

Analyzing learners' activity beyond the MOOC

Diego Sapunar-Opazo¹, Ronald Pérez-Álvarez¹, Jorge Maldonado-Mahauad¹, Carlos Alario-Hoyos², Mar Pérez-Sanagustín¹

¹ Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 486, Santiago, Chile
{dasapunar, jjmaldonado, raperez13, mar.perez}@uc.cl

² Universidad Carlos III de Madrid, Avda. de la Universidad, 30, Leganés, España
calario@it.uc3m.es

Abstract. Research on help seeking in MOOCs has mainly focused on analyzing learners' traces within the course forum, or in external social tools which are directly associated to the course. However, little research has been done on the external supplementary websites and digital resources that learners consult outside the MOOC as a way of help seeking. In this working paper, we present the results of an exploratory study with 61 learners from 3 MOOCs in which we analyzed what type of information learners visit outside the MOOC during their study sessions. The results show that learners spent 2% of the time in their study sessions outside the MOOC, being social networking sites, search engines and sites related to the course content the most visited.

Keywords: Learning Analytic, Massive Open Online Courses, MOOCs, Exploratory study.

1 Introducción

De acuerdo con la bibliografía de autorregulación de los últimos 30 años, saber buscar ayuda cuando lo necesitas es una de las estrategias más importantes para lograr sus objetivos de aprendizaje [6] [11]. Esta ayuda puede provenir tanto de otras personas, como también de fuentes de información (búsqueda de información). Debido a la falta de guía por parte de un profesor en los Cursos Masivos en Línea (del inglés *Massive Open Online Courses*), la habilidad de buscar ayuda por parte del estudiante para enfrentar dificultades y lograr los objetivos de aprendizaje es crítica [8].

Investigadores han estudiado la búsqueda de ayuda por parte de los alumnos de MOOCs mediante dos perspectivas: (1) búsqueda de ayuda desde otras personas, y (2) búsqueda de ayuda desde fuentes de información. Respecto a la primera perspectiva, hay estudios que se centran en analizar las interacciones entre los distintos estudiantes dentro del MOOC, generalmente mediante el foro de discusión del curso. Por ejemplo, los autores en [5], proponen diferentes métodos para investigar el intercambio de conocimiento que ocurre en los foros de discusión de un MOOC de Coursera, con el objetivo de ver cómo la estructura de comunicación va cambiando con el transcurso del tiempo. Los autores en [14], analizaron quiénes son los estudiantes más influyentes en los foros

de discusión de un Curso Masivo en Línea (MOOC), y como sus intervenciones producen un efecto positivo en la discusión. También existen trabajos que se centran en el uso de herramientas sociales externas al curso utilizadas por los estudiantes. Por ejemplo, autores en [1] y en [2], compararon el uso de las herramientas sociales internas y externas de un Curso Masivo en Línea (MOOC), concluyendo que el foro de discusión interno del curso es el preferido por los estudiantes para buscar ayuda. Autores en [13] estudiaron el uso de Twitter en dos Cursos Masivos en Línea (MOOCs), concluyendo que es una valiosa herramienta social para enfrentar preguntas y respuestas relacionadas con el contenido del curso. Esta conclusión la corrobora otro estudio cuyos resultados indican la importancia de las conexiones digitales que ocurren en los recursos sociales externos al curso [12]. Sin embargo, uno de los problemas que tiene el utilizar datos de herramientas sociales externas al curso es la identificación de los estudiantes; autores en [3] indican que solamente lograron identificar al 5% de los estudiantes de 18 Cursos Masivos en Línea (MOOCs) y las 5 plataformas digitales sociales más populares.

Con el fin de lidiar con esta limitación, otros autores se han enfocado en el estudio de la búsqueda de ayuda en Cursos Masivos en Línea (MOOCs) usando otras fuentes de información. La mayoría de estos estudios se centran en el análisis de las interacciones de los estudiantes con el contenido del curso combinada con información auto-reportada en el proceso de autorregulación del estudiante, para poder extraer patrones de comportamiento y relacionarlos con el desempeño del estudiante. Por ejemplo, autores en [8] muestran que personas con perfiles de baja y alta autorregulación del aprendizaje se comportan de manera distinta en Cursos Masivos en Línea (MOOCs), y que la motivación de los estudiantes influye en cómo estos interactúan con el contenido del curso. Además, en [7] identificaron que las estrategias de autorregulación del aprendizaje, tales como el establecimiento de objetivos claros y la planificación estratégica, predicen un logro en los objetivos personales en el curso, mientras que la búsqueda de ayuda (desde personas) se asocia a un menor logro.

