros idiomas de Wikipedia y otros sitios web que emplean MediaWiki, es necesario continuar con la internacionalización del código y mejorar, si cabe, la versatilidad de éste.

Otra línea de trabajo futuro es el desarrollo de herramientas de análisis de wikis para fines docentes. En la actualidad se están llevando a cabo diversos proyectos educativos en la Universidad de Cádiz en los que los alumnos crean conocimiento libre en clase [15]. Por ejemplo, con el apoyo de su Oficina de Software Libre y Conocimiento Abierto se lanzó WikiHaskell, un wiki para la asignatura "Programación Funcional" de Ingeniería Informática (en el que los alumnos generan documentación libre sobre bibliotecas del lenguaje de programación Haskell), o los proyectos fin de carrera WikiUNIX y Wikijuegos [16]. Estos proyectos suelen tener la modificación de páginas restringida a los alumnos que trabajan en ellos en sus primeras fases, pero una vez está terminada la experiencia docente se abren a cualquier usuario de Internet, por lo que el uso de AVBOT les ayudaría a protegerse del vandalismo. Además, para la evaluación del trabajo de los alumnos en los

wikis se ha desarrollado la herramienta StatMediaWiki [14], en la que el autor de este artículo participa como principal desarrollador y cuya aplicación se extiende al análisis estadístico de cualquier wiki en general.



Finalistas del III Concurso Universitario de Software Libre: de izquierda a derecha, en la primera fila: Francisco Salido Ruiz, Andoni Morales Alastruey, Luis Antonio González Jaime y Daniel García Moreno; en la segunda fila: José Francisco Lupion González, Emilio José Rodríguez Posada, Jose Manuel Cordeiro Rodríguez, Ricardo Juan Palma Durán y David Castellanos Serrano.

Referencias

- [1] AVBOT. Donde descargar el software, documentación y estar al tanto de los últimos avances <<http://avbot.blogspot.com>>.
- [2] Forja de AVBOT. Para seguir su desarrollo <<http://code.google.com/p/avbot>>.
- [3] Demostración de AVBOT. Se puede ver el bot en funcionamiento las 24 horas del día <<http://es.wikipedia.org/wiki/Especial:Contribuciones/AVBOT>>.
- [4] Emilio José Rodríguez Posada. *Detector y corrector automático de ediciones maliciosas en Wikipedia*. Memoria del proyecto fin de carrera <<http://emijrp.googlepages.com/mpfc.pdf>>.
- [5] Wikipedia. El proyecto de enciclopedia libre <<http://www.wikipedia.org>>.
- [6] MediaWiki. Software Wiki para sitios web <<http://www.mediawiki.org>>.
- [7] Python. <<http://www.python.org>>.
- [8] pywikipediabot. Framework para el desarrollo de bots para sitios MediaWiki <<http://pywikipediabot.sourceforge.net>>.
- [9] python-irc-lib. Framework para IRC en Python <<http://python-irc-lib.sourceforge.net>>.
- [10] Canal de IRC de cambios recientes de Wikipedia en español. <<http://irc.wikimedia.org/es.wikipedia>>.
- [11] Wikipedia. Tablón con los últimos informes de vandalismo reincidente <http://es.wikipedia.org/wiki/Wikipedia:Vandalismo_en_curso>.
- [12] Votación sobre AVBOT. <http://es.wikipedia.org/wiki/Wikipedia:Bot/Autorizaciones/Archivo_2008#AVBOT>.
- [13] Emilio José Rodríguez Posada. Contador de ediciones de los proyectos Wikimedia <<http://toolsserver.org/~emijrp/wikimediacounter>>.
- [14] StatMediaWiki. Herramienta de análisis estadístico de wikis MediaWiki <<http://statmediawiki.forja.rediris.es>>.
- [15] Manuel Palomo, Inmaculada Medina, Emilio José Rodríguez, Noelia Sales. Tecnologías wiki y conocimiento abierto en la Universidad. *Actas de la V Conferencia Internacional en Software Libre, Cáceres 2009* pp. 16-19, ISBN 978-84-692-8739-2.
- [16] Wikis @ OSLUCA. Algunos de los wikis apoyados por la Oficina de Software Libre y Conocimiento Abierto de la Universidad de Cádiz <<http://osl.uca.es/wikis>>.

Nota

¹ Un bot, en el mundo de los wikis, es un programa que realiza tareas repetitivas de manera automática sobre los contenidos del sitio web.