

Contribuciones a la identificación de patrones de navegación de estudiantes en
cursos de Open edX



FABIÁN ANDRÉS ANACONA ANACONA

Informe
Maestría en Ingeniería Telemática

Director:
Ph.D. Mario Fernando Solarte Sarasty

Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Telemática
Maestría en Ingeniería Telemática
Línea de investigación: Aplicaciones y Servicios sobre Internet
Popayán, abril de 2020

FABIÁN ANDRÉS ANACONA ANACONA

Contribuciones a la identificación de patrones de navegación de estudiantes en
cursos de Open edX

Tesis presentada a la Facultad de Ingeniería
Electrónica y Telecomunicaciones de la
Universidad del Cauca, para la obtención del
Título de

Magíster en:
Ingeniería Telemática

Director:
Ph.D. Mario Fernando Solarte Sarasty

Popayán
2020

Página de Aceptación

Ph.D. Mario Fernando Solarte Sarasty
Director-Tutor

Presidente del Jurado

Jurado

Jurado

Popayán, abril de 2020

*A todos los que contribuyeron para alcanzar este gran objetivo. Mi eterna gratitud.
Fabián.*

Resumen Estructurado

Los Cursos en Línea Abiertos y Masivos -MOOC, permiten la participación de un gran número de estudiantes, generando de esta forma una gran cantidad de información de la interacción de los estudiantes con el curso. En el presente trabajo de investigación se analizan cuatro conjuntos de datos, archivos Excel, que contienen la interacción de los estudiantes, administradores, profesor y monitor, con el curso Astronomía Cotidiana, ofertado en la Universidad del Cauca en el primer periodo académico de 2017, el curso MPOC fue soportado por una instancia de la plataforma de Open edX.

Los conjuntos de datos, se limpian y adecuan, además se procesan con el modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento. Con el algoritmo LexRank con Umbral se procesa cada conjunto de datos, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad del usuario, y la identificación de la matriz estacionaria. Con los resultados de la matriz estacionaria se realiza, la grafica de la centralidad del patrón de navegación EF-ISF por sesión del usuario.

Se desarrolla análisis e interpretación de los resultados, en el administrador, profesor, monitor y en tres estudiantes con numero de sesiones bajas, media y alta, debido al gran número de estudiantes por conjuntos de datos. Por último, se describen las conclusiones del presente trabajo de investigación y el trabajo futuro.

Palabras claves: conjuntos de datos, estudiante, usuario, archivo Excel, sesión, LexRank con Umbral, patrones de navegación EF-ISF por sesión, centralidad, MPOC.

Structured Summary

The Massive Online Open Courses -MOOC, allow the participation of a large number of students, thus generating a large amount of information on the interaction of students with the course. In the present research work four sets of data are analyzed, Excel files, which contain the interaction of students, administrators, professor and monitor, with the course Astronomía Cotidiana, offered at the University of Cauca in the first academic period of 2017, the MPOC course was supported by an instance of the Open edX platform.

Data sets are cleaned and adjusted, and processed with the vector space model and weighting based on the relative frequency of an event. The LexRank With Threshold algorithm, each data set is processed to identify the EF-ISF navigation patterns by the user's most central session, and to identify the stationary matrix. With the results of the stationary matrix, the graph of the centrality of the EF-ISF navigation pattern by user session is made.

Analysis and interpretation of the results is developed in the administrator, teacher, monitor and in three students with low, medium and high number of sessions, due to the large number of students per data set. Finally, the conclusions of the present research work and the future work are described.

Keywords: data sets, student, user, Excel file, session, LexRank with Threshold, EF-ISF navigation patterns per session, centrality, MPOC.

Tabla de contenido

Resumen Estructurado.....	iii
Structured Summary.....	iv
Lista de Tablas.....	viii
Lista de Figuras.....	xi
CAPITULO 1	1
1 INTRODUCCIÓN.....	1
1.1 Planteamiento del Problema	1
1.2 Objetivos	3
1.2.1 Objetivo General.....	3
1.2.2 Objetivos específicos.....	3
1.3 Resultados Obtenidos	4
CAPITULO 2	6
2 MARCO TEÓRICO	6
2.1 Open edX.....	6
2.1.1 Archivo Tracking.log	8
2.2 Patrones de estudiantes en MOOC.....	11
2.3 Navegación de estudiantes en MOOC	14
2.4 Patrones de Navegación de los estudiantes en MOOC	16
2.5 Generación Automática de Resúmenes de Texto	19
2.5.1 Definición.....	19
2.6 Representación de los Documentos.....	19
2.6.1 Modelo de Espacio Vectorial	19
2.6.2 Técnicas de ponderación de términos	20
2.6.3 Medidas de Similitud.....	20
2.6.4 Representación de Documentos por Medio de Matrices	21
2.7 Algoritmo LexRank con Umbral para generación automática de resúmenes de múltiples documentos	22
2.8 Conclusiones del Marco Teórico	26
CAPITULO 3	28

3	APROXIMACIÓN ARQUITECTÓNICA PARA LA IDENTIFICACIÓN DE LOS PATRONES DE NAVEGACIÓN EF-ISF POR SESIÓN DE MAYOR CENTRALIDAD DE LOS USUARIOS.....	28
3.1	Conjuntos de Datos.....	29
3.2	Pre procesamiento	32
3.2.1	Conjuntos de Datos Grupo A.....	32
3.2.2	Conjuntos de Datos Grupo B.....	36
3.2.3	Archivo de texto plano	40
3.2.4	Representación de Usuarios.....	41
3.2.5	Ponderación basada en la frecuencia relativa de un evento.....	41
3.2.6	Medida de Similitud	42
3.2.7	Representación del conjunto de datos por medio de Matrices	42
3.3	Algoritmo	43
3.3.1	Algoritmo LexRank con Umbral	43
3.3.2	Patrón de Navegación EF-ISF por sesión de mayor centralidad	47
	CAPITULO 4	48
4	PROCESO PARA LA IDENTIFICACIÓN DE LOS PATRONES DE NAVEGACIÓN EF-ISF POR SESIÓN DE MAYOR CENTRALIDAD EN CONJUNTOS DE DATOS.....	48
4.1	Parámetros del algoritmo LexRank con Umbral	48
4.2	Paso a paso para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios	49
4.3	Grafica de la centralidad del patrón de navegación EF-ISF por sesión del usuario	50
	CAPITULO 5	52
5	RESULTADOS Y ANÁLISIS	52
5.1	Astronomía Cotidiana 2017-I Grupo A	52
5.1.1	Estudiantes	53
5.1.2	AdministradorGA	61
5.1.3	ProfesorGA.....	62
5.1.4	Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5	65
5.2	Astronomía Cotidiana 2017-I Grupo B	70

5.2.1	Estudiantes	71
5.2.2	Administrador.....	78
5.2.3	ProfesorGB.....	81
5.2.4	MonitorGB.....	83
5.2.5	Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5	86
5.3	Conclusiones.....	93
CAPITULO 6		97
6	CONCLUSIONES Y TRABAJO FUTURO	97
6.1	Conclusiones.....	97
6.2	Trabajo Futuro.....	98
REFERENCIAS		100

Lista de Tablas

Tabla 1. Descripción de campos. Tomado de (Palta & Vásquez, 2016).	10
Tabla 2. Campos del evento context. Tomado de (Palta & Vásquez, 2016).	11
Tabla 3. Sub campos adicionales del evento context. Tomado de (Palta & Vásquez, 2016).	11
Tabla 4. Evaluación con ROUGE de los algoritmos con DUC2005. Tomado de (Anacona et al., 2015).	24
Tabla 5. Mejora en % de LexRank con Umbral (ver Tabla 4). Tomado de (Anacona et al., 2015).	24
Tabla 6. Resultados de ROUGE de los algoritmos con DUC2005. Adaptado de (Alguliyev et al., 2015).	25
Tabla 7. Descripción de los Conjuntos de Datos. Adatado de (Jaramillo & Mario, 2017).	29
Tabla 8. Descripción de campos de la primera fila de los conjuntos de datos. Adatado de (Jaramillo & Mario, 2017).	30
Tabla 9. Descripción de actividades. Adatado de (Jaramillo & Mario, 2017).	30
Tabla 10. Descripción de los Conjuntos de Datos Grupo A. Adatado de (Jaramillo & Mario, 2017).	33
Tabla 11. Codificación de Caracteres Grupo A. Adaptado de (Merino, 2012)	33
Tabla 12. Descripción de Identificadores de YouTube de los Cursos. Fuente Propia.	34
Tabla 13. Registros eliminados del Excel Grupo A. Fuente propia.	34
Tabla 14. Descripción del conjunto de datos del Grupo A. Fuente Propia.	35
Tabla 15. Descripción del conjunto de datos del Grupo A U1T1 a U1T5. Fuente Propia.	36
Tabla 16. Descripción de los Conjuntos de Datos Grupo B. Adatado de (Jaramillo & Mario, 2017).	37
Tabla 17. Codificación de Caracteres Grupo B. Adaptado de (Merino, 2012)	37
Tabla 18. Descripción de Identificadores de YouTube del Grupo B. Fuente Propia.	38
Tabla 19. Registros eliminados del Excel Grupo B. Fuente propia.	38
Tabla 20. Descripción del conjunto de datos del Grupo B. Fuente Propia.	39
Tabla 21. Descripción del conjunto de datos del Grupo B U1T1 a U1T5. Fuente Propia.	40
Tabla 22. Parámetros del LexRank con Umbral. Adaptado de (Anacona et al., 2015).	48
Tabla 23. Descripción de campos de resultados del algoritmo. Fuente propia.	52
Tabla 24. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante1GA. Fuente propia.	53
Tabla 25. Sesión s3, del estudiante1GA. Adaptado de (Jaramillo & Mario, 2017). 53	

Tabla 26. Patrones de navegación del estudiante1GA. Fuente propia.	54
Tabla 27. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante2GA. Fuente propia.....	55
Tabla 28. Sesión s5, del estudiante2GA. Adaptado de (Jaramillo & Mario, 2017).55	
Tabla 29. Patrones de navegación del estudiante2GA. Fuente propia.	56
Tabla 30. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante3GA. Fuente propia.....	57
Tabla 31. Sesión s16, del estudiante3GA. Adaptado (Jaramillo & Mario, 2017). ..	59
Tabla 32. Patrones de navegación del estudiante3GA. Fuente propia.	59
Tabla 33. Centralidad de los patrones de navegación EF-ISF por sesión del administradorGA. Fuente propia.	61
Tabla 34. Sesión s2, del administradorGA. Adaptado (Jaramillo & Mario, 2017)..	61
Tabla 35. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGA. Fuente propia.....	63
Tabla 36. Sesión s14, del profesorGA. Adaptado (Jaramillo & Mario, 2017).....	64
Tabla 37. Patrones de navegación del ProfesorGA. Fuente propia.	64
Tabla 38. Centralidad de los patrones de navegación EF-ISF por sesión del estudianteGAT1T5. Fuente propia.	66
Tabla 39. Sesión s27, del estudianteGAT1T5. Adaptado (Jaramillo & Mario, 2017).	66
Tabla 40. Patrones de navegación del estudianteGAT1T5. Fuente propia.	67
Tabla 41. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGAT1T5. Fuente propia.....	68
Tabla 42. Sesión, s12, del profesorGAT1T5. Adaptado (Jaramillo & Mario, 2017).	69
Tabla 43. Patrones de navegación del profesorGAT1T5. Fuente propia.	69
Tabla 44. Descripción de campos de resultados del algoritmo. Fuente propia.	70
Tabla 45. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante1GB. Fuente propia.....	71
Tabla 46. Sesión s5, del estudiante1GB. Adaptado (Jaramillo & Mario, 2017).	71
Tabla 47. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante2GB. Fuente propia.....	73
Tabla 48 . Sesión, s2, del estudiante2GB. Adaptado (Jaramillo & Mario, 2017). ..	73
Tabla 49. Patrones de navegación del estudiante2GB. Fuente propia.	74
Tabla 50. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante3GB. Fuente propia.....	75
Tabla 51. Sesión s15, del estudiante3GB. Adaptado (Jaramillo & Mario, 2017). ..	76
Tabla 52. Patrones de navegación del estudiante3GB. Fuente propia.	77
Tabla 53. Centralidad de los patrones de navegación EF-ISF por sesión del administradorAGB. Fuente propia.	79

Tabla 54. Sesión s5, del administradorAGB. Adaptado (Jaramillo & Mario, 2017).	79
Tabla 55. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGB. Fuente propia.....	81
Tabla 56. Sesión s56, del profesorGB. Adaptado (Jaramillo & Mario, 2017).....	82
Tabla 57. Patrones de navegación del profesorGB. Fuente propia.....	82
Tabla 58. Centralidad de los patrones de navegación EF-ISF por sesión del monitorGB. Fuente propia.....	84
Tabla 59. Sesión s9, del monitorGB. Adaptado (Jaramillo & Mario, 2017).....	84
Tabla 60. Patrones de navegación del monitorGB. Fuente propia.....	85
Tabla 61. Centralidad de los patrones de navegación EF-ISF por sesión del estudianteGBT1T5. Fuente propia.....	87
Tabla 62. Sesión s25, del estudianteGBT1T5. Adaptado (Jaramillo & Mario, 2017).	87
Tabla 63. Patrones de navegación del estudianteGBT1T5. Fuente propia.....	88
Tabla 64. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGBT1T5. Fuente propia.....	89
Tabla 65. Sesión s10, del profesorGBT1T5. Adaptado (Jaramillo & Mario, 2017).	89
Tabla 66. Patrones de navegación del profesorGBT1T5. Fuente propia.....	90
Tabla 67. Centralidad de los patrones de navegación EF-ISF por sesión del monitorGBT1T5. Fuente propia.....	91
Tabla 68. Sesión s8, del monitorGBT1T5. Adaptado (Jaramillo & Mario, 2017). ..	91
Tabla 69. Patrones de navegación del monitorGBT1T5. Fuente propia.....	92

Lista de Figuras

Figura 1. Arquitectura de Open edX. Tomado de (edX, 2018).	7
Figura 2. Imagen del LMS del Curso Astronomía Cotidiana Grupo B 2017-I. Tomado de (Jaramillo & Mario, 2017).	8
Figura 3. Estructura general del archivo traking.log de Open edX. Tomado de (Anacona et al., 2017).	10
Figura 4. Aproximación arquitectónica para el descubrimiento de patrones de navegación. Tomado de (Anacona et al., 2017).	18
Figura 5. Aproximación arquitectónica para la identificación de la centralidad de los patrones de navegación EF-ISF por sesión. Adaptado de (Anacona et al., 2017).	28
Figura 6. Transformación del conjunto de datos en el modelo de espacio vectorial. Adaptado de (Anacona et al., 2017).	32
Figura 7. Algoritmo para la identificación de patrones de navegación EF-ISF por sesión de mayor centralidad. Adaptado de (Anacona et al., 2017).	43
Figura 8. archivo Excel de usuarios representado en el modelo de espacio vectorial. Fuente Propia.	49
Figura 9. Identificación del patrón de navegación EF-ISF por sesión de mayor centralidad. Fuente propia.	50
Figura 10. Grafica de la centralidad del patrón de navegación EF-ISF por sesión del usuario <i>i</i> . Fuente propia.	51
Figura 11. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante1GA. Fuente propia.	54
Figura 12. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante2GA. Fuente propia.	57
Figura 13. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante3GA. Fuente propia.	60
Figura 14. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del administradorGA. Fuente propia.	62
Figura 15. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGA. Fuente propia.	65
Figura 16. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudianteGAT1T5. Fuente propia.	67
Figura 17. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGAT1T5. Fuente propia.	70
Figura 18. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante1GB. Fuente propia.	72
Figura 19. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante2GB. Fuente propia.	75
Figura 20. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante3GB. Fuente propia.	78

Figura 21. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del administradorAGB. Fuente propia.	80
Figura 22. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGB. Fuente propia.	83
Figura 23. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del monitorGB. Fuente propia.	86
Figura 24. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudianteGBT1T5. Fuente propia.	88
Figura 25. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGBT1T5. Fuente propia.....	90
Figura 26. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del monitorGBT1T5. Fuente propia.....	93
Figura 27. Graficas de centralidad de los patrones de navegación EF-ISF de los usuarios analizados. Fuente propia.....	94

CAPITULO 1

En este capítulo se describe el planteamiento del problema, objetivo general, objetivos específicos y resultados obtenidos.

1 INTRODUCCIÓN

1.1 Planteamiento del Problema

Los Cursos en Línea Abiertos y Masivos (MOOC, Massive Open Online Courses) es una propuesta basada en el Conectivismo¹, para universalizar la educación y ofrecer educación gratuita y de calidad a las personas desde cualquier lugar en el mundo (Siemens, 2004; Zhang & Lin, 2018). Su principal objetivo, es brindar alternativas educativas soportadas en tecnologías de Internet de acceso abierto (Guàrdia, Maina, & Sangrà, 2013; Hernández, 2013; Lin & Kao, 2018; T. Liyanagunawardena, Williams, & Adams, 2013; Solarte, Ramírez, & Jaramillo, 2017). Los MOOC presentan ciertas características similares como: empleo de material de aprendizaje (mapas de conocimiento, videoconferencias, textos y entre otros) (Fasihuddin, Skinner, & Athauda, 2015; Tang, Xing, & Pei, 2018); por lo general son de corta duración (T. R. Liyanagunawardena, Adams, & Williams, 2013); atraen a miles de estudiantes de diferentes lugares del mundo (Kizilcec, 2013; Rizvi, Rienties, & Rogaten, 2018); los estudiantes inscritos a estos cursos virtuales tienen la oportunidad de aprender con los docentes de diferentes Universidades (Rizvi et al., 2018; Sonwalkar, 2013) y algunos emiten un certificado de aprobación, el cual puede tener un valor monetario (Kizilcec, 2013; McAuley, Stewart, Siemens, & Cormier, 2010).