Por lo tanto, las investigaciones plantean dos problemáticas relevantes: (1) la importancia de analizar como los estudiantes de Curso Masivo en Línea (MOOC) usan herramientas sociales externas al curso para buscar ayuda con otros estudiantes; y (2) la importancia de analizar como los estudiantes de Curso Masivo en Línea (MOOC) complementan el contenido del curso con información externa a este.

Este artículo se centra en la segunda problemática para ayudar a entender qué ocurre más allá del Curso Masivo en Línea (MOOC) cuando los estudiantes buscan complementar sus sesiones de estudio con información externa a este. Hay una necesidad de expandir el conocimiento acerca de la búsqueda de ayuda, y estudiarla como un proceso que no solo ocurre a través de personas en herramientas sociales, sino un proceso que ocurre en toda la Internet. En este contexto surgen dos preguntas de investigación: (RQ1) ¿Cómo invierten su tiempo los estudiantes de Curso Masivo en Línea (MOOC) (dentro y fuera del curso)? Y (RQ2) ¿Qué recursos adicionales están consultando los estudiantes fuera del Curso Masivo en Línea (MOOC)?

A través de estas dos preguntas de investigación, este artículo contribuye a ampliar el conocimiento acerca de las estrategias de búsqueda de ayuda en los Cursos Masivos en Línea (MOOCs). Para ello, se presentan los resultados de un estudio exploratorio

que analiza el comportamiento de 61 estudiantes fuera del MOOC durante sus sesiones de estudio en el curso.

2 Estudio Exploratorio

2.1 Contexto

El año 2016 la Pontificia Universidad Católica de Chile desarrollo la aplicación web llamada *NoteMyProgress* [10]. Esta herramienta ha sido diseñada para apoyar la autorregulación de los estudiantes en Cursos Masivos en Línea (MOOCs). *NoteMyProgress* provee a los estudiantes la funcionalidad de tomar notas y una serie de visualizaciones interactivas que muestran información acerca del desempeño del estudiante en el curso, tales como tiempo invertido dentro y fuera del curso, o el tiempo invertido en cada una de las actividades del curso. Esta aplicación está compuesta por dos partes principales: (1) un *plugin* de Google Chrome que se activa automáticamente cuando algún usuario registrado comienza una sesión de estudio en Coursera, (2) y una aplicación web que siempre está disponible mediante un *plugin* de Google Chrome. El *plugin* también ofrece una visualización del tiempo invertido dentro y fuera del curso, y la funcionalidad de tomar notas. Mientras que la aplicación web presenta al estudiante información detallada acerca de sus sesiones de estudio. *NoteMyProgress* no sólo recopila la información de las actividades que hace el estudiante dentro del curso, sino que también las direcciones web (URLs) externas y sitios webs que los estudiantes visitan durante sus sesiones de estudio. Es importante mencionar que una sesión de estudio es un periodo de tiempo, el cual comienza al registrarse actividad continua en Coursera, con intervalos de inactividad no más largos que 45 minutos. Esta definición de sesión de estudio ya ha sido utilizada en estudios anteriores [9].

Actualmente, *NoteMyProgress* solo funciona en la plataforma Coursera, sin embargo, será adaptada a cualquier otra plataforma MOOC, o cualquier otro sistema de administración de aprendizaje. *NoteMyProgress* fue lanzado durante 2 meses y 4 días, desde abril a mayo del año 2017, durante dos ediciones distintas de tres MOOCs: Uno correspondiente a educación (“Aula Constructivista” – Curso 1), otro en administración de organizaciones (“Gestión de organizaciones Efectivas” – Curso 2) y en electrónica (“Electrones en Acción – Curso 3). Por responsabilidad ética y aspectos de privacidad de datos, la instalación y uso de esta aplicación era voluntaria, y cada usuario tenía que aceptar los términos de condiciones de uso, donde se indicaba qué tipo de información se recopila y que su uso es únicamente de propósito académico.