Los MOOC se clasifican de manera general en: cMOOC, cuyo objetivo es la generación de conocimiento de forma colaborativa, basándose en la cantidad de aportes generados por sus participantes; los xMOOC, tienen en cuenta la presentación de contenidos a través de videos cortos, documentos de apoyo y evaluaciones (Guàrdia et al., 2013; Hernández, 2013; Wang, Law, Hemberg, & O'Reilly, 2019); el termino SPOC (*Small Private Online Courses* - Pequeños Cursos Privados en Línea), aparece en el año 2013, estos son cursos en línea soportados en plataforma MOOC, pero con la diferencia de no ser abiertos a todo el público y con una cantidad menor de estudiantes (Fox, 2013); en el año 2014, aparecen los MPOC (*Massive Private Online Courses* - Masivos Cursos Privados en Línea), igualmente soportados en plataforma MOOC, con la posibilidad de ser masivos,

¹ Una teoría educativa propuesta por George Siemens; concibe al aprendizaje como el proceso creativo y social de la conexión de los nodos de conocimiento.

pero sin estar abiertos a todo el público y con una cantidad mayor de estudiantes (W. Guo, 2014).

El primer curso estilo MOOC orientado fue "Conectivismo y conocimiento colectivo" ofrecido en el año 2008 por Stephen Downes, de la Universidad de Manitoba de Canadá (Mackness, Mak, & Williams, 2010); sin embargo, el mayor reconocimiento de la estrategia MOOC se consigue por el curso de "Introducción a la Inteligencia Artificial" orientado por Sebastián Thrun, profesor de la Universidad de Stanford de EE.UU y Peter Norvig, Investigador de Google, en el cual se matricularon aproximadamente 160.000 estudiantes (Martin, 2012). Según un estudio realizado, en el año 2018 más de 101 millones de estudiantes se han matriculado en MOOC ofertados por diferentes universidades (Central, 2018).

Los MOOC consolidan su fama mundial en el año 2012, con la aparición de diversas plataformas como: Udacity, Coursera, Open edX y entre otras, en las que se ofrecieron cursos de diferentes Universidades (Hernández, 2013; Zhang & Lin, 2018). En la plataforma Open edX, se ofrecen xMOOC, su estructura está principalmente centrada en el profesor, quien crea la mayor parte del material de aprendizaje que generalmente es organizado en el curso, como una estructura lineal de arriba hacia abajo, en secciones, subsecciones y unidades, el cual es mostrado a través de la plataforma al estudiante, según la organización del curso que plantea el docente (P. J. Guo & Reinecke, 2014; Rizvi et al., 2018).

Los xMOOC, difieren de los cursos presenciales y del aprendizaje tradicional, ya que los estudiantes pueden aprender a su propio ritmo, repetir u omitir lecciones y conducir el proceso de aprendizaje de forma más independiente (P. J. Guo & Reinecke, 2014; J. J. Maldonado et al., 2017); sin embargo, los xMOOC son criticados por replicar la enseñanza tradicional, basada en conferencias o clases magistrales y no tener presente las diferentes estrategias de aprendizaje, como la preferencia del estudiante por una estructura más o menos lineal en la organización del material de aprendizaje y en la estructura de la plataforma (P. J. Guo & Reinecke, 2014; Rizvi et al., 2018).

En el contexto del aprendizaje en línea, existen estudiantes que prefieren seguir un camino de aprendizaje definido externamente, como lo impone un profesor o el entorno de aprendizaje en línea. Hay otros estudiantes, que se sienten seguros en entornos de aprendizaje no lineales, es decir en la definición de sus propios caminos de aprendizaje, lo que indica que navegan libremente sin seguir necesariamente lo sugerido por los creadores de contenido o el entorno de aprendizaje en línea (P. J. Guo & Reinecke, 2014; Rizvi et al., 2018; Zhang & Lin, 2018).

En la literatura se encuentran algunos estudios referente a MOOC en minería de datos educativos, en los cuales se evidencian análisis de archivos o conjuntos de datos² (Korhonen & Multisilta, 2017), para caracterizar patrones de navegación EF-ISF³ (F. Anacona, M. Solarte, & G. González, 2018), encontrar patrones, navegaciones y patrones de navegación de los estudiantes, a través de herramientas y/o métodos estadísticos y/o algoritmos, entre los cuales se encuentran: la participación activa en el curso (Yang, Wen, Kumar, Xing, & Rose, 2014); los espectadores (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2014); los videos vistos (Kim et al., 2014), la navegación de los estudiantes en el curso MOOC (Jorge J Maldonado et al., 2016; Sonwalkar, 2013) y patrones de navegación, acceso a los problemas de videos, acceso a wiki y entre otros (Zhang & Lin, 2018).

Teniendo en cuenta que los estudiantes generan diferentes navegaciones en el curso MPOC y hasta el momento no se encuentran trabajos que identifiquen los patrones de navegación EF-ISF por sesión de mayor centralidad, en conjuntos de datos que contengan la interacción de los estudiantes con MPOC de la plataforma Open edX, por esta razón se plantea la siguiente pregunta de investigación: ¿Cómo verificar la existencia de los patrones de navegación EF-ISF por sesión de mayor centralidad de los estudiantes, en conjuntos de datos de MPOC de Open edX?

1.2 Objetivos

A continuación, se presentan los objetivos desarrollados en el presente trabajo de investigación.

1.2.1 Objetivo General

Proponer un mecanismo para la identificación de patrones de navegación EF-ISF por sesión de mayor centralidad, en conjuntos de datos que contienen la interacción de los estudiantes, de un MPOC alojado en una instancia de la plataforma Open edX.

1.2.2 Objetivos específicos

- 1) Realizar pre procesamiento de conjuntos de datos que contienen la interacción de los estudiantes, de un MPOC de la plataforma Open edX.
- 2) Adaptar el algoritmo LexRank con Umbral, para la identificación de la centralidad de los patrones de navegación EF-ISF por sesión de los estudiantes.
- 3) Desarrollar prueba de concepto con el algoritmo LexRank con Umbral en los conjuntos de datos procesados, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los estudiantes.

² Contienen algunas interacciones de los estudiantes con el MOOC o xMOOC.

³ Ponderación basada en la frecuencia relativa de un evento.

1.3 Resultados Obtenidos

Con el desarrollo de la propuesta de investigación se lograron los siguientes resultados:

1) **Monografía del trabajo de grado**, con la siguiente estructura:

- **CAPÍTULO 1:** se describe el planteamiento del problema, objetivo general, objetivos específicos y resultados obtenidos.
- **CAPÍTULO 2:** se describe la plataforma Open edX, artículos de patrones, navegaciones y patrones de navegación de los estudiantes en MOOC, además del proceso general en generación automática de resúmenes de múltiples documentos y por ultimo las conclusiones de este capítulo.
- **CAPÍTULO 3:** contiene la aproximación arquitectónica para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios, la cual consta de los siguientes módulos: conjunto de datos, pre procesamiento, algoritmo LexRank con Umbral, patrón de navegación EF-ISF por sesión de mayor centralidad.
- **CAPÍTULO 4:** Se describe el proceso general para la identificación de los patrones de navegación EF-ISF por sesión en conjuntos de datos, los parámetros del algoritmo LexRank con Umbral y la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del usuario i .
- **CAPÍTULO 5:** Se describen los resultados y análisis de la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios y la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del usuario i , en algunos usuarios de los cuatro archivos Excel y por ultimo las conclusiones de este capítulo.
- **CAPÍTULO 6:** Se describen las conclusiones del presente trabajo de investigación y trabajos futuros.
- **REFERENCIAS:** contiene las referencias de los artículos utilizados en el presente trabajo de investigación.

2) **Anexos:** información que no está descrita en la monografía.

- **Anexo A:** artículo titulado “Descubrimiento de patrones de navegación en Open edX – una aproximación arquitectónica”, publicado en la Revista de Ingeniería e Innovación de la Universidad de Córdoba – Colombia, revista nacional indexada por COLCIENCIAS categoría C en la fecha de publicación,

el documento puede ser consultado en los Anexos y en la siguiente dirección: <https://revistas.unicordoba.edu.co/index.php/rii/article/view/1103>.

- **Anexo B:** ponencia en el Seminario Internacional de Innovaciones Educativas y MOOC SINNEM18, realizada en la Universidad del Cauca, Popayán, Colombia, el día viernes 5 de octubre de 2018; el título de la ponencia fue “LexRank con Umbral para identificar patrones de navegación EF-ISF en xMOOC de Open edX”. El artículo se encuentra en las actas del evento en la siguiente dirección: <https://www.unicauca.edu.co/moocmaker/sinnem/memorias.html> y en los Anexos, el artículo está en proceso de evaluación en la revista de Ingenierías de la Universidad de Medellín, Colombia.
 - **Anexo C:** ponencia en la Segunda Conferencia Internacional MOOC MAKER 2018, realiza el día jueves 11 de octubre de 2018, en el Hotel TRYP de la ciudad de Medellín, Colombia; el título de la ponencia fue “Modelo de Espacio Vectorial con ponderación basada en frecuencia relativa de eventos de navegación en una instancia de Open edX para caracterización del estudiantado”. El artículo de la ponencia se encuentra en las actas del evento y puede ser consultado en la siguiente dirección: <http://ceur-ws.org/Vol-2224/10.pdf> y en los Anexos y en los Anexos.
 - **Anexo D:** descripción de las metodologías utilizadas y diagrama de clases que se implementó, para el desarrollo del presente trabajo de investigación.
- 3) **Código fuente** que permite la evaluación del mismo con los archivos Excel o conjuntos de datos, que contienen la interacción de los usuarios con el curso de Astronomía Cotidiana 2017-I, MPOC, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios. El software fue implementado en lenguaje de programación C# en el entorno de desarrollo VS.NET 2015.

CAPITULO 2

En este capítulo se describe el Sistema de Gestión de Aprendizaje (LMS - Learning Management System), el archivo tracking.log, de la aplicación web Open edX, los artículos que identifican patrones, navegación y patrones de navegación de los estudiantes en cursos MOOC, además, la descripción de generación automática de resúmenes de múltiples documentos y las conclusiones del marco teórico.

2 MARCO TEÓRICO

2.1 Open edX

Edx es una organización sin ánimo de lucro fundada por las Universidades de Harvard, el Instituto Tecnológico de Massachusetts (MIT) y la componen diferentes instituciones a nivel mundial. Edx tiene un sitio de aprendizaje en línea en la Red Informática mundial (WWW - World Wide Web) en la siguiente dirección: <https://www.edx.org>, en la cual se provee MOOC de alta calidad de Universidades e Instituciones a estudiantes de diferentes lugares en el mundo (edX, 2017).

Open edX es una plataforma de código abierto y está disponible para su descarga e instalación de manera gratuita en la siguiente dirección: https://edx.readthedocs.io/projects/edx-installing-configuring-and-running/en/open-release-qinkgo.master/installation/fullstack/install_fullstack.html, la plataforma permite crear cursos xMOOC, herramientas de aprendizaje y aportar nuevas funciones, creando soluciones innovadoras para beneficiar a los docentes y estudiantes de todo el mundo (edX, 2017).

En la plataforma se pueden encontrar varias características, como las siguientes: la posibilidad de mostrar lecciones grabadas en video con subtítulos e indexación sobre los propios subtítulos, añadir materiales de estudio (libros, notas), realizar evaluaciones, crear laboratorio virtual con interfaz interactivo, foros de discusión, wikis, informes de progreso, sistema para la implementación de analítica de aprendizaje, diferentes tipos de evaluación de tareas (evaluación entre pares, auto evaluación y evaluación automática mediante técnicas de aprendizaje automático) (Palta & Vásquez, 2016).

Open edX cuenta con una arquitectura de múltiples componentes, los cuales se ejecutan de manera independiente para ofrecer servicios a la plataforma y a los usuarios (Palta & Vásquez, 2016), la arquitectura de Open edX se muestra en la Figura 1.

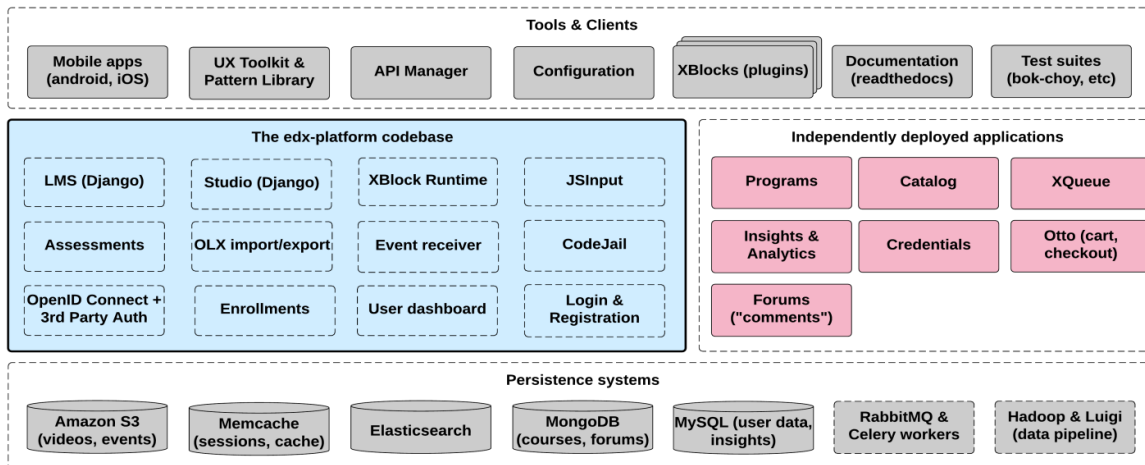


Figura 1. Arquitectura de Open edX. Tomado de (edX, 2018).

En el presente trabajo de investigación, se encontrarán los patrones de navegación EF-ISF por sesión de mayor centralidad de los estudiantes, en conjuntos de datos, ver secciones 3.2.1 y 3.2.2, que contienen la interacción con el LMS, por esta razón se describirá este componente.

El LMS de Open edX permite la interacción en el curso de los estudiantes, docentes y entre otros, con cuestionarios, evaluaciones, foros de discusión, aplicaciones interactivas, también, muestra el contenido del curso a través de secciones, subsecciones y unidades; el instructor cuenta con un centro de mando, en el cual puede producir informes y administrar el curso, entre otras funciones; además el LMS utiliza bases de datos como: MogoDB, almacena los cursos y las referencias a los videos de YouTube y Amazon CDN; MySQL, guarda los datos de los estudiantes (Jaramillo & Mario, 2017; Palta & Vásquez, 2016).

En la Figura 2 se muestra una imagen de LMS, del curso Astronomía Cotidiana del Grupo B, del año 2017 primer periodo académico de la Universidad del Cauca, curso creado en una instancia de la plataforma Open edX y puesto en línea a través del servidor llamado Selene.



Figura 2. Imagen del LMS del Curso Astronomía Cotidiana Grupo B 2017-I. Tomado de (Jaramillo & Mario, 2017).

En las Sección(es), por lo general se publica el tema(s) principal del curso, en las Subsección(es) los sub temas del curso y en las unidades pueden ir cuatro tipos de material de aprendizaje, foros, ejercicios, videos y Html (Lenguaje de marcas de hipertexto - HyperText Markup Language), estos se describen a continuación (Jaramillo & Mario, 2017):

- **Foro:** permite añadir una discusión o abrir un debate referente a la Unidad temática, para publicar preguntas y permitir a los estudiantes responder e interactuar entre ellos y el profesor. El tema de discusión estará disponible en la estructura del curso y también se puede acceder en la pestaña de discusión en el LMS
- **Ejercicios:** se visualizan en el curso diferentes tipos de ejercicios interactivos de calificación automática, acorde a los propuestos por el docente(s). Por lo general los ejercicios disponibles son: Checkbox, Lista desplegable y Opción múltiple.
- **Videos:** se muestran los videos de la unidad temática correspondiente, siendo referenciados en la plataforma de YouTube o en otras plataformas como Amazon S3; además el estudiante puede pulsar sobre un enlace para descargar el vídeo desde la localización alternativa.
- **Html:** a través de código de Html se puede adicionar libros digitales, pdf, texto, listas, enlaces, presentaciones e imágenes.

2.1.1 Archivo Tracking.log

El tracking.log es un archivo de Notación de Objetos de JavaScript (JSON - JavaScript Object Notation) que guarda un registro cada vez que se realiza alguna actividad o evento por parte del estudiante(s) en el LMS. Las actividades pueden

ser las siguientes (F. Anacona et al., 2018; Anacona, Solarte, & Ramírez, 2017; Palta & Vásquez, 2016; Yadav, 2014):

- Hacer clic en un enlace.
- Ver o descargar un video.
- Leer o descargar documentos, pdf, ppt y entre otros.