2.2 Muestra

Durante los 2 meses del estudio exploratorio se registraron un total de 816 estudiantes activos en los tres MOOCs. Desde esta muestra, 85 estudiantes descargaron *NoteMyProgress*, de los cuales 61 (N=61) lo usaron al menos una vez. Ese grupo de 61 alumnos es el universo de este estudio exploratorio. La Tabla 1 muestra el número de

estudiantes activos por curso (quienes completaron al menos una evaluación), el número de personas que usaron *NoteMyProgress*, y también el número de usuarios que utilizaron *NoteMyProgress* y aprobaron el curso.

Table 1. Universo de estudio (N=61). NMP corresponde a NoteMyProgress

Número del curso	Estudiantes activos	Estudiantes usando NMP	Aprobados usando NMP
1	107	11	2
2	237	19	11
3	472	31	11
Total	816	61	21

2.3 Metodología

Para el análisis se utilizaron dos fuentes de información: (1) los *logs files* recolectados por *NoteMyProgress*, los cuales contienen información acerca de los a URLs accedidos por el estudiante durante sus sesiones de estudio; y (2) los *logs files* de Coursera de aquellos estudiantes utilizando *NoteMyProgress*, los cuales contienen información de las interacciones de los estudiantes con el contenido del curso.

Se siguió un proceso de 3 pasos para responder las preguntas de investigación. Primero, se trabajaron los *logs files* de acuerdo a nuestra unidad básica de análisis: sesión de estudio. Luego se creó un log file único mezclando los *logs files* mencionados anteriormente: Coursera y *NoteMyProgress*. El log file resultante y que se usó para el análisis contiene la siguiente información: ID del usuario, ID de la sesión, Duración de la sesión en segundos, Completado (Aprobado / Desaprobado), Nota obtenida en el curso, Contenido (correspondiente al recurso de Coursera o al URL consultado fuera del MOOC) y el tiempo invertido.

En segundo lugar, para responder la RQ1 sobre cómo los estudiantes de MOOCs están invirtiendo su tiempo en las sesiones de estudio, se analizaron las sesiones de estudio de los estudiantes. Por cada sesión de estudio se calculó el porcentaje de tiempo invertido por el estudiante dentro y fuera del curso. Se repitió este cálculo para cada curso y para todos los cursos juntos.

Finalmente, para responder la segunda pregunta de investigación RQ2 acerca de que recursos adicionales están consultando los estudiantes fuera del contenido del MOOC, se categorizaron los diferentes URLs accedidos por los estudiantes fuera del MOOC utilizando el servicio entregado por Cyren¹. Para clasificar los diferentes URLs, se utilizó el nombre del sitio web (*hostname*) como unidad básica de medida. Por ejemplo, si un URL era: “www.facebook.com” y el usuario navegaba a un enlace con el nombre: “www.facebook.com/profile”, ambos son contados como “www.facebook.com” y se cuenta como si hubiera sido consultado dos veces.

Se identificaron un total de 31 categorías, siendo 8 de ellas las más comunes (aproximadamente el 99% de las consultas correspondió a estas 8 categorías): (1) Redes So-

¹ Página web: <https://www.cyren.com/security-center/url-category-check>

ciales, sitios como Facebook, Twitter, etc.; (2) Educación, sitios relacionados con instituciones educativas, como la sitios oficiales de las universidades; (3) Portales y Motores de Búsqueda, como Google o Yahoo!; (4) Página web basada en correo electrónico, tal como Gmail; (5) Entretenimiento, como Google Play o Miniclip; (6) Computadores y Tecnología, sitios web relacionados con componentes tecnológicos o computación, como Arduino; (7) Gobierno, correspondiente a sitios gubernamentales; y (8) Negocios, correspondiente a sitios webs de negocios. El resto de los URLs, corresponden al 1% de los sitios webs consultados por los estudiantes, los cuales se clasificaron como “Otros”. Esta categoría incluye sitios webs tales como sitios personales, política, compras, noticias o servicios de traducción. Después de clasificar los URLs, se realizaron dos análisis. Primero, se calculó la frecuencia de cada una de las categorías, esto quiere decir el número de veces que una persona accede un URL por categoría, luego, se calculó el porcentaje de la frecuencia en que esta categoría fue accedida para detectar las categorías más frecuentes. En segundo lugar, se calculó el tiempo invertido por los alumnos en cada uno de los URLs. Esto permite calcular el porcentaje del tiempo invertido por parte de los estudiantes en cada una de las categorías.