Existen varios beneficios de almacenar registros del LMS, entre los que se encuentran: monitoreo del progreso del estudiante, verificar el uso del material de aprendizaje en el LMS por parte del estudiante; comprobar el uso de LMS por los estudiantes, como han utilizado los estudiantes el material de aprendizaje y el foro en LMS; consultas personalizadas, responder a preguntas a partir de la información almacenada en el archivo de texto (Yadav, 2014).

La entrada de registro, es una estructura de datos que almacena pocos campos, según el tipo de entrada o tipo de actividad, en el que la actividad también se conoce como evento. La entrada de registro es clave - valor (key - value) para los campos y sus valores, el cual se registra en formato JSON (Yadav, 2014).

El archivo tracking.log en la plataforma Open edX se encuentra en la ruta /edx/var/log/tracking/ y guarda los registros de entrada de los tipos de eventos generados por el servidor, navegador o dispositivo móvil; a pesar de que dichos eventos pueden representar acciones distintas, el registro de los eventos en el archivo tiene un conjunto de campos comunes, la estructura general se presenta a continuación (Anacona et al., 2017):

```
{
  "username": "",
  "event_source": "",
  "name": "",
  "accept_language": "",
  "time": "",
  "agent": "",
  "page": "",
  "host": "",
  "session": "",
  "referer": "",
  "context":
  {
    "user_id":,
    "org_id": "",
    "course_id": "",
    "path": ""
  },
}
```

```

    "ip": "",
    "event": "",
    "event_type": ""
  }

```

Figura 3. Estructura general del archivo tracking.log de Open edX. Tomado de (Anaconda et al., 2017).

En las siguientes tablas se describen los eventos.

Tabla 1. Descripción de campos. Tomado de (Palta & Vásquez, 2016).

Campo	Tipo	Descripción
accept_lenguaje	string	Determina el lenguaje por defecto del navegador.
agent	string	Reconoce al principal agente de usuario que inicia el evento.
context	object	Tiene un conjunto de sub campos que describen información del evento, los sub campos pueden ser comunes para todos los eventos como se muestra en la Tabla 2; sin embargo, pueden contener campos adicionales, ver Tabla 3.
event	object	Reconoce los campos específicos de cada evento iniciado, los campos pueden variar dependiendo del tipo de evento, los cuales pueden ser: de navegación y de inscripción, entre otros.
event_source	string	Determina el origen que genero el evento, los cuales pueden ser: "browser", "mobile", "server" y "task".
event_type	string	Identifica el tipo de evento registrado acorde al campo de "event_source".
host	string	Registra el sitio visitado por el usuario.
ip	string	La dirección "ip" del usuario cuando genera el evento.
name	string	Registra el tipo de evento lanzado.
page	string	Registra la URL de la página visitada al momento de generar el evento.
referer	string	Registra la cabecera HTTP del cual llego el evento.

session	string	Registra una cadena de 32 caracteres de la sesión del usuario.
---------	--------	--

Tabla 2. Campos del evento context. Tomado de (Palta & Vásquez, 2016).

Campo	Tipo	Descripción
course_id	string	Registra el curso en el que se generó el evento.
org_id	string	Registra la organización que oferta el curso.
path	string	identifica la URL que genera el evento.
user_id	number	Registra el usuario que género el evento.

Tabla 3. Sub campos adicionales del evento context. Tomado de (Palta & Vásquez, 2016).

Campo	Tipo	Descripción
course_user_tags	object	Registra la clave y el valor de la tabla user_api_usercoursetag para el usuario en el curso.
module	object	Registra información de los componentes que se involucran en el evento generado por el servidor.

2.2 Patrones de estudiantes en MOOC

En los MOOC, la interacción de los estudiantes con el material de aprendizaje, proporciona un registro de las actividades sin precedentes, los cuales son almacenados en conjuntos de datos. En el desarrollo de la propuesta del artículo, trabajan con conjuntos de datos de Coursera. Proponen una taxonomía del comportamiento individual, en el cual examinan diferentes patrones de compromiso de estudiantes de alto y bajo rendimiento, clasificándolos como: *Viewers*, *Solvers*, *All-rounders*, *Collectors*, *Bystanders*; e investigaron cómo la participación en los foros se relaciona con otros componentes del curso. De igual forma, reportaron los efectos del despliegue de una serie de incentivos para la participación en los MOOC, incluyendo experimentos aleatorios, en los cuales la presentación de dichos distintivos fue variada en las subpoblaciones, encontrando un aumento en la participación en el foro. Como trabajo futuro proponen lo siguiente: modelos predictivos del comportamiento y las calificaciones de los estudiantes; mecanismos de personalización y recomendación para ayudar a aumentar la participación y el aprendizaje de los usuarios, entre otros (Anderson et al., 2014).

Los MOOC se distinguen por su alcance global y asíncrono, por esta razón es importante entender las formas en cómo los estudiantes interactúan con estos

entornos. Los foros de discusión reúnen a miles de personas, siendo un componente clave para investigar cómo se producen las interacciones y determinar si se puede ofrecer nuevas perspectivas en lo pedagógico y en el potencial de los MOOC. Crearon un conjunto de datos a partir de encuestas pre y post curso y las interacciones sobre el foro, luego utilizaron Análisis de Redes Sociales⁴ para visualizar e interpretar los datos, en el cual encuentran tres tendencias: primero, los participantes del foro tienden a ser jóvenes adultos de la parte occidental del mundo; segundo, los estudiantes se reúnen y se dispersan no como las comunidades tradicionales; por último, los que tienen mayor participación en el foro son los de mayor rendimiento (Gillani & Eynon, 2014).

Además, los MOOC representan una alternativa educativa con poca barrera de entrada, al permitir la inscripción a cualquier persona. Los cursos presentan bajas tasas de finalización, entre el 5 al 10% de los estudiantes reciben un certificado de aprobación. Poco se ha investigado sobre los patrones de deserción y cómo éstos pueden variar en los diferentes grupos de estudiantes. Analizan conjuntos de datos de Coursera, el cual contiene información de exámenes desarrollados y videos vistos por los estudiantes, encuentran dos grupos de estudiantes (Auditores) los que miran videos y (Participante) los que miran videos y desarrollan exámenes; la deserción en los cursos es similar a pesar de, duración, dificultad, variación de temas; o factores externos como falta de tiempo, dificultades en el acceso a la información y presión externa (Glance, Barrett, & Hugh, 2014).

En otro estudio realizado en tres MOOC, utilizan conjuntos de datos de Coursera para analizar el comportamiento de abandono de los estudiantes en los cursos. Encuentran veinte subcomunidades emergentes a partir de la discusión de los hilos en los foros, luego utilizan un modelo de supervivencia, para medir la participación de los estudiantes en los foros, hallan que de dos a cuatro subcomunidades se asocian con tasas altas o bajas de abandono en promedio. Un análisis cualitativo post-hoc, evidencia como los modelos de aprendizaje se pueden utilizar para entender los valores y los focos de discusión en las subcomunidades, con lo cual encuentran los siguientes patrones: inicialmente se identificó una motivación expresada para participar activamente en el curso; luego se muestra un compromiso cognitivo con el material en las subcomunidades con menor deserción; presentándose lo contrario con las de mayor deserción. Como trabajo a futuro plantean analizar la interacción de los estudiantes con diferentes entornos sociales como: correo electrónico, Facebook y entre otros (Yang et al., 2014).

La gran cantidad de estudiantes que ven videos en los cursos MOOC, generan información de interacción, lo cual es una oportunidad para entender cómo ellos

⁴ Librería de Python NetwokX

aprenden por medio de este material de aprendizaje, al analizar en donde hay máxima o mínima audiencia. Para ello hacen uso de un conjunto de datos que contiene la interacción de los estudiantes en videos de cuatro cursos MOOC de la plataforma Open edX; para realizar el procesamiento de los datos los autores utilizaron Insights y calcularon la tasa de abandono en los videos con la siguiente formula: $1 - \text{contadorVisulizacion}(n) / \text{contadorVisulizacion}(0)$, encuentran mayores tasas de abandono en videos largos, videos vistos vs primera vez, tutoriales (explicaciones en video) vs conferencias; audiencia máxima en, videos vistos vs primera vez, tutoriales vs conferencias; al analizar la audiencia máxima a través de una variación del algoritmo TwitInfo, identificaron los siguientes patrones: ver el video desde el inicio, volver a un punto de fallo en el video, seguir un tutorial, repetir parte del video y repetir un tutorial (Kim et al., 2014).

El material de aprendizaje de los MOOC archivados (cursos que ya se han dado) pueden ser utilizado como material de estudio por los estudiantes, frente a las nuevas versiones de cursos o MOOC activos (cursos que se estén dando y estén orientados por un docente); con lo cual se puede comparar el uso del material de aprendizaje por parte de los estudiantes en ambos cursos. Para realizar el análisis utilizan: datos de las encuestas realizadas a los estudiantes previos al curso, información de demografía y razones de la inscripción; los conjuntos de datos de Coursera que contiene: videos a los que accede, evaluaciones enviadas, publicaciones en foros y los clics de acciones realizados por los estudiantes. Luego comparan la demografía, propósito y los comportamientos de los estudiantes, con lo cual encuentran: los estudiantes que hacen uso de los MOOC archivados están interesados en los MOOC activos, además de los patrones de uso: número y secuencia de videos que vieron, numero de evaluaciones que completaron, comportamiento autorregulado y la participación en el foro fue similar en ambos cursos. También determinaron que los estudiantes se basan en el MOOC archivado como material de referencia (Campbell, Gibbs, Najafi, & Severinski, 2014).

En cinco MOOC de Coursera, con más de 400.000 estudiantes de la Universidad de Maryland, utilizan Tableau⁵ como herramienta de análisis visual para explorar la participación. Para esto realizan dos observaciones: la primera observación es el grado por actividad, numero de clics, número de páginas vistas y número de visitas al foro; esto les permite identificar dos patrones: el primero, hay tres grupos de estudiantes, en los cuales los autores presumen que se corresponden con las trayectorias prototípicas: auditoria (*at the bottom*), detrás (*in the middle*) y rastro (*on top*); el segundo, la forma en que varían los grupos según la actividad de los estudiantes; la segunda observación, examina la forma en que los perfiles de los estudiantes se correlacionan con el rendimiento, sin encontrar resultados

⁵ Herramienta comercial para examinar datos

concluyentes. Sin embargo, los autores recomiendan realizar un trabajo mayor para confirmar los resultados (Xu, Goldwasser, Bederson, & Lin, 2014).

En un estudio realizado en un MOOC, proponen identificar el estilo de aprendizaje de los alumnos, a través de las interacciones de los estudiantes con el material de aprendizaje del curso. Con el fin de determinarlos, aplican el método basado en literatura para entornos de aprendizaje abiertos, utilizan patrones definidos por Felder y Silverman, entre otros autores y para los valores de umbral. La identificación de los estilos de aprendizaje, es basado en el Modelo de estilo de aprendizaje de Felder y Silverman (FSLSM), en el que el estilo de aprendizaje se refiere a la forma en que un alumno recibe y procesa la información. Para el desarrollo de la propuesta trabajan con el método FSLSM, en la creación del contenido en el curso MOOC; como método para el cálculo de patrones utilizan la metodología de literatura; para identificar el estilo de cada estudiante usan identificación de estilos de aprendizaje (ILS) y para comparar los resultados obtenidos lo realizan con Precisión. Encontraron que el uso del método basado en literatura para identificar estilos de aprendizaje en cursos MOOC, es útil para el desarrollo de un marco adaptativo (Fasihuddin et al., 2015).

2.3 Navegación de estudiantes en MOOC

En los cursos MOOC los instructores proponen una estructura lineal de arriba hacia abajo con un orden cronológico de semanas y secuencias de aprendizaje en la forma de publicar los contenidos, con el fin de ser seguida por los estudiantes. Por otro lado, los investigadores proponen determinar cómo los estudiantes navegan a través del contenido del curso, para ello utilizan un conjunto de datos de MOOC alojados en la plataforma Open edX. Realizaron un análisis de regresión lineal múltiple, ANOVA F, valores-p y coeficientes-b; mostraron que, a pesar de la estructura lineal impuesta, los estudiantes navegan en el contenido del curso de una manera no lineal, en el cual omiten el 22% de las secuencias de aprendizaje y realizan una gran cantidad de saltos hacia atrás, de evaluaciones a contenidos anteriores (P. J. Guo & Reinecke, 2014).

En un MOOC, realizan un estudio exploratorio desde una perspectiva orientada al proceso, para estudiar si los estudiantes con diferentes estilos de aprendizaje y perfiles de autorregulación (SRL, Self Regulation Learning) muestran diferencias en la navegación a través de un MOOC de Open edX. Utilizaron técnicas de minería de procesos para analizar archivos que registran el comportamiento de estudiantes en videos, lectura y evaluación, combinado con datos de cuestionarios de SRL y estilos de aprendizaje. Con la herramienta de Disco⁶ - Disco Tool, basado en minería de Fuzzy, mostraron que los estudiantes con diferentes perfiles de SRL siguen

⁶ Software comercial: <https://fluxicon.com/disco/>

caminos similares de navegación; pero hay diferencias en los estudiantes por su estilo de aprendizaje (Jorge J Maldonado et al., 2016).

En otro curso MOOC, hacen uso del modelo de aula invertida, para que los estudiantes utilicen el curso a su propio ritmo de trabajo; sin embargo, los alumnos aprenden a un ritmo diferente y usan varios recursos durante su aprendizaje, lo que requiere de ellos la capacidad de autorregularse. En el documento proponen un estudio piloto e investiga cómo los estudiantes con diferentes perfiles de autorregulación navegan a través de un MOOC, cuando se usa como parte de la metodología del Aula Invertida. Mediante minería de procesos con la Herramienta de Disco, realizan análisis sobre los registros log de un curso MOOC de Open edX. Los resultados mostraron que los estudiantes con perfiles diferentes de autorregulación del aprendizaje –SRL, siguen rutas de navegación distintas, además, se evidenció que los estudiantes asociados con Aula Invertida obtuvieron mejores resultados con respecto a los estudiantes orientados de forma tradicional. Como trabajo futuro plantean realizar el mismo estudio para un semestre y complementar el MOOC (J. J. Maldonado et al., 2017).

Los cursos en línea abiertos y masivos –MOOC permiten la matrícula de una gran cantidad de estudiantes, sin embargo, la permanencia de los alumnos ha sido uno de los problemas más importantes asociados con el aprendizaje MOOC, se pueden encontrar algunos estudios sobre la permanencia de los estudiantes en los cursos; sin embargo, poco se sabe sobre la razón en el nivel de participación o la progresión de los alumnos en las actividades de aprendizaje en los MOOC, es decir, las trayectorias o navegación de aprendizaje de los alumnos. En la Universidad Abierta, se ofreció un curso MOOC a través de la plataforma FutureLearn. Analizan el conjunto de datos que contiene registros de identificación anónima de los alumnos, número de semana, tipo de actividad de aprendizaje, actividad de aprendizaje y marcas de tiempo; mediante el uso de métodos exploratorios asociados con la minería de procesos educativos (EPM), a través de Herramienta de Disco. Las actividades de los alumnos se mapearon para identificar las navegaciones de aprendizaje comunes y distintas, en dos grupos de estudiantes, los que completaron el 50% de las actividades y los que no. Usaron medida de frecuencias relativas, para comparar los comportamientos participativos de ambos grupos con el comportamiento de aprendizaje esperado para todos los tipos de actividades de aprendizaje. También, exploran el rendimiento semanal típico, identifican y mapean las navegaciones de aprendizaje temporal más significativas del subgrupo de estudiantes. Los resultados indicaron que existía al menos una navegación principal y dominante; pero las navegaciones dominantes en los subgrupos de estudiantes permanecieron notablemente distintas, además, concluyen, para mejorar la

generalización de los hallazgos del estudio, se requiere una investigación profunda entre más diseños y plataformas MOOC (Rizvi et al., 2018).

2.4 Patrones de Navegación de los estudiantes en MOOC

Los MOOC están liderando una revolución educativa, reuniendo recursos educativos de diferentes partes del mundo y reestructurando el entorno de aprendizaje. La plataforma MOOC registra los clicks del usuario, lo que proporciona información valiosa sobre la forma en que el usuario interactúa con el curso. Los autores analizan un conjunto de datos de un curso de cinco semanas de 3914 usuarios, a nivel de sesión del usuario, la sesión la definen como la interacción del estudiante cuando ingresa al MOOC, si en los registros de los datos no encuentra actividad del estudiante por un periodo de 30 minutos termina la sesión. El conjunto de datos contiene las siguientes interacciones: problema, trabajar en tareas del curso; video, ver videos del curso; acceso, acceso a materiales del curso, excepto videos y cuestionarios; Wiki, acceso al curso wiki; discusión, acceso al foro del curso; navegar, navegando a otra parte del curso y cierre de página, cierre la página web. Proponen una métrica para medir la similitud entre trayectorias de navegación a nivel de sesión del usuario, adaptando el algoritmo de la Subsecuencias comunes más largas (LCS - Longest Common Subsequences) y construir un modelo de agrupación sin supervisión para capturar los patrones de navegación del usuario en los MOOC, con el algoritmo de agrupamiento jerárquico aglomerativo (AHC - Agglomerative Hierarchical Clustering), con lo cual identifican trece tipos de patrones de navegación del usuario. Concluyen que la participación difiere en los diferentes tipos de actividades en el curso, como: trabajar en tareas y ver videos, entre otros. Como trabajo futuro proponen explorar y analizar los grupos en detalle para identificar otros patrones de navegación (Zhang & Lin, 2018).

A medida que los MOOC crecen en popularidad, las tasas relativamente bajas de finalización de los alumnos ha sido una crítica central. Sin embargo, este enfoque en las tasas de finalización refleja una visión monolítica de la deserción que no permite a los diseñadores de MOOC enfocarse en intervenciones o desarrollar características de cursos adaptativos para subpoblaciones particulares de estudiantes. Para abordar esto, trabajan con tres conjuntos de datos de tres MOOC de ciencias de la computación. Generan una descripción aproximada de la participación individual de cada estudiante por curso con video conferencias y evaluaciones. Con el algoritmo k-means encuentran subpoblaciones en las descripciones de participación, con lo cual identifican cuatro patrones de trayectorias de compromiso longitudinal, Auditoria - Auditing, Completando - Completing, Desacoplamiento - Disengaging y Muestreo - Sampling. Comparan los estudiantes en cada trayectoria y curso mediante pruebas estadísticas formales, ANOVA, n^2 y Cohen's d ; con información de datos demográficos,

participación en foros, acceso a videos e informes de experiencia general. Los autores concluyen que los estudiantes agrupados en tres de las cuatro trayectorias identificadas se considerarían en otros estudios como deserción del curso, en cambio los estudiantes que se encuentran en el patrón de trayectoria Completado son aquellos que completarían el curso. Aunque encuentran patrones de alto nivel, la gran cantidad de información disponible debe permitir descubrir tendencias más sutiles y profundas que permitan inspirar el diseño de herramientas, adaptadas a las necesidades de tipos particulares de estudiantes (Kizilcec, Piech, & Schneider, 2013).

Los foros de discusión son cada vez más importantes para los cursos en línea abiertos y masivos -MOOC. La participación activa en el foro se relaciona positivamente con los logros de los estudiantes en el sentido de que más publicaciones producen un mejor rendimiento, sin embargo, la participación en el foro varía temporalmente entre los estudiantes con diferentes motivaciones al realizar MOOC, con lo cual las trayectorias longitudinales varían. Para proporcionar apoyo oportuno a los alumnos a que se mantengan comprometidos, es importante comprender la variación temporal de la participación en los foros y cómo las diferentes motivaciones explican la variación. Realizan un análisis en un conjunto de datos, el cual se derivó de un MOOC ofrecido por una universidad en el noreste de los Estados Unidos - northeastern United States. Con el algoritmo K-means longitudinal -kml, identifican distintas trayectorias longitudinales de la participación del foro de estudiantes en MOOC, lo que permite nuevas perspectivas para proporcionar apoyo oportuno para que los alumnos se mantengan comprometidos con los MOOC. La investigación revela que los alumnos intrínsecamente motivados tienen más probabilidades de participar longitudinalmente en los MOOC, además, el estudio confirma que los alumnos que participan constantemente en foros de discusión con el tiempo tienen más probabilidades de sobresalir en los MOOC. Como trabajo futuro proponen extender el estudio a un conjunto(s) de datos que contengan mayor información de diversas plataformas MOOC (Tang et al., 2018).

Los MOOC se han convertido en una estrategia educativa disruptiva en la última década, caracterizada por dar soporte a procesos formativos a gran escala y de forma gratuita; a pesar de su éxito igualmente tienen debilidades como la rigidez pedagógica pues generalmente las actividades de aprendizaje son las mismas sin importar el perfil y número de los estudiantes. Desde hace algunos años se aplican técnicas de adaptación; pero sin un referente claro sobre cómo realizarla. En el trabajo presentan una aproximación a una arquitectura para descubrir patrones de navegación en los estudiantes en MOOC soportados en la plataforma Open edX. La aproximación arquitectónica consta de los siguientes módulos: archivo tracking.log, se encarga de almacenar los eventos realizados en el curso xMOOC

de la plataforma Open edX en un formato JSON; pre procesamiento, limpia los datos recolados de información que no aporten para ser representados en un archivo de texto plano y posteriormente son procesados en modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento –EF-ISF; algoritmo, proponen utilizar el LexRank con Umbral para la identificación de los patrones de navegación EF-ISF. Como trabajo a futuro plantean la implementación de la propuesta realizada. En la Figura 4, se muestra la propuesta arquitectónica (Anacona et al., 2017).

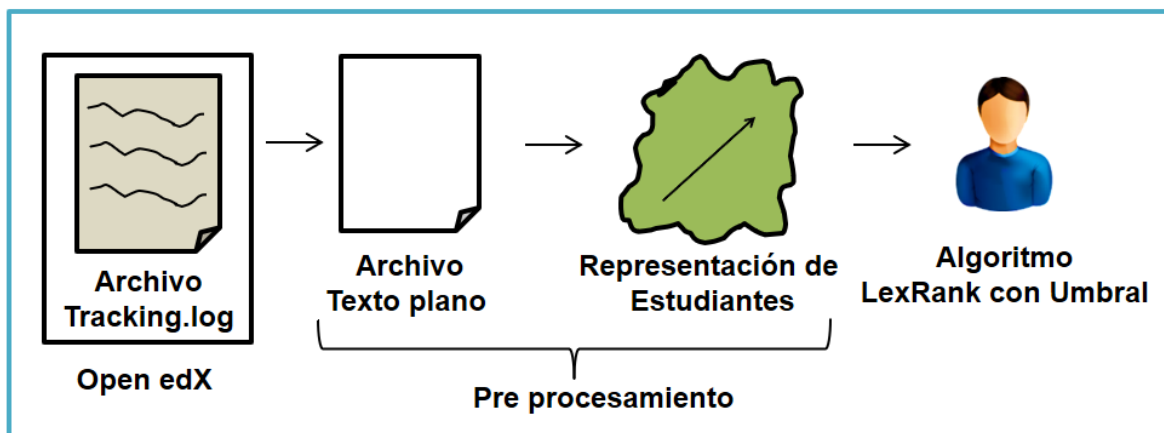


Figura 4. Aproximación arquitectónica para el descubrimiento de patrones de navegación. Tomado de (Anacona et al., 2017).

El uso de los Cursos en Línea Abiertos y Masivos, MOOC, está creciendo en los últimos años, debido a que diversas Universidades en el mundo los están ofreciendo. Esto ha permitido generar una gran cantidad de datos por las actividades desarrolladas de los estudiantes en los cursos. Estas actividades generan patrones que son registrados en el archivo tracking.log para xMOOC de la plataforma Open edX. Hasta el momento no hay un referente claro en la forma de procesar los patrones de navegación de este archivo. Por tal razón, en el documento proponen, caracterizar los patrones de navegación EF-ISF de los estudiantes en cursos xMOOC de la plataforma Open edX, a través del modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento; además, describen los pasos generales de cuatro algoritmos de diferentes metodologías para la identificación de los patrones de navegación EF-ISF. Como trabajo a futuro plantean la implementación de la propuesta con los diferentes algoritmos (F. Anacona et al., 2018).

El uso de la plataforma Open edX para ofrecer cursos xMOOC por diferentes Universidades en el mundo, ha llevado a un crecimiento en la participación de los estudiantes en los cursos, generando de esta forma un conjunto de patrones de navegación al interactuar con los xMOOC, los cuales se registran en el archivo tracking.log. Los autores proponen un primer estudio que identifica patrones de

navegación EF-ISF del conjunto de patrones, a través del algoritmo LexRank con Umbral en dos registros de interacción de un estudiante en curso xMOOC ofertado en la Universidad del Cauca, primer periodo académico del año 2017. Como trabajo a futuro proponen la implementación del Algoritmo LexRank con Umbral, para la identificación de patrones de navegación EF-ISF registrados en el archivo tracking.log del xMOOC de la plataforma Open edX (F. A. Anacona, M. Solarte, & G. R. González, 2018).

2.5 Generación Automática de Resúmenes de Texto

En esta subsección se describe una alternativa para generar resúmenes automáticos de múltiples documentos, con el modelo de espacio vectorial, la ponderación basada en la frecuencia relativa de un término y el algoritmo LexRank con Umbral (Anacona, Cobos, & Mendoza, 2015; Erkan & Radev, 2004).

2.5.1 Definición

Las técnicas que realizan generación automática de resúmenes de texto buscan crear en esencia un resumen, con una breve pero exacta representación del contenido de uno o varios documentos (Mendoza & Leon, 2015; Steinberger & Ježek, 2009). Los resúmenes deben conservar la información importante y deben ser cortos (Das & Martins, 2007) o estar sujetos a un tamaño definido por la persona(s) (Anacona et al., 2015).

2.6 Representación de los Documentos

2.6.1 Modelo de Espacio Vectorial

Un conjunto de documentos se puede representar como, $D = \{d_1, d_2, \dots, d_o\}$ donde o es el número de documentos. Cada documento d_j contiene un conjunto de oraciones $d_j = \{s_1, s_2, \dots, s_p\}$, en el que p es el número de oraciones en d_j y $j = 1, \dots, o$. De esta forma la colección de documentos se representa como el conjunto de todas las oraciones de la colección, es decir, $D = \{s_1, s_2, \dots, s_n\}$, s_j denota la oración j -ésima en D y n es el número de oraciones en la colección, $s_j \in D$ si y solo si $s_i \in d_j \in D$. Sea $T = \{t_1, t_2, \dots, t_m\}$ representa todos los términos que aparecen en D , donde m , es el número de los diferentes términos (Alguliev, Aliguliyev, Hajirahimova, & Mehdiyev, 2011; Anacona et al., 2015).

Por lo anterior, se considera el modelo de espacio vectorial de una oración \vec{s}_j , donde cada una es representada por uno o más términos t_i ; los términos pueden ser ponderados (w_{ij}) o no ponderados de acuerdo a su importancia. En este caso, cada oración \vec{s}_j es representada por un m -vector de la siguiente forma (Salton, Wong, & Yang, 1975):

$$\vec{s}_{mj} = (w_{1j}, w_{2j}, \dots, w_{ij}, \dots, w_{mj}) \quad (2-1)$$

2.6.2 Técnicas de ponderación de términos

2.6.2.1 Frecuencia del término

La frecuencia del término (TF) otorga mayor relevancia a los términos con mayor frecuencia, evaluando el número de veces que el término aparece en la oración, como se muestra en la ecuación (2-2) (Montiel Soto, García-Hernández, Ledeneva, & Cruz Reyes, 2009):

$$w_{ij} = f_{ij} \quad (2-2)$$

Donde f_{ij} es la frecuencia del término i en la oración j .

2.6.2.2 Frecuencia Inversa de la Oración

TF presenta problemas para definir el peso cuando un término aparece en casi todas las oraciones de la colección, por lo tanto, el término no es útil para determinar las oraciones relevantes (Montiel Soto et al., 2009). Como una alternativa surge la frecuencia inversa de la oración (ISF), que evalúa la importancia de los términos de una oración \vec{s}_j , definido por la ecuación (2-3) (Erkan & Radev, 2004):

$$w_{ij} = \log \frac{N}{n_j} \quad (2-3)$$

Donde N es el número total de oraciones y n_j es el número de oraciones en las que aparece el término i .

2.6.2.3 Ponderación basada en la frecuencia relativa de un término

La frecuencia relativa de un término (TF-ISF), combina las definiciones de TF e ISF para asignar un peso a cada término de la oración \vec{s}_j . Otorgando mayor relevancia a los términos menos frecuentes en el conjunto de oraciones, pero a su vez más frecuentes en la oración, como se observa en la siguiente ecuación (2-4) (Manning, Raghavan, & Schütze, 2009):

$$w_{ij} = (f_{ij}) \left(\log \frac{N}{n_j} \right) \quad (2-4)$$

Donde w_{ij} representa el peso del término i –ésimo de la oración \vec{s}_j .

2.6.3 Medidas de Similitud

Para el cálculo de similitud entre las oraciones, cada una de ellas se debe representar como un vector en el modelo de espacio vectorial (Alguliev et al., 2011).

Teniendo las oraciones como vectores es posible calcular su semejanza, por medidas de similitud (Salton et al., 1975).

2.6.3.1 Similitud de Coseno

Sean \vec{s}_i, \vec{s}_j dos m -vectores diferentes del vector cero, donde el peso de los términos se calcula con TF-ISF. Entonces el ángulo φ o la similitud entre \vec{s}_i y \vec{s}_j se define en el intervalo $[0, 1]$ de acuerdo a la ecuación (2-5) (Alguliev et al., 2011; Grossman, 2001):

$$\text{simcos}(\vec{s}_i, \vec{s}_j) = \frac{\sum_{k=1}^m w_{ki} w_{kj}}{\sqrt{\sum_{k=1}^m w_{ki}^2 \sum_{k=1}^m w_{kj}^2}} \quad (2-5)$$

2.6.4 Representación de Documentos por Medio de Matrices

La representación en el espacio multidimensional del conjunto de vectores de oraciones y la similitud de las mismas, se puede hacer con la Matriz de Términos por Oraciones y la Matriz de similitud de Cosenos.

2.6.4.1 Matriz de Términos por Oraciones

La matriz TF-ISF es una matriz dispersa de $m \times n$ pesos dispuestos en m términos (filas) y n oraciones (columnas). El elemento ij de la Matriz es denotado por w_{ij} corresponde al peso del término i en la oración j , este peso es calculado de acuerdo a la ecuación TF-ISF; la Matriz se representa por la siguiente ecuación (2-6) (Gong & Liu, 2001; Grossman, 2001):

$$\text{Matriz TF - ISF} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1j} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2j} & \dots & w_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ w_{i1} & w_{i2} & \dots & w_{ij} & \dots & w_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ w_{m1} & w_{m2} & \dots & w_{mj} & \dots & w_{mn} \end{pmatrix} \quad (2-6)$$

2.6.4.2 Matriz de Similitud de Cosenos

La matriz de similitud de cosenos es una matriz cuadrada de $n \times n$, donde n es el número de oraciones, el elemento ij de la matriz denotado por a_{ij} , representa la similitud entre las oraciones, calculada con la similitud de cosenos, lo anterior se representa de acuerdo a la ecuación (2-7) (Grossman, 2001):

$$MatrizDeSimilitud = \begin{pmatrix} 1_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & 1_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & 1_{ij} & \dots & a_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & 1_{nn} \end{pmatrix} \quad (2-7)$$

2.7 Algoritmo LexRank con Umbral para generación automática de resúmenes de múltiples documentos

En generación automática de resúmenes de múltiples documentos, el algoritmo LexRank con Umbral es de los más referenciados en métodos de grafos. El algoritmo permite ver un grupo de documentos como una red de oraciones relacionadas. Para generar el resumen automático, se realiza pre procesamiento en los conjuntos de datos que contienen los documentos a procesar: segmentación, construir las oraciones de los documentos; Tokenización, encontrar los términos en las oraciones y entre otros. Con lo cual, se dejan las oraciones de los documentos a procesar y seguidamente se hace una representación en el modelo de espacio vectorial con las matrices TF-ISF y similitud de cosenos (Anacona et al., 2015; Erkan & Radev, 2004).

En los pseudo-códigos siguientes se muestran el procesamiento a llevar a cabo para encontrar el resumen automático. El Algoritmo 1 se describe el pseudo-código llamado método de potencia, para calcular la distribución estacionaria, es decir la solución de la cadena de Markov (Anacona et al., 2015; Erkan & Radev, 2004).

Entrada: Una matriz M estocástica, irreducible y aperiódica

Entrada: Tolerancia de error ϵ

Salida: Vector propio p

01 $p_0 = \frac{1}{N} \mathbf{1}$

02 $t = 0$

03 repita

04 $t = t + 1$

05 $p_t = M^T p_{t-1}$

06 $\delta = ||p_t - p_{t-1}||$

07 hasta que $\delta < \epsilon$;

08 retorne p_t

Algoritmo 1. Método de Potencia para el cálculo de la distribución estacionaria de una Cadena de Markov, Adaptado de (Erkan & Radev, 2004).

El Algoritmo 2 resume cómo calcular las puntuaciones LexRank con Umbral para un determinado conjunto de oraciones (Anacona et al., 2015; Erkan & Radev, 2004).

```

Entrada: Arreglo S de n oraciones, umbral t, valor de amortiguamiento dampingFactor
Salida: Arreglo L con los scores definidos por LexRank para cada oración
Arreglo CosineMatriz[n][n];
Arreglo L[n];
01 Para i=1 hasta n haga
02     suma=0
03     Para j=1 hasta n haga
04         CosineMatriz[i][j] = idf-modified-cosine(S[i],S[j]);
05         Si CosineMatriz[i][j] > t haga
06             CosineMatriz[i][j] = 1;
07             suma++;
08         Si No
09             CosineMatriz[i][j] = 0;
10         Fin Si
11     Fin Para
15 Fin Para
16 Para i=1 hasta n haga
17     Para j=1 hasta n haga
18         CosineMatriz[i][j] = CosineMatriz[i][j] / suma;
19     Fin Para
20 Fin Para
21 Para i=1 hasta n haga
22     Para j=1 hasta n haga
23         CosineMatriz[i][j] = (dampingFactor/n) + (1- dampingFactor)* CosineMatriz[i][j];
24     Fin Para
25 Fin Para
26 L = PowerMethod(CosineMatriz, n, ε ); //Algoritmo 1 previamente presentado
27 retorne L;

```

Algoritmo 2. Cálculo de puntuación de LexRank con Umbral, Adaptado de (Erkan & Radev, 2004).