3 Resultados

Respecto al **tiempo invertido por los estudiantes fuera y dentro del curso (RQ1)**, la Tabla 2 muestra que los estudiantes invierten el 98,35% de su tiempo dentro del MOOC, y 1,65% fuera. Analizando los resultados curso por curso, se observa que: en el Curso 1, 98,45% del tiempo fue invertido dentro del curso, mientras 1,57% del tiempo fuera de este; para el Curso 2 se observa que el 99,15% del tiempo fue invertido dentro del curso y 0,85% del tiempo fuera; para el Curso 3, se observa que el 97,46% del tiempo fue invertido dentro del curso y 2,54% del tiempo fuera del curso. Debido a la naturaleza del experimento, invertir menos tiempo fuera del MOOC es coherente, ya que *NoteMyProgress* sólo captura la actividad del estudiante durante sesiones de estudio, cuando supuestamente el estudiante debería estar trabajando en el contenido del curso.

Table 2. Porcentaje del tiempo invertido por estudiantes dentro y fuera del curso.

Curso	Dentro del Curso	Fuera del Curso
1	98,45%	1,57%
2	99,15%	0,85%
3	97,46%	2,54%
Media	98,35%	1,65%

Respecto a **qué recursos digitales adicionales están consultando los estudiantes fuera del curso (RQ2)**, la Figura 1 muestra una lista de sitios consultados fuera del curso por parte de los alumnos en sus sesiones de estudio (~2% del tiempo total). En esta figura se observa que las categorías más frecuentes (Educación y Portales y Motores de Búsqueda) son las mismas para los tres cursos. La frecuencia del resto de las categorías varía dependiendo del curso.

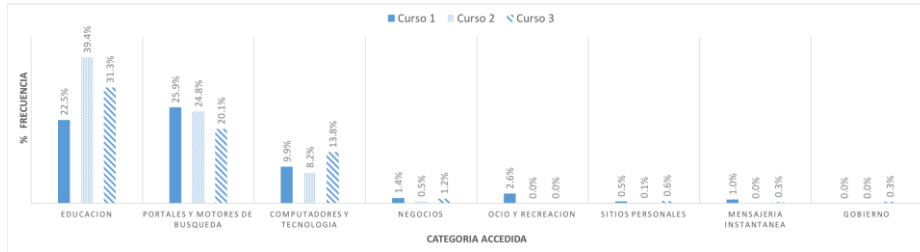


Fig. 1. Dentro del ~2% del tiempo invertido por los estudiantes en recursos externos al curso, el porcentaje de frecuencia de los sitios visitado en cada categoría.

Sin embargo, analizando el tiempo invertido en cada una de las categorías, el resultado es distinto. Como se puede observar en la Figura 2, en la mayoría de los casos, existe una relación directa entre las categorías con mayor tiempo invertido por parte de los estudiantes y la temática del curso. Esto se refleja claramente en el tercer curso, donde la categoría con mayor tiempo invertido es Computadores y Tecnología.

También se observa que en la Figura 2 que una gran parte del tiempo se invierte en Portales y Motores de Búsqueda. Esto puede condicionar los resultados obtenidos, sin embargo, por limitaciones del procedimiento utilizado, se consideran los motores de búsqueda como un recurso más.

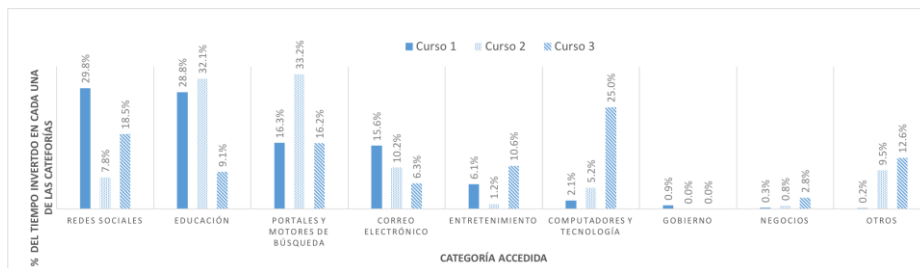


Fig. 2. Dentro del ~2% del tiempo invertido por los estudiantes en recursos externos al curso, el porcentaje del tiempo invertido por los estudiantes en cada categoría.

4 Conclusiones y Trabajos Futuros

Este artículo presenta un trabajo en progreso en el cual se realizó un estudio exploratorio para analizar qué tipo de información accedieron 61 estudiantes fuera del contenido del MOOC en un proceso de búsqueda de ayuda. Este estudio tiene varias limitaciones que se espera mejorar en trabajos futuros. Por un lado, el número de participantes del experimento es poco como para extraer conclusiones generalizables. Actualmente, se están realizando nuevos experimentos con una muestra mayor. Por otro lado, la metodología de análisis de los datos es solamente exploratorio. Próximamente, se espera realizar un análisis más sofisticado, con el objetivo de encontrar una relación entre el tipo de información que consulta un estudiante y su desempeño en el curso.