Al realizar pruebas con el conjunto de documentos DUC2005⁷ y evaluación automática con las medidas ROUGE-1⁸, ROUGE-2 y ROUGE-SU4, con los siguientes algoritmos: Máxima Cobertura Mínima Redundancia (MCMR*)⁹, método meta heurístico; Análisis Semítico Latente (LSA), método algebraico y LexRank con Umbral, método de grafo; se evidenció que el LexRank con un Umbral de 0.1 obtuvo mejores resultados en las métricas, como se observa en la Tabla 4; para analizar el rendimiento de los algoritmos se usó mejora relativa, como se muestra en la Tabla 5 (Anaconda et al., 2015).

La fórmula de mejora relativa se muestra en la ecuación (2-8).

⁷ La Conferencia de Entendimiento del Documento (Document Understanding Conference), ofrece un conjunto de documentos para realizar pruebas.

⁸ ROUGE (Recall-Oriented Understudy for Gisting Evaluation), es un estándar para la evaluación automática de resúmenes de textos

⁹ El MCMR* con el MCMR de la Tabla 6 son diferentes, porque generan el resumen automático de diferente manera.

$$\frac{(\text{método propuesto} - \text{otro método})}{\text{otro método}} \times 100 \quad (2-8)$$

En la Tabla 4, se muestran los resultados de la evaluación de los algoritmos con las medidas de ROUGE y el conjunto de datos DUC2005.

Tabla 4. Evaluación con ROUGE de los algoritmos con DUC2005. Tomado de (Anacona et al., 2015).

DU2005				
Algoritmo	ROUGE-1	ROUGE-2	ROUGE-SU4	Posición
LexRank con Umbral	0,3660	0,0633	0,1207	1
MCMR*	0,3320	0,0451	0,1014	2
LSA	0,3052	0,0361	0,0907	3

En la Tabla 5, se muestran los resultados de evaluación de los algoritmos con mejora relativa.

Tabla 5. Mejora en % de LexRank con Umbral (ver Tabla 4). Tomado de (Anacona et al., 2015).

Mejora Relativa %			
Algoritmo	ROUGE-1	ROUGE-2	ROUGE-SU4
LexRank con Umbral	-	-	-
MCMR*	10,2409	40,3547	19,0335
LSA	19,9213	75,3462	33,0760

La Tabla 5 permite observar que LexRank con Umbral es el mejor algoritmo, porque mejora el rendimiento en comparación con los otros algoritmos, en las tres medidas de ROUGE, también se puede clasificar el orden de los algoritmos de acuerdo a los resultados obtenidos de forma ascendente como se muestra en la Tabla 4, siendo el LexRank con Umbral el que ocupa la primera posición, para el conjunto de datos DUC2005 en las medidas de ROUGE (Anacona et al., 2015).

Por otra parte, en la Tabla 6 se muestra los resultados de los mejores dieciocho (18) algoritmos en generación automática de resúmenes de múltiples documentos, en el cual los resúmenes fueron generados del conjunto de datos DUC2005 y sus resultados se comparan con métricas de ROUGE (Alguliyev, Aliguliyev, & Isazade, 2015). En la tabla también se incluyen los resultados de los algoritmos LexRank con Umbral, MCMR* y LSA con el fin de comparar y determinar las posiciones que

ocupan con los algoritmos del estado del arte, con lo cual se tendrían veintiún (21) algoritmos con sus respectivos resultados. La tabla incluye el nombre del algoritmo, la métrica de ROUGE y la posición que ocupa de acuerdo al resultado de la métrica, estos son ordenados de forma ascendente, siendo la posición uno el algoritmo de mejor resultado.

Tabla 6. Resultados de ROUGE de los algoritmos con DUC2005. Adaptado de (Alguliyev et al., 2015).

Algoritmo	ROUGE-1	Pos	ROUGE-2	Pos	ROUGE-SU4	Pos
CDDS	0.3956	(1)	0.0852	(1)	0.1423	(1)
DESAMC + DocSum	0.3937	(2)	0.0822	(3)	0.1418	(2)
MR&MR-Sum (cos + overlap)	0.3932	(3)	0.0834	(2)	0.1411	(3)
MR&MR-Sum (Jaccard + overlap)	0.3925	(4)	0.0822	(4)	0.1397	(5)
MR&MR-Sum (cos)	0.3917	(5)	0.0816	(5)	0.1391	(7)
MR&MR-Sum (Jaccard)	0.3909	(6)	0.0808	(6)	0.1385	(8)
MR&MR-Sum (overlap)	0.3896	(8)	0.0804	(7)	0.1376	(9)
MCMR	0.3891	(9)	0.0790	(10)	0.1392	(6)
ILFP	0.3905	(7)	0.0804	(8)	0.1403	(4)
MMR	0.3479	(19)	0.0601	(19)	0.1134	(19)
iRANK	0.3880	(10)	0.0802	(9)	0.1373	(10)
Qs-MRC	0.3868	(11)	0.0779	(11)	0.1366	(11)
SVR	0.3849	(14)	0.0757	(12)	0.1335	(15)
TranSumm	0.3857	(13)	0.0755	(13)	0.1366	(12)
Biased LexRank	0.3861	(12)	0.0753	(14)	0.1363	(13)
QEA	0.3748	(17)	0.0749	(15)	0.1333	(16)
Content-term	0.3814	(15)	0.0718	(16)	0.1338	(14)
TMR + TF	0.3775	(16)	0.0715	(17)	0.1304	(17)
LexRank con Umbral	0,3660	(18)	0,0633	(18)	0,1207	(18)

MCMR*	0,3320	(20)	0,0451	(20)	0,1014	(20)
LSA	0,3052	(21)	0,0361	(21)	0,0907	(21)

Nomenclatura: Posición = Pos

Con base en los resultados de la Tabla 6, se puede concluir que el LexRank con Umbral ocupa la posición 18, el MCMR* ocupa la posición 20 y el LSA ocupa la posición 21, para ROUGE-1, ROUGE-2 y ROUGE-SU4 respectivamente. Con lo cual se puede concluir que el LexRank con Umbral esta entre los 18 mejores algoritmos para generación automática de resúmenes de múltiples documentos, para conjuntos de datos DUC2005 evaluadas con métricas de ROUGE.

2.8 Conclusiones del Marco Teórico

En este apartado se presentan las conclusiones de la documentación revisada en la sección 2, Marco Teórico.

- En la sección 2.1 se describió el LMS y el archivo tracking.log de la plataforma Open edX. El LMS, permite la interacción de estudiantes, docentes, profesores y entre otros, con la plataforma; entre tanto, el archivo tracking.log registra los eventos realizados por parte de los usuarios en el curso.
- En los artículos de patrones, navegación y patrones de navegación de los estudiantes en MOOC, secciones 2.2, 2.3 y 2.4 respectivamente, los diferentes autores realizan análisis en conjuntos de datos, que contienen los registros de las actividades realizadas por parte de los estudiantes con el material de aprendizaje del curso(s), ofertado por una plataforma MOOC. El análisis es realizado para encontrar diferentes patrones, navegaciones, patrones de navegación, a través de herramientas y/o métodos estadísticos y/o algoritmos.
- En la sección de patrones de navegación de los estudiantes en MOOC, sección 2.4, se describen propuestas para identificar patrones de navegación EF-ISF, como: propuesta arquitectónica para identificar patrones de navegación EF-ISF, caracterizar patrones de navegación EF-ISF a través del modelo de espacio vectorial y EF-ISF, además, la identificación del patrón de navegación EF-ISF a través del algoritmo LexRank con Umbral en dos registros de un estudiante. Las tres propuestas plantean como trabajo futuro la implementación de las mismas.
- En las secciones 2.5 y 2.7 se describe una forma de generar resúmenes automáticos para múltiples documentos, la cual consta de los siguientes pasos generales: pre procesamiento del conjunto de documentos, pesado de oraciones TF-ISF, representación en el modelo de espacio vectorial de los documentos y algoritmo(s), en el cual, el LexRank con Umbral esta entre los mejores algoritmos de generación automática de resúmenes de múltiples documentos, evaluados con conjuntos de datos DUC2005 y métricas de ROUGE.
- Los conjuntos de datos, que contienen registros de la interacción de los estudiantes en los MOOC, con los que realizan las diferentes investigaciones no son públicos.

- Los diferentes artículos de investigación, no evidencian pre procesamiento del conjunto de datos, que contiene las sesiones realizadas por el estudiante(s) al interactuar con el material de aprendizaje del MOOC, utilizando el modelo de espacio vectorial y EF-ISF.
- En las diferentes investigaciones realizadas en MOOC, no identifican la centralidad de los patrones de navegación EF-ISF por Sesión de los estudiantes, con el algoritmo LexRank con Umbral.
- En los artículos no se evidencia la identificación de la centralidad de los patrones de navegación EF-ISF por sesión que involucren, secciones, unidades y material de aprendizaje.
- En los artículos no se evidencia la identificación del patrón de navegación EF-ISF por sesión de mayor centralidad del estudiante.
- No hay un estándar para evaluar los patrones de navegación en MOOC.

Con base en lo anterior, a continuación, se describe la hipótesis del presente trabajo de investigación: el algoritmo LexRank con Umbral permitirá identificar los patrones de navegación EF-ISF por sesión de mayor centralidad de los estudiantes, en conjuntos de datos que contengan la interacción de los estudiantes de MPOC.

CAPITULO 3

En este capítulo se describe la aproximación arquitectónica para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios, la cual consta de los siguientes módulos: conjunto de datos, pre procesamiento, algoritmo LexRank con Umbral y patrón de navegación EF-ISF por sesión de mayor centralidad.

3 APROXIMACIÓN ARQUITECTÓNICA PARA LA IDENTIFICACIÓN DE LOS PATRONES DE NAVEGACIÓN EF-ISF POR SESIÓN DE MAYOR CENTRALIDAD DE LOS USUARIOS.

En el presente trabajo de investigación, se tiene como objetivo identificar los patrones de navegación EF-ISF por sesión de mayor centralidad de los estudiantes. Por consiguiente, se adapta la propuesta para el Descubrimiento de patrones de navegación en Open edX – una aproximación arquitectónica, la cual consta de los siguientes módulos: archivo tracking.log, pre procesamiento y algoritmo LexRank con Umbral (Anacona et al., 2017), porque no se va a trabajar con archivos tracking.log, si no, con conjuntos de datos que se encuentran en formato Excel, los cuales contienen la interacción de los estudiantes, administradores, profesor y monitor con el MPOC de la plataforma Open edX, por tal razón, los módulos a desarrollar son los siguientes: conjunto de datos, pre procesamiento de datos, algoritmo LexRank con Umbral y los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios. Esto se muestra en la Figura 5.

El uso del algoritmo LexRank con Umbral, el cual pertenece a los métodos de grafos, es porque está entre los mejores algoritmos de generación automática de resúmenes de múltiples documentos, ver sección 2.7.

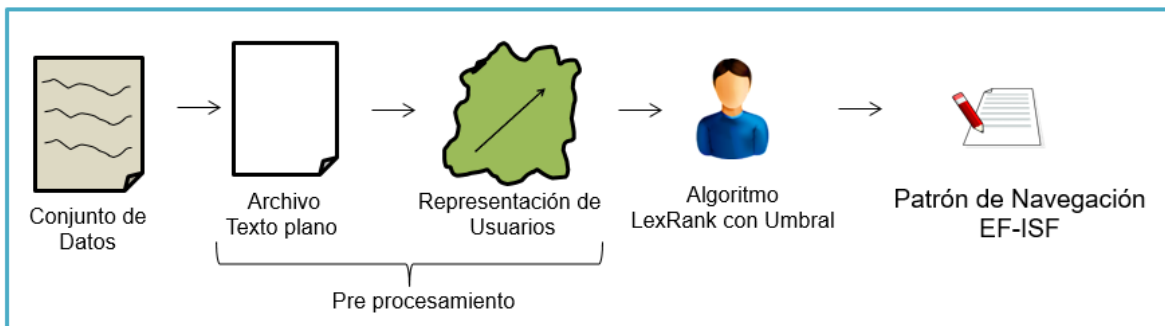


Figura 5. Aproximación arquitectónica para la identificación de la centralidad de los patrones de navegación EF-ISF por sesión. Adaptado de (Anacona et al., 2017).

Cada uno de los elementos de la Figura 5, se describen en las siguientes secciones.

3.1 Conjuntos de Datos

Los conjuntos de datos del MPOC son obtenidos, del curso Astronomía Cotidiana el cual se oferto en el año 2017 primer periodo académico de la Universidad del Cauca, el curso se dividió en Grupo A y Grupo B, ambos cursos fueron soportados por una instancia de la plataforma de Open edX en un servidor con nombre Selene. Los conjuntos de datos contienen algunas interacciones con el curso respectivo, las interacciones fueron realizadas por usuarios, estos pueden ser: estudiante, administrador, docente y monitor. Esta información fue registrada en archivos Excel (Jaramillo & Mario, 2017).

Además, de los archivos del Grupo A y B se generaron otros dos archivos de la Unidad Uno Temática Uno –U1T1 a la Unidad Uno Temática Cinco –U1T5 respectivamente, porque en ellos hay más del 50% de los datos de cada grupo. La forma en cómo se generan el archivo Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5 y Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5 se describe en las secciones 3.2.1.2 y 3.2.2.2 respectivamente. En la Tabla 7, se muestra una descripción de los conjuntos datos.

Tabla 7. Descripción de los Conjuntos de Datos. Adatado de (Jaramillo & Mario, 2017).

Conjunto de datos del curso	Administradores	Profesores	Monitores	estudiantes	Registros	Columnas
Astronomía Cotidiana 2017-I Grupo A	1	1	0	187	39610	11
Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5	1	1	0	168	19788	11
Astronomía Cotidiana 2017-I Grupo B	2	1	1	263	65582	11
Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5	1	1	1	241	32866	11

Los conjuntos de datos contienen los mismos once campos en la primera fila, estos se describen en la Tabla 8.

Tabla 8. Descripción de campos de la primera fila de los conjuntos de datos. Adatado de (Jaramillo & Mario, 2017).

Campos	Descripción
Idregistros	Identificador único de la actividad realizada por el usuario.
Usuario	Nombre y apellidos del usuario.
Correo	Correo del usuario.
Curso	Nombre del curso.
Hora	Hora en la que el usuario realizo la actividad.
Fecha	Fecha en la que el usuario realizo la actividad.
Sesión	Identificador único de 32 caracteres para la sesión del usuario generado por la plataforma.
Actividad	Contiene los nombres de las actividades realizadas por el usuario en la plataforma, ver Tabla 9.
Complemento1	El contenido del campo varia acorde a la actividad, ver Tabla 9.
Complemento2	El contenido del campo varia acorde a la actividad, ver Tabla 9.
Complemento3	El contenido del campo varia acorde a la actividad, ver Tabla 9.

El contenido de los campos Actividad, Complemento1, Complemento2 y Complemento3, de cada fila de los archivos Excel varia acorde a la actividad realizada por el usuario en el curso respectivo. En la Tabla 9 se describen las actividades desarrolladas por los usuarios.

Tabla 9. Descripción de actividades. Adatado de (Jaramillo & Mario, 2017).

Actividad	Complemento1	Complemento2	Complemento3
play_video: La acción de reproducir un video desde el principio se registra cuando: el usuario inicia un video por primera vez, el usuario se devuelve al inicio del video, el video terminó completamente y el usuario lo volvió a reproducir desde cero.	Sección en la que ingresó el usuario.	Unidad en la que ingresó el usuario.	Identificador del video de YouTube o Nombre del video en YouTube. Aparece el nombre cuando se cambia el identificador por el nombre respectivo, en el proceso de limpieza y

			adecuación de los conjuntos de datos.
pause_video: Las pausas que el usuario realizó durante la visualización del video.	Sección en la que ingresó el usuario.	Unidad en la que ingresó el usuario.	Identificador del video de YouTube o Nombre del video en YouTube. Aparece el nombre cuando se cambia el identificador por el nombre respectivo, en el proceso de limpieza y adecuación de los conjuntos de datos.
stop_video: Se registra cuando el video llega hasta el final.	Sección en la que ingresó el usuario.	Unidad en la que ingresó el usuario.	Identificador del video de YouTube o Nombre del video en YouTube. Aparece el nombre cuando se cambia el identificador por el nombre respectivo, en el proceso de limpieza y adecuación de los conjuntos de datos.
problem_check: Se registra cuando el usuario realiza click en el botón "enviar" para enviar las respuestas de un examen.	Sección en la que ingresó el usuario.	Unidad en la que ingresó el usuario.	NULL
edx.forum.comment.created: Comentario hecho a una publicación que ya ha tenido una respuesta en un foro. Es el tercer nivel, teniendo en cuenta. Foro -> Respuesta -> Comentario.	Título de la publicación principal en el foro.	NULL	NULL
edx.forum.response.created: Respuesta hecha a una publicación en un foro. Es el segundo nivel, teniendo en cuenta. Foro -> Respuesta -> Comentario.	Sintitulo	NULL	NULL
edx.forum.thread.created: Es una publicación en un foro. Es el primer	Título de la publicación	NULL	NULL

nivel, teniendo en cuenta. Foro -> respuesta -> Comentario.	principal en el foro.		
edx.forum.thread.voted: Son los votos que los usuarios dan a buenas publicaciones en el foro y respuestas.	A quién se le dio un voto.	NULL	NULL
Contenidos: Es la navegación en el curso. Por ejemplo: cuando un usuario pasa de la Sección x a la Unidad y, esto se registra como "Contenidos".	Sección en la que ingresó el usuario.	Unidad en la que ingresó el usuario.	NULL

3.2 Pre procesamiento

El Pre procesamiento se realiza en los datos sin procesar, esta etapa tiene como objetivo, limpiar los datos recolectados de información que no aporte y preparar los datos para el análisis por el LexRank con Umbral (Anacona et al., 2017), esto se ilustra en la Figura 6, en las siguientes sub secciones se describirá el proceso.

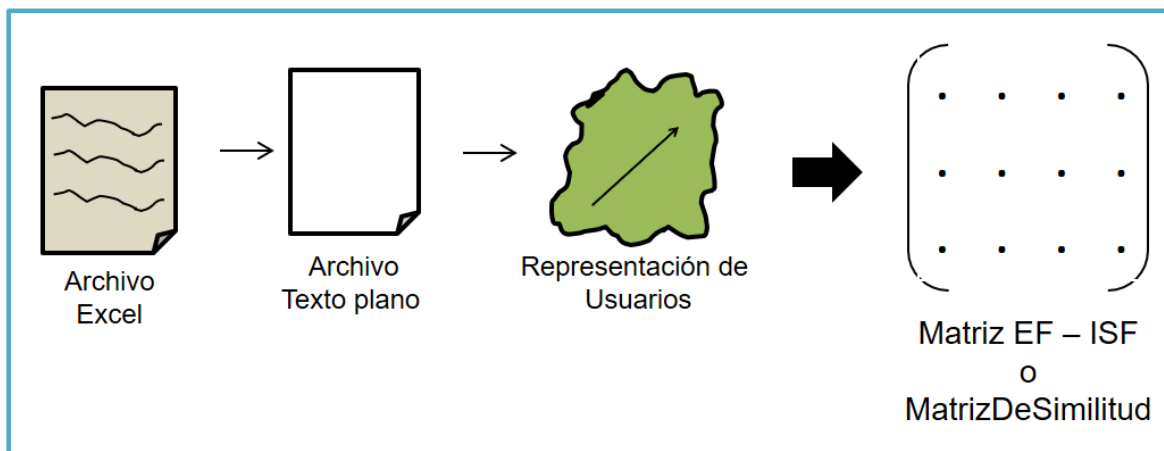


Figura 6. Transformar el conjunto de datos al modelo de espacio vectorial. Adaptado de (Anacona et al., 2017).

3.2.1 Conjuntos de Datos Grupo A

En las siguientes dos sub secciones describe la limpieza y una primera adecuación del conjunto de datos, Astronomía Cotidiana 2017-I Grupo A y la generación del conjunto de datos Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5, este primer tratamiento de los datos se realiza en el archivo Excel.

3.2.1.1 Astronomía Cotidiana 2017-I Grupo A

El conjunto de datos Astronomía Cotidiana 2017-I Grupo A se encuentra en formato Excel, contiene interacciones de los usuarios con el curso Astronomía Cotidiana

2017-I Grupo A. El archivo contiene 39610 registros incluyendo la primera fila que contiene el nombre de cada campo, en la sección 3.1 se describe en mayor detalle la información. En la Tabla 10 se muestra la información del archivo.

Tabla 10. Descripción de los Conjuntos de Datos Grupo A. Adatado de (Jaramillo & Mario, 2017).

Conjunto de datos del curso	Administradores	Profesores	estudiantes	Registros	Columnas
Astronomía Cotidiana 2017-I Grupo A	1	1	181	39610	11

Al archivo Astronomía Cotidiana 2017-I Grupo A, se le realizó un pre procesamiento en la columna de Complemento 1 y Complemento 3. La columna de complemento 1, contiene algunos textos con acentos y signos de interrogación, los cuales estaban codificados en Notación Unicode Java, estos fueron reemplazados por los caracteres en latino, en la Tabla 11 se muestran los caracteres que fueron reemplazados.

Tabla 11. Codificación de Caracteres Grupo A. Adaptado de (Merino, 2012)

Codificación Latino	Codificación Java
Á	\u00c1
É	\u00c9
Í	\u00cd
Ó	\u00d3
Ú	\u00da
á	\u00e1
é	\u00e9
í	\u00ed
ó	\u00f3
ú	\u00fa
¿	\u00bf

En el archivo, en la columna de complemento 3, en algunos registros cuando el usuario uso el video aparecen los identificadores de YouTube, estos fueron

reemplazados por los nombres respectivos. En Tabla 12 se realiza una descripción de los identificadores.

Tabla 12. Descripción de Identificadores de YouTube de los Cursos. Fuente Propia.

Conjunto de datos del curso	Total Identificadores YouTube	Total de Identificadores reemplazados por el nombre de video	Total de Identificadores no disponibles en la Plataforma YouTube	Identificadores de los videos no disponibles en la plataforma YouTube
Astronomía Cotidiana 2017-I Grupo A	108	104	4	IJ65SZqdaWI afP-P4TjEq0 80dR8vjZu9w pCeyjPuMnuw

Al momento de la consulta de los identificadores de los videos para obtener los nombres en la plataforma de YouTube, no se encontraban disponibles cuatro identificadores, por esta razón se dejaron estos identificadores en el conjunto de datos; sin embargo, los nombres de 104 identificadores encontrados se reemplazaron en el Conjuntos de Datos.

Al archivo Excel se le realiza un pre procesamiento para eliminar registros que no se relacionan con los demás registros, esto se describe en la Tabla 13.

Tabla 13. Registros eliminados del Excel Grupo A. Fuente propia.

Eliminados	Numero Registros	Motivo
Ingresos	4119	Los ingresos no tienen sesión asignada, el valor es NULL, lo que indica ingresos fallidos a la plataforma.
Courseware	81	La información no se puede relacionar con los demás registros del archivo Excel. No indica en que parte del curso se encuentra el usuario.
Total Registros eliminados	4200	

Registro total del archivo Excel	39610 – 4200 = 35410
Estudiantes	Al eliminar los registros de Ingresos y Courseware, tuvo como consecuencia la eliminación de 6 estudiantes. Lo que permite concluir, la información contenida en el conjunto de datos sobre estos estudiantes no tenía relación, lo que afectaría la identificación del patrón de navegación EF-ISF por sesión de mayor centralidad.

El archivo Excel con el que se trabaja, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad, tiene un total de 35410 registros, incluido el encabezado que es la primera fila. En la Tabla 14 se muestra la distribución de la información.

Tabla 14. Descripción del conjunto de datos del Grupo A. Fuente Propia.

Conjunto de datos del curso	Usuarios	Sesiones	Registros
Astronomía Cotidiana 2017-I Grupo A	1 Administrador	8	37
	1 Profesor	64	197
	181 Estudiantes	2838	35175
	Total	2910	35409
	Columnas	11	
	Promedio Sesiones por usuario	15	
	Registros total del archivo Excel	35409 + 1 del encabezado = 35410	

3.2.1.2 Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5

El archivo Excel Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5 es generado a partir del archivo Astronomía Cotidiana 2017-I Grupo A, porque de la Unidad Uno Temática Uno –U1T1 a la Unidad Uno Temática Cinco –U1T5, hay un total de sesiones, 1617 y registros, 19787 que representan el 55.57% y 55.88%

respectivamente de sesiones y registros del Grupo A, cuentan con más del 50% de las interacciones de los estudiantes con el curso.

El archivo Excel con el que se trabaja, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad, tiene un total de 19788 registros, incluido el encabezado que es la primera fila. En la Tabla 15 se muestra la distribución de la información.

Tabla 15. Descripción del conjunto de datos del Grupo A U1T1 a U1T5. Fuente Propia.

Conjunto de datos del curso	Usuarios	Sesiones	Registros
Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5	1 Administrador	2	15
	1 Profesor	28	87
	168 Estudiantes	1587	19685
	Total	1617	19787
	Columnas	11	
	Promedio Sesiones por usuario	10	
	Registros total del archivo Excel	19787 + 1 del encabezado = 19788	

3.2.2 Conjuntos de Datos Grupo B

En las siguientes dos sub secciones describe la limpieza y una primera adecuación del conjunto de datos, Astronomía Cotidiana 2017-I Grupo B y la generación del conjunto de datos Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5, este primer tratamiento de los datos se realiza en el archivo Excel.

3.2.2.1 Astronomía Cotidiana 2017-I Grupo B

El conjunto de datos Astronomía Cotidiana 2017-I Grupo B se encuentra en formato Excel, contiene interacciones de los usuarios con el curso Astronomía Cotidiana 2017-I Grupo B. El archivo contiene 65582 registros incluyendo la primera fila que contiene el nombre de cada campo, en la sección 3.1 se describe en mayor detalle la información. En la Tabla 16 se muestra la información del archivo.

Tabla 16. Descripción de los Conjuntos de Datos Grupo B. Adatado de (Jaramillo & Mario, 2017).

Conjunto de datos del curso	Administradores	Profesores	Monitores	estudiantes	Registros	Columnas
Astronomía Cotidiana 2017-I Grupo B	2	1	1	255	65582	11

Al archivo Astronomía Cotidiana 2017-I Grupo B, se le realizó un pre procesamiento en la columna de Complemento 1 y Complemento 3. En la columna de complemento 1, contiene algunos textos con acentos y signos de interrogación, los cuales estaban codificados en Notación Unicode Java, estos fueron remplazados por los caracteres en latino, en la Tabla 17 se muestran los caracteres que fueron remplazados.

Tabla 17. Codificación de Caracteres Grupo B. Adaptado de (Merino, 2012)

Codificación Latino	Codificación Java
Á	\u00c1
É	\u00c9
Í	\u00cd
Ó	\u00d3
Ú	\u00da
á	\u00e1
é	\u00e9
í	\u00ed
ó	\u00f3
ú	\u00fa
¿	\u00bf

En el archivo, en la columna de complemento 3, en algunos registros cuando el usuario uso el video aparecen los identificadores de YouTube, estos fueron remplazados por los nombres respectivos. En Tabla 18 se realiza una descripción de los identificadores.

Tabla 18. Descripción de Identificadores de YouTube del Grupo B. Fuente Propia.

Conjunto de datos del curso	Total Identificadores YouTube	Total de Identificadores reemplazados por el nombre de video	Total de Identificadores no disponibles en la Plataforma YouTube	Identificadores de los videos no disponibles en la plataforma YouTube
Astronomía Cotidiana 2017-I Grupo B	110	106	4	IJ65SZqdaWI afP-P4TjEq0 80dR8vjZu9w pCeyjPuMnuw

Al momento de la consulta de los identificadores de los videos para obtener los nombres en la plataforma de YouTube, no se encontraban disponibles cuatro identificadores, por esta razón se dejaron estos identificadores en el conjunto de datos; sin embargo, los nombres de 106 identificadores encontrados se reemplazaron en el Conjuntos de Datos.

Al archivo Excel se le realiza un pre procesamiento para eliminar registros que no se relacionan con los demás registros, esto se describe en la Tabla 19.

Tabla 19. Registros eliminados del Excel Grupo B. Fuente propia.

Eliminados	Numero Registros	Motivo
Ingresos	5888	Los ingresos no tienen sesión asignada, el valor es NULL, lo que indica ingresos fallidos a la plataforma.
Courseware	152	La información no se puede relacionar con los demás registros del archivo Excel. No indica en que parte del curso se encuentra el usuario.
Total Registros eliminados	6040	
Registro total del archivo Excel	65582 – 6040 = 59542	

Estudiantes	Al eliminar los registros de Ingresos y Courseware, tuvo como consecuencia la eliminación de 8 estudiantes. Lo que permite concluir, la información contenida en el conjunto de datos sobre estos estudiantes no tenía relación, lo que afectaría la identificación del patrón de navegación EF-ISF por sesión de mayor centralidad.
--------------------	--

El archivo Excel con el que se trabaja, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad, tiene un total de 59542 registros, incluido el encabezado que es la primera fila. En la Tabla 20 se muestra la distribución de la información.

Tabla 20. Descripción del conjunto de datos del Grupo B. Fuente Propia.

Conjunto de datos del curso	Usuarios	Sesiones	Registros
Astronomía Cotidiana 2017-I Grupo B	1 Administrador A	11	33
	1 Administrador B	2	5
	1 Profesor	56	140
	1 Monitor	19	54
	255 Estudiantes	3940	59309
	Total	4028	59541
	Columnas	11	
	Promedio Sesiones por usuario	16	
	Registros total del archivo Excel	59541 + 1 del encabezado = 59542	

3.2.2.2 Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5

El archivo Excel Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5 es generado a partir del archivo Astronomía Cotidiana 2017-I Grupo B, porque de la Unidad Uno Temática Uno –U1T1 a la Unidad Uno Temática Cinco –U1T5, hay un total de sesiones, 2236 y registros, 32865 que representan el 55.51% y 55.20%

respectivamente de sesiones y registros del Grupo B, cuentan con más del 50% de las interacciones de los estudiantes con el curso.

El archivo Excel con el que se trabaja, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad, tiene un total de 32866 registros, incluido el encabezado que es la primera fila. En la Tabla 21 se muestra la distribución de la información.

Tabla 21. Descripción del conjunto de datos del Grupo B U1T1 a U1T5. Fuente Propia.

Conjunto de datos del curso	Usuarios	Sesiones	Registros
Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5	1 Administrador A	1	1
	1 Profesor	10	20
	1 Monitor	11	43
	241 Estudiantes	2214	32801
	Total	2236	32865
	Columnas	11	
	Promedio Sesiones por usuario	9	
	Registros total del archivo Excel	32865 + 1 del encabezado = 32866	

3.2.3 Archivo de texto plano

En un curso, una sesión se define como un período de actividad continua con el LMS por parte de un usuario(s) durante un tiempo máximo de 45 minutos (Palta & Vásquez, 2016). En el presente trabajo de investigación se tendrán en cuenta todas las sesiones generadas por los usuarios, sin importar el tiempo de interacción registrado en el conjunto de datos, archivos Excel.

El archivo tracking.log tiene un campo llamado "session", como se describió en la sección 2.1.1, el campo "session" es de tipo String y registra una cadena de 32 caracteres de la sesión del usuario, es decir, cuando un estudiante(s) ingresa a la plataforma sin ningún inconveniente con las credenciales respectivas, nombre de usuario y contraseña, la plataforma le asigna una cadena de 32 caracteres al campo

“session”. La cadena es una clave que identifica la sesión del usuario(s) de manera única (edX, 2016). El identificador de sesión del usuario en los archivos Excel está registrado en la columna de Sesión, ver sección 3.1.

Ahora bien, en este proceso de limpieza y adecuación, el archivo Excel, sección 3.2.1 o 3.2.2, se procesa para separar los usuarios en archivos de texto plano con sus respectivos registros, organizándolos por sesiones, se crean los archivos de sesiones por usuario. Con este proceso se elimina la estructura de formato del archivo Excel. Cada archivo de sesiones por usuario, se procesa con la ponderación basada en frecuencia relativa de un evento y se representa en el modelo de espacio vectorial con la Matriz de similitudes, el proceso de transformación al modelo de espacio vectorial se describe en la sección 3.2.4.

3.2.4 Representación de Usuarios

En este documento se propone trabajar con el Modelo de Espacio Vectorial (Salton et al. 1975), porque permite encontrar la relación de los eventos de navegación de los usuarios en un curso MPOC (Anacona et al., 2017).

3.2.4.1 Modelo de Espacio Vectorial

Un usuario se puede representar como un conjunto de sesiones, $u_i = \{s_1, s_2, \dots, s_n\}$ donde s_j corresponde a la sesión j -ésima del usuario i y n es el número total de sesiones del usuario i . Además, cada s_j de u_i se puede representar como un conjunto de eventos, $s_j = \{e_1, e_2, \dots, e_m\}$, donde e_i corresponde al evento i -ésimo de la sesión s_j y m es el número total de eventos de la sesión j .