A pesar de las limitaciones del trabajo, este artículo presenta una aproximación innovadora para estudiar cómo los estudiantes de un MOOC gestionan sus sesiones de trabajo con recursos que van más allá de los ofrecidos directamente por el curso. Este tipo de información podría utilizarse en futuro para, por ejemplo, detectar cuándo un estudiante está copiando, o para desarrollar un sistema recomendador de sitios externos para MOOCs.

5 Agradecimientos

Este trabajo es apoyado por FONDECYT (Chile) bajo el proyecto N 11150231, el proyecto MOOC-Maker (561533-EPP-1-2015-1-ES-EPPKA2-CBHE-JP), el Proyecto LALA (586120-EPP-1-2017-1-ES-EPPKA2-CBHE-JP), Universidad de Costa Rica, la Comisión Nacional de Investigación Científica – CONICYT/ DOCTORADO NACIONAL 2016/21160081, CONICYT/ DOCTORADO NACIONAL 2017/21170467

Referencias

1. Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., Muñoz-Organero, M., & Rodríguez-de-las-Heras, A.: Analyzing the impact of built-in and external social tools in a MOOC on educational technologies. In European Conference on Technology Enhanced Learning, pp. 5-18. Springer, Berlin, Heidelberg (2013).
2. Alario-Hoyos, C., Pérez-Sanagustín, M., Delgado-Kloos, C., & Munoz-Organero, M.: Delving into participants' profiles and use of social tools in MOOCs. *IEEE Transactions on Learning Technologies*, 7(3), 260-266 (2014).
3. Chen, G., Davis, D., Lin, J., Hauff, C., & Houben, G. J.: Beyond the MOOC platform: gaining insights about learners from the social web. In Proceedings of the 8th ACM Conference on Web Science, pp. 15-24. ACM, Hannover, Germany, (2016).
4. Hecking, T., Hoppe, H. U., & Harrer, A.: Uncovering the structure of knowledge exchange in a MOOC discussion forum. In Advances in Social Networks Analysis and Mining (ASONAM), 2015 IEEE/ACM International Conference, pp. 1614-1615. IEEE (2015).
5. Karabenick, S. A. (Ed.): Strategic help seeking: Implications for learning and teaching. Routledge (1998).
6. Karabenick, S. A., & Newman, R. S. (Eds.): Help seeking in academic settings: Goals, groups, and contexts. Routledge (2013).
7. Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J.: Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & education*, 104, 18-33 (2017).
8. Littlejohn, A., Hood, N., Milligan, C., & Mustain, P.: Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Internet and Higher Education*, 29, 40-48 (2016).
9. Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M.: Penetrating the black box of time- on-task estimation. In International Conference on Learning Analytics and Knowledge, pp. 184– 193. ACM, NY, USA (2015).
10. Pérez-Álvarez, R., Maldonado-Mahauad, J. J., Sapunar-Opazo, D., & Pérez-Sanagustín, M.: NoteMyProgress: A Tool to Support Learners' Self-Regulated Learning Strategies in

- MOOC Environments. In *European Conference on Technology Enhanced Learning*, pp. 460-466. Springer, Cham (2017).
11. Schunk, D. H., & Zimmerman, B. (Eds.): *Handbook of self-regulation of learning and performance*. Taylor & Francis (2011).
 12. van Treeck, T., & Ebner, M.: How useful is twitter for learning in massive communities? An analysis of two MOOCs. *Twitter & Society*, pp. 411-424, (2013).
 13. Veletsianos, G., Collier, A., & Schneider, E.: Digging deeper into learners' experiences in MOOCs: Participation in social networks outside of MOOCs, notetaking and contexts surrounding content consumption. *British Journal of Educational Technology*, 46(3), 570-587 (2015).
 14. Wong, J. S., Pursel, B., Divinsky, A., & Jansen, B. J.: An analysis of MOOC discussion forum interactions from the most active users. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, pp. 452-457. Springer, Cham (2015).
 15. Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., & Muñoz-Gama, J.: Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses. *Computers in Human Behavior*, 80, 179-196 (2018).