Por lo anterior, se considera el modelo de espacio vectorial de una sesión \vec{s}_j , donde cada una es representada por uno o más eventos e_i ; los eventos pueden ser ponderados (w_{ij}) o no ponderados de acuerdo a su importancia. Esto se representa por un m -vector de la siguiente forma, ver ecuación (3-1) (Anacona et al., 2017; Salton et al., 1975):

$$\vec{s}_{mj} = (w_{1j}, w_{2j}, \dots, w_{ij}, \dots, w_{mj}) \quad (3-1)$$

3.2.5 Ponderación basada en la frecuencia relativa de un evento

Para el peso o ponderación de los eventos de la sesión \vec{s}_j , se usa la Ponderación basada en la Frecuencia Relativa de un Evento (EF-ISF), la cual combina las definiciones de frecuencia del evento ($EF = f_{ij}$) y frecuencia inversa de la sesión ($ISF = \log \frac{N}{n_j}$), en el que N es el número total de sesiones y n_j es el número de sesiones en el que aparece el evento i . Con lo cual, se otorga mayor relevancia a los eventos menos frecuentes en el conjunto de sesiones; pero a su vez más frecuentes en cada sesión, como se observa en la ecuación (3-2), donde w_{ij}

representa el peso del evento i –ésimo de la sesión \vec{s}_j (Anacona et al., 2017; Manning et al., 2009):

$$w_{ij} = (f_{ij}) \left(\log \frac{N}{n_j} \right) \quad (3-2)$$

3.2.6 Medida de Similitud

En el cálculo de similitud entre las sesiones, estas se deben representar como vectores en el modelo de espacio vectorial. Teniendo las sesiones como vectores es posible calcular su semejanza, con medidas de similitud, como la de coseno u otras. Con similitud de cosenos sería: sean \vec{s}_i, \vec{s}_j dos m -vectores diferentes del vector cero, donde el peso de los eventos se calcula con EF-ISF (Ver ecuación (3-2)). Entonces el ángulo ϕ o la similitud entre \vec{s}_i y \vec{s}_j se define en el intervalo $[0, 1]$ de acuerdo a la ecuación (3-3) (Alguliev et al., 2011; Anacona et al., 2017; Grossman, 2001; Singhal, 2001):

$$\text{simcos}(\vec{s}_i, \vec{s}_j) = \frac{\sum_{k=1}^m w_{ki} w_{kj}}{\sqrt{\sum_{k=1}^m w_{ki}^2 \sum_{k=1}^m w_{kj}^2}} \quad (3-3)$$

3.2.7 Representación del conjunto de datos por medio de Matrices

La representación en el espacio multidimensional del conjunto de vectores de sesiones y la similitud de los mismos, se puede hacer con la Matriz de Eventos por Sesiones y la Matriz de similitud de Cosenos.

3.2.7.1 Matriz de Eventos por Sesiones

La *Matriz EF – ISF* $_{m \times n}$ es una matriz dispersa de pesos dispuestos en m eventos (filas) y n sesiones (columnas), el elemento ij de la Matriz es denotado por w_{ij} , que corresponde al peso del evento i en la sesión j , este peso es calculado de acuerdo a la ecuación (3-2), esta se muestra en la ecuación (3-4) (Anacona et al., 2017; Gong & Liu, 2001; Grossman, 2001).

$$\text{Matriz EF – ISF} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1j} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2j} & \dots & w_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ w_{i1} & w_{i2} & \dots & w_{ij} & \dots & w_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ w_{m1} & w_{m2} & \dots & w_{mj} & \dots & w_{mn} \end{pmatrix} \quad (3-4)$$

3.2.7.2 Matriz de Similitud de Cosenos

La *MatrizDeSimilitud*_{*n*×*n*} de cosenos es una matriz cuadrada, donde *n* es el número de sesiones, el elemento *ij* de la matriz denotado por *a_{ij}*, se calcula con la ecuación (3-3), es la similitud entre la sesión \vec{s}_i y \vec{s}_j que aparece en la fila *i* y la columna *j*, los elementos de diagonal principal son iguales a 1, ya que es el cálculo de cada sesión con el mismo, \vec{s}_i con \vec{s}_i cuando *i* = *j*, esta se muestra en la ecuacion (3-5) (Anacona et al., 2017; Gong & Liu, 2001; Grossman, 2001).

$$MatrizDeSimilitud = \begin{pmatrix} 1_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & 1_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & 1_{ij} & \dots & a_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & 1_{nn} \end{pmatrix} \quad (3-5)$$

3.3 Algoritmo

Al tener las sesiones de los usuarios en matrices de similitudes, ecuación (3-5), estas son procesadas por el algoritmo LexRank con Umbral, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad, como se ilustra en la Figura 7. (F. Anacona et al., 2018; Anacona et al., 2017).

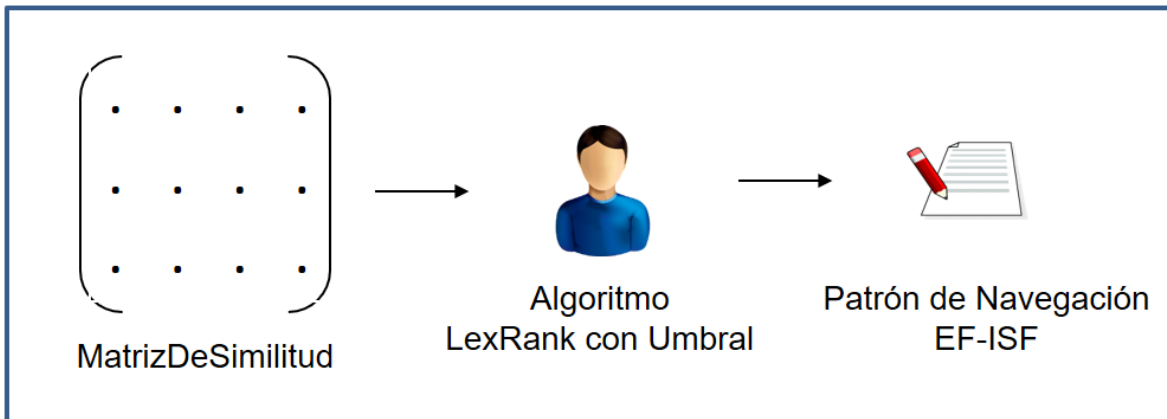


Figura 7. Algoritmo para la identificación de patrones de navegación EF-ISF por sesión de mayor centralidad. Adaptado de (Anacona et al., 2017).

3.3.1 Algoritmo LexRank con Umbral

El Algoritmo LexRank con Umbral es basado en el concepto de las redes sociales. Una red social es un mapa de las relaciones entre las entidades (usuarios, organizaciones y documentos, entre otros) que interactúan. Las redes sociales se representan comúnmente en forma de grafos, donde los nodos representan las entidades y los enlaces representan las relaciones entre los nodos (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

Ahora bien, un usuario puede verse como una red de sesiones relacionadas; algunas son más similares entre sí, mientras otras pueden compartir poca información con el resto de las sesiones. Si una sesión es muy similar a las demás sesiones, esta se puede considerar como la más central. Es por ello que hay dos puntos clave a soportar esta definición de centralidad, primero, cómo definir la similitud entre dos sesiones y segundo, cómo calcular la centralidad de una sesión dada su similitud con otras sesiones (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

Para definir la similitud, inicialmente se representa el conjunto de sesiones en el modelo de espacio vectorial, Matriz EF-ISF, ecuación (3-4). La similitud entre dos sesiones se define por la similitud de coseno que se muestra en la ecuación (3-3); luego el conjunto de sesiones se representa como un grafo a través de una matriz de adyacencia, MatrizDeSimilitud ecuación (3-5), cada valor de la matriz corresponde a la similitud de coseno entre dos sesiones (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

Posteriormente, para la centralidad, el nodo sesión debe tener en cuenta los votos de cada nodo y de donde vienen esos votos. Esto se puede considerar, si cada nodo tiene un valor de centralidad distribuido entre el nodo mismo y sus vecinos; como se muestra en la ecuación (3-6) (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

$$p(u) = \sum_{v \in adj[u]} \frac{p(v)}{\deg(v)} \quad (3-6)$$

Donde $p(u)$ es la centralidad del nodo u , $adj[u]$ es el conjunto de nodos que son adyacentes a u , y $\deg(v)$ es el grado del nodo v . Esta ecuación se puede escribir en notación matricial de la siguiente manera, ver ecuación (3-7) (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

$$p = B^T p, \text{ que es igual a: } p^T B = p^T \quad (3-7)$$

Donde la matriz B se obtiene de la matriz de adyacencia del grafo de similitud dividiendo cada elemento por la suma de la fila correspondiente, ver ecuación (3-8) (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

$$B(i, j) = \frac{A(i, j)}{\sum_k A(i, k)} \quad (3-8)$$

La suma de una fila es igual al grado del nodo correspondiente y cada sesión es similar al menos a sí misma, por tal razón, las sumas de las filas son distintas de

cero. La ecuación (3-7) establece que p^T es el vector propio izquierdo de la matriz B con el valor propio correspondiente a 1, para garantizar la existencia de un vector propio que pueda ser identificado y calculado de forma única, se necesita tener en cuenta lo siguiente (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

Una matriz estocástica X , es la matriz de transición de una cadena de Markov, un elemento $X(i, j)$ especifica la probabilidad de transición de un estado i a un estado j . Por los axiomas de probabilidad, todas las filas de una matriz estocástica deben sumar 1. $x^n(i, j)$, es la probabilidad del estado i para alcanzar el estado j en n transiciones. Una cadena de Markov con la matriz estocástica X converge a una distribución estacionaria sí (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

$$\lim_{n \rightarrow \infty} X^n = 1^T r \quad (3-9)$$

Donde $1 = (1, 1, \dots, 1)$ y el vector r se llama la distribución estacionaria de la cadena de Markov; cada elemento del vector r da la probabilidad asintótica de terminar en el estado correspondiente a largo plazo, independientemente del estado de partida. Una cadena de Markov es irreducible si cualquier estado es accesible desde cualquier otro estado, es decir, para todos i, j existe un n de tal manera que $X^n(i, j) \neq 0$. Una cadena de Markov es aperiódica si para todo i , $\gcd \{n : X^n(i, i) > 0\} = 1$. Por el teorema de Perron-Frobenius, una cadena de Markov irreducible y aperiódica converge a una distribución estacionaria única (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

Dado que la matriz de similitud B en la ecuación (3-7) satisface las propiedades de una matriz estocástica, se puede tratar como una cadena de Markov. El vector de centralidad P corresponde a la distribución estacionaria de B . Sin embargo, se debe asegurar que la matriz de similitud sea irreducible y aperiódica. Para resolver esto, se guarda una baja probabilidad para saltar a cualquier nodo en el grafo, lo que hace que el grafo sea irreducible y aperiódico. Si se asigna una probabilidad uniforme para saltar a cualquier nodo en el grafo, se obtiene la siguiente versión modificada de la ecuación (3-6), que se conoce como el algoritmo LexRank con Umbral, ver ecuación (3-10) (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

$$p(u) = \frac{d}{N} + (1 - d) \sum_{v \in \text{adj}[u]} \frac{p(v)}{\text{deg}(v)} \quad (3-10)$$

Donde N es el número total de nodos en el grafo y d es un “factor de amortiguamiento”, que por lo general se elige en el intervalo $[0.1, 0.2]$. La ecuación

(3-10), se puede escribir en forma matricial como se muestra en la ecuación (3-11) (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

$$p = [dU + (1 - d)B]^T p \quad (3-11)$$

Donde U es una matriz cuadrada de $N \times N$ con todos los elementos iguales a $\frac{1}{N}$. El núcleo de transición $[dU + (1 - d)B]$ de la cadena de Markov resultante es una mezcla de dos núcleos U y B . Un recorrido aleatorio en esta cadena de Markov elige uno de los estados adyacentes del estado actual con probabilidad $1 - d$, o salta a cualquier estado en el grafo, incluyendo el estado actual, con probabilidad d (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

A continuación, se muestra el pseudo-código del LexRank con Umbral para el cálculo de centralidad de las sesiones del usuario en un conjunto de sesiones, es decir, la centralidad de los patrones de navegación EF-ISF por sesión del usuario, esto se describe en el Algoritmo 3 (Anacona et al., 2015; Anacona et al., 2017; Erkan & Radev, 2004).

Entrada: Matriz MatrizDeSimilitud de n sesiones, umbral t, valor de amortiguamiento dampingFactor.
Salida: Matriz L[n] con la centralidad de los patrones de navegación ef-isf por sesión.

```

01 Para i=1 hasta n haga
02     grado=0
03     Para j=1 hasta n haga
04         Si MatrizDeSimilitud[i][j] > t haga
05             MatrizDeSimilitud[i][j] = 1;
06             grado ++;
07         Si No
08             MatrizDeSimilitud[i][j] = 0;
09     Fin Si
10 Fin Para
11 Para j=1 hasta n haga
12     MatrizDeSimilitud[i][j] = MatrizDeSimilitud[i][j] / grado;
13 Fin Para
14 Fin Para
15 Para i=1 hasta n haga
16     Para j=1 hasta n haga
17         MatrizDeSimilitud[i][j] = (dampingFactor/n) + (1- dampingFactor)*
MatrizDeSimilitud[i][j];
18     Fin Para
19 Fin Para
20 L = MetodoDePotencia(MatrizDeSimilitud,n, ε ); // Ver Algoritmo 4 Método de Potencia.
21 retorne L;
```

Algoritmo 3: centralidad de los patrones de navegación ef-isf por sesión. Adaptado de (Erkan & Radev, 2004).

El Método de Potencia describe como se calcula la matriz estacionaria de una Cadena de Markov, este se muestra en el Algoritmo 4.

Entrada: Una matriz M estocástica, irreducible y aperiódica

Entrada: Tamaño de la matriz N , Error de Tolerancia ϵ

Salida: Vector propio p con la centralidad de los patrones de navegación ef-isf por sesión

```
01  $p_0 = \frac{1}{N} \mathbf{1}$ 
02  $t = 0$ 
03 repita
04      $t = t + 1$ 
05      $p_t = M^T p_{t-1}$ 
06      $\delta = \|p_t - p_{t-1}\|$ 
07 hasta que  $\delta < \epsilon$ ;
08 retorne  $p_t$ 
```

Algoritmo 4: Método de Potencia. Adaptado de (Erkan & Radev, 2004).

3.3.2 Patrón de Navegación EF-ISF por sesión de mayor centralidad

Los valores de probabilidad $\in \mathbb{R}$ de la matriz estacionaria, son la centralidad de los patrones de navegación EF-ISF por sesión del conjunto de sesiones del usuario. El número mayor de la matriz estacionaria, p_j , es el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario i , además, el índice j , indica cual es la sesión \vec{s}_j del conjunto de sesiones realizadas por el usuario i , la cual contiene la información más relevante para el usuario en el curso.

CAPITULO 4

En este capítulo se describen los parámetros necesarios para el algoritmo LexRank con Umbral, para la identificación de los patrones de navegación en conjuntos de datos, los cuales contiene la interacción de estudiantes, administradores, profesores y monitores con el curso. Además, la descripción del proceso para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad en los conjuntos de datos y la forma de graficar la centralidad de los patrones de navegación EF-ISF por sesión de usuario.

4 PROCESO PARA LA IDENTIFICACIÓN DE LOS PATRONES DE NAVEGACIÓN EF-ISF POR SESIÓN DE MAYOR CENTRALIDAD EN CONJUNTOS DE DATOS

4.1 Parámetros del algoritmo LexRank con Umbral

En esta investigación se trabajan con cuatro conjuntos de datos, obtenidos del curso Astronomía Cotidiana, el cual se oferto en el año 2017 primer periodo académico de la Universidad del Cauca. El curso se dividió en Grupo A y Grupo B; del Grupo A se obtuvieron dos archivos, Astronomía Cotidiana 2017-I Grupo A y Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5; del Grupo B se obtuvieron dos archivos, Astronomía Cotidiana 2017-I Grupo B y Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5. Los archivos contienen interacción de usuarios con el curso respectivo, entre los usuarios se encuentran: administrador, profesor, monitor y estudiante, ver sección 3.1.

Por lo anterior, en este trabajo a parte de los estudiantes, también se va a trabajar con la información de los datos de, administradores, profesores y monitores; con el fin de identificar los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios. Para ello es necesario que cada usuario i sea representado en la matriz de similitudes en modelo de espacio vectorial, ver sección 3.2.7.2. Cada matriz del usuario i , es parámetro de entrada del algoritmo LexRank con Umbral, en la Tabla 22 se describen los parámetros del algoritmo.

Tabla 22. Parámetros del LexRank con Umbral. Adaptado de (Anaconda et al., 2015).

Parámetro	Valor	Descripción
Factor de amortiguamiento	0,15	Valor para realizar desplazamientos en los nodos de sesión del usuario i .
Umbral	0,1	Valor para clasificar las sesiones del usuario i , si el ángulo entre dos sesiones del usuario i es superior a este umbral.

Error de Tolerancia	0,1	Valor para terminar el método de potencia cuando se está calculando la matriz estacionaria.
Matriz de Similitudes	n	Es una matriz de $n \times n$, donde n el número de sesiones del usuario i .

4.2 Paso a paso para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios

Para comenzar es importante aclarar, la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios, se realiza en cada archivo Excel por separado. A continuación, se muestra un paso a paso del proceso.

Al archivo Excel, se le realiza un pre procesamiento para la limpieza y adecuación, ver secciones 3.1, 3.2.1, 3.2.2, luego, se separan los usuarios en archivos de texto plano con sus respectivos registros, organizándolos por sesiones, se crean los archivos de sesiones por usuario i , ver sección 3.2.3, seguidamente, para cada usuario i , se procesan las sesiones con la ecuación (3-2) ponderación basada en la frecuencia relativa de un evento, y se representan en la *Matriz EF – ISF* $m \times n$ de m eventos por n sesiones del usuario, ver sección 3.2.7.1. La Matriz EF-ISF se transforma en la *MatrizDeSimilitud* $n \times n$ de n sesiones del usuario i , ver sección 3.2.7.2. En la Figura 8 se muestran el proceso descrito.

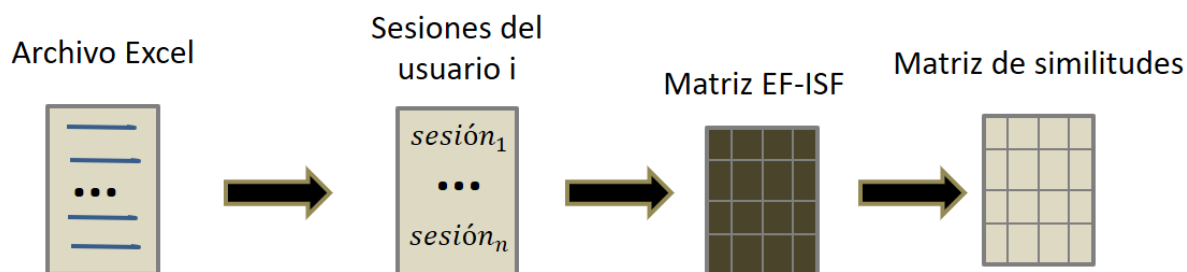


Figura 8. Archivo Excel de usuarios representado en el modelo de espacio vectorial.
Fuente Propia.

La Matriz de Similitudes es recibida por el algoritmo LexRank con Umbral, la transforma a una Matriz Estocástica que sea irreducible y aperiódica, como se describió en la sección 3.3.1, seguidamente se encuentra la Matriz Estacionaria, en la cual se encuentran la centralidad de los patrones de navegación EF-ISF por Sesión. El número mayor $\in \mathbb{R}$ en la Matriz Estacionaria, p_j , es el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario i , por último se

muestra la sesión j del usuario i relacionada con el valor p_j . En la Figura 9 se muestra el proceso.

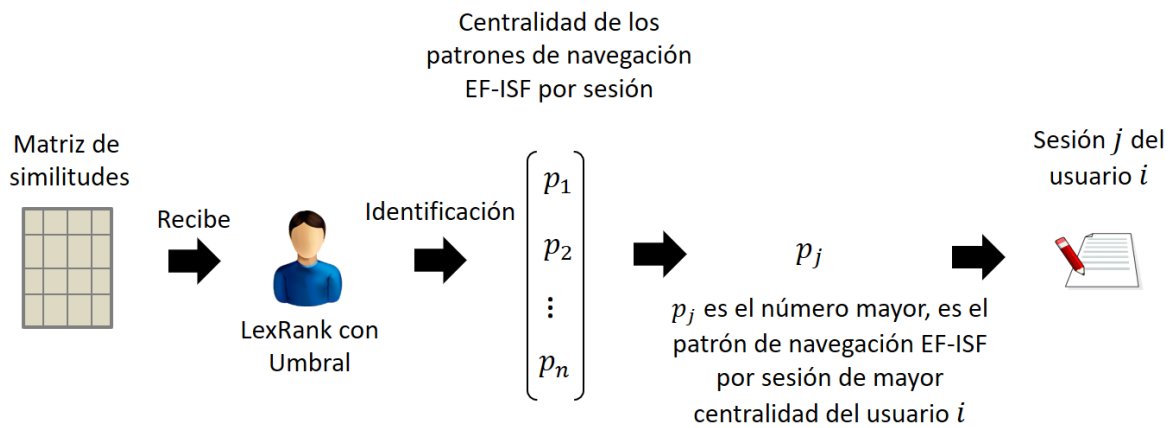


Figura 9. Identificación del patrón de navegación EF-ISF por sesión de mayor centralidad. Fuente propia.

4.3 Grafica de la centralidad del patrón de navegación EF-ISF por sesión del usuario

En esta sección se propone como construir la gráfica del patrón de navegación del usuario, esta se puede graficar a partir de los resultados de la matriz estacionaria, obtenidos a través del procesamiento del archivo Excel con el Algoritmo LexRank con Umbral. En la literatura de MOOC, hasta el momento no hay graficas de patrones de navegación EF-ISF por sesión del usuario i , por tal razón, se la llamara, grafica de la centralidad del patrón de navegación EF-ISF por sesión del usuario i .

La forma de representar la gráfica, es en el plano cartesiano, el *eje x* representa las sesiones del usuario i , ordenadas por fecha y hora de menor a mayor; entre tanto, el *eje y* representa las centralidades de los patrones de navegación EF-ISF por sesión del usuario i . La grafica se ubica en el primer cuadrante del plano cartesiano, porque se está trabajando con números reales positivos, ver Figura 10.

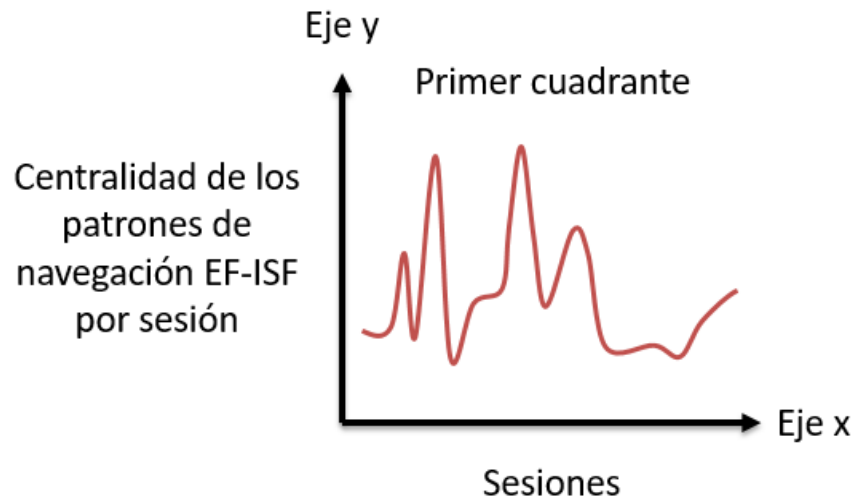


Figura 10. Grafica de la centralidad del patrón de navegación EF-ISF por sesión del usuario *i*. Fuente propia.

En la gráfica se puede observar como el usuario ha realizado la navegación y la interacción en el curso, además, se puede identificar el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario. La grafica es continua y dependiendo de la interacción del usuario con el curso, puede ser: creciente, decreciente, constante, con presencia de máximos y mínimos.

CAPITULO 5

En este capítulo se describen los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios, la sesión a la cual le corresponde la centralidad y la gráfica de la centralidad del patrón de navegación EF-ISF por sesión. Para la presentación de los resultados y análisis, se describen los resultados del administrador, profesor y monitor, además de 3 estudiantes con sesiones bajas, medias y altas por cada archivo Excel, teniendo en cuenta el número de sesiones promedio de los usuarios; esto es debido a la gran cantidad de estudiantes por cada archivo. También, se realiza anonimización en la información, para no identificar al usuario que realizó las actividades en el curso y por último conclusiones de este capítulo.

5 RESULTADOS Y ANÁLISIS

5.1 Astronomía Cotidiana 2017-I Grupo A

En esta sección se realiza un análisis de algunos usuarios del archivo Excel, Astronomía Cotidiana 2017-I Grupo A, el cual contiene interacciones de los usuarios con el grupo A, ver sección 3.2.1. Los resultados obtenidos por el LexRank con Umbral, se mostrarán en tablas con campos que se describen en la Tabla 23.

Tabla 23. Descripción de campos de resultados del algoritmo. Fuente propia.

Nombre	Descripción
Hora	Hora en la que el usuario realizó la actividad.
Fecha	Fecha en la que el usuario realizó la actividad.
Sesión Id	Identificador de la sesión del usuario.
Centralidad	Centralidad del patrón de navegación EF-ISF por sesión j , calculado por el algoritmo LexRank con Umbral.
Registros	Número de registros realizados por el usuario en la sesión j

Los resultados de la sesión j , a la que le corresponde el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario, se mostrarán en una tabla, con algunos campos del archivo Excel que se analizan, los campos son: Hora, Fecha, Actividad, Complemento1, Complemento2 y Complemento3; la descripción detallada de cada campo, se puede consultar en la sección 3.1.

5.1.1 Estudiantes

Estudiante1GA

El estudiante1GA, al interactuar con el curso realizo en total, 5 sesiones y 62 registros, a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 24.

Tabla 24. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante1GA. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
02:34:14 am	24/02/2017	s1	0,1858	3
05:07:07 pm	28/02/2017	s2	0,1858	31
05:50:41 pm	7/03/2017	s3	0,2708	4
06:02:49 pm	7/03/2017	s4	0,1575	14
01:39:26 am	17/03/2017	s5	0,2000	10
Total Sesiones		5	Total Registros	62

En la Tabla 24, en el campo centralidad el número mayor es el 0,2708, este valor representa, el patrón de navegación EF-ISF por sesión de mayor centralidad, s3 es la sesión a la que le corresponde este valor. En la Tabla 25 se muestra en detalle la sesión realizada por el estudiante1GA.

Tabla 25. Sesión s3, del estudiante1GA. Adaptado de (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
5:50:41 pm	7/03/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Examen en Línea	NULL
5:50:59 pm	7/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Simuladores	NULL
5:51:25 pm	7/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Examen en Línea	NULL
6:02:03 pm	7/03/2017	play_video	U1T2: Los movimientos del planeta tierra	Videos	C02-v01 los movimientos del planeta Tierra

Por el valor cuantitativo 0,2708, se realiza una interpretación cualitativa en la sesión 3, con la ponderación basada en la frecuencia relativa de un evento y la centralidad. De la Tabla 25, se obtienen los patrones de navegación del estudiante1GA, ver Tabla 26.

Tabla 26. Patrones de navegación del estudiante1GA. Fuente propia.

Hora	Patrones de navegación
5:50:41 pm	U1T1: Ubicación de nuestro planeta en el universo -> Examen en Línea.
	U1T2: Los movimientos del planeta tierra -> Simuladores.
-	U1T2: Los movimientos del planeta tierra -> Examen en Línea.
6:02:03 pm	U1T2: Los movimientos del planeta tierra -> Videos -> C02-v01 los movimientos del planeta Tierra -> play_video.

El estudiante1GA, tiene cuatro patrones de navegación, por lo general ingresa al curso en horas de la tarde, observa la evaluación; pero no la envía, utiliza diferente material de aprendizaje para el desarrollo del curso, tiene una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje. El material de aprendizaje que más utiliza es video sin interrupciones, exámenes en línea y simuladores, por lo anterior, se infiere que el estudiante está mirando los exámenes en línea para determinar respuestas correctas e incorrectas, este comportamiento se clasifica como una deshonestidad académica. Ahora bien, con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante1GA, esta se ilustra en la Figura 11.

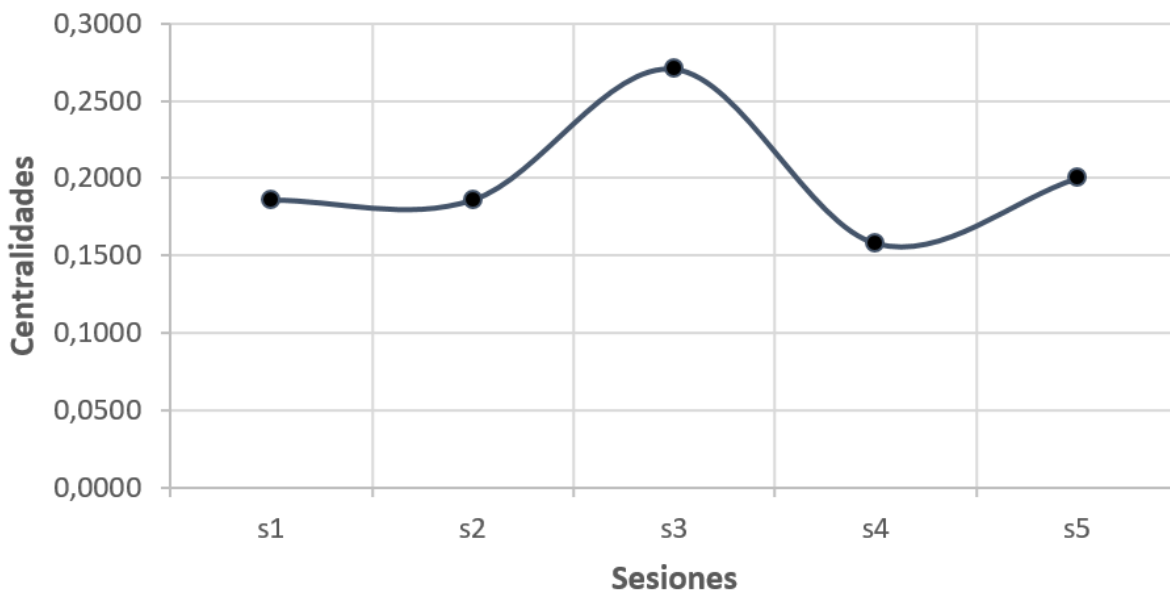


Figura 11. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante1GA. Fuente propia.

En la gráfica se observa cómo fue la interacción del estudiante en el curso, el estudiante1GA, empieza a interactuar lentamente con el curso, después incrementa su interacción llegando a su punto más alto en la sesión 3, luego disminuye su

interacción alcanzando su punto más bajo, posteriormente tiene una leve recuperación en la interacción con el curso. El estudiante tiene una navegación aleatoria. En la gráfica, el patrón de navegación EF-ISF por sesión de mayor centralidad, es el punto (s3 , 0,2708).

Estudiante2GA

El estudiante2GA, al interactuar con el curso realizo en total, 15 sesiones y 63 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 27.

Tabla 27. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante2GA. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
04:32:20 pm	8/02/2017	s1	0,0423	1
02:55:32 am	14/02/2017	s2	0,0483	8
12:27:31 am	3/03/2017	s3	0,0449	2
01:03:15 am	3/03/2017	s4	0,0840	15
01:40:54 am	3/03/2017	s5	0,1002	6
11:42:40 pm	9/03/2017	s6	0,0824	2
11:59:06 pm	9/03/2017	s7	0,0685	6
01:39:11 am	10/03/2017	s8	0,0566	5
11:48:28 pm	16/03/2017	s9	0,0858	1
01:32:33 am	17/03/2017	s10	0,0706	3
01:30:45 am	24/03/2017	s11	0,0653	4
06:09:39 am	31/03/2017	s12	0,0527	3
01:57:44 am	21/04/2017	s13	0,0854	4
02:35:50 am	21/04/2017	s14	0,0566	1
02:06:27 am	28/04/2017	s15	0,0566	2
Total sesiones		15	Total Registros	63

En la Tabla 27, en el campo centralidad el 0,1002 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s5 es la sesión a la que le corresponde este valor, en la Tabla 28 se muestra en detalle la sesión realizada por el estudiante2GA.

Tabla 28. Sesión s5, del estudiante2GA. Adaptado de (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
1:40:54 am	3/03/2017	Contenidos	Introducción	Examen en línea - simulacro	NULL
1:42:33 am	3/03/2017	Contenidos	discussion	NULL	NULL
1:43:43 am	3/03/2017	Contenidos	U1T1: Ubicación de nuestro	Examen en Línea	NULL

			planeta en el universo		
2:05:20 am	3/03/2017	problem_check	U1T1: Ubicación de nuestro planeta en el universo	Examen en Línea	NULL
2:05:28 am	3/03/2017	problem_check	U1T1: Ubicación de nuestro planeta en el universo	Examen en Línea	NULL
2:05:36 am	3/03/2017	problem_check	U1T1: Ubicación de nuestro planeta en el universo	Examen en Línea	NULL

Por el valor cuantitativo 0,1002, se realiza una interpretación cualitativa en la sesión 5, con la ponderación basada en la frecuencia relativa de un evento y la centralidad. De la Tabla 28, se obtienen los patrones de navegación del estudiante2GA, ver Tabla 29.

Tabla 29. Patrones de navegación del estudiante2GA. Fuente propia.

Hora	Patrones de Navegación
1:40:54 am	Introducción -> Examen en línea - simulacro
	discussion
2:05:36 am	U1T1: Ubicación de nuestro planeta en el universo -> Examen en Línea
	U1T1: Ubicación de nuestro planeta en el universo (3) -> Examen en Línea (3) -> problem_check (3)

El estudiante2GA, tiene cuatro patrones de navegación, por lo general ingresa al curso en horas de la madrugada, revisa foros de discusión, simulacros de las evaluaciones y evaluaciones en línea y los envía, el estudiante está buscando las respuestas correctas e incorrectas de los exámenes, sin cumplir las actividades de estudio planteadas por el docente, con lo cual, se puede inferir un comportamiento deshonesto en las actividades académicas. Navega en diferentes secciones, unidades y utiliza diferente material de aprendizaje, por tal razón, el estudiante tiene una navegación aleatoria. Ahora bien, con los resultados de la matriz estacionaria, se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante2GA, esta se ilustra en la Figura 12.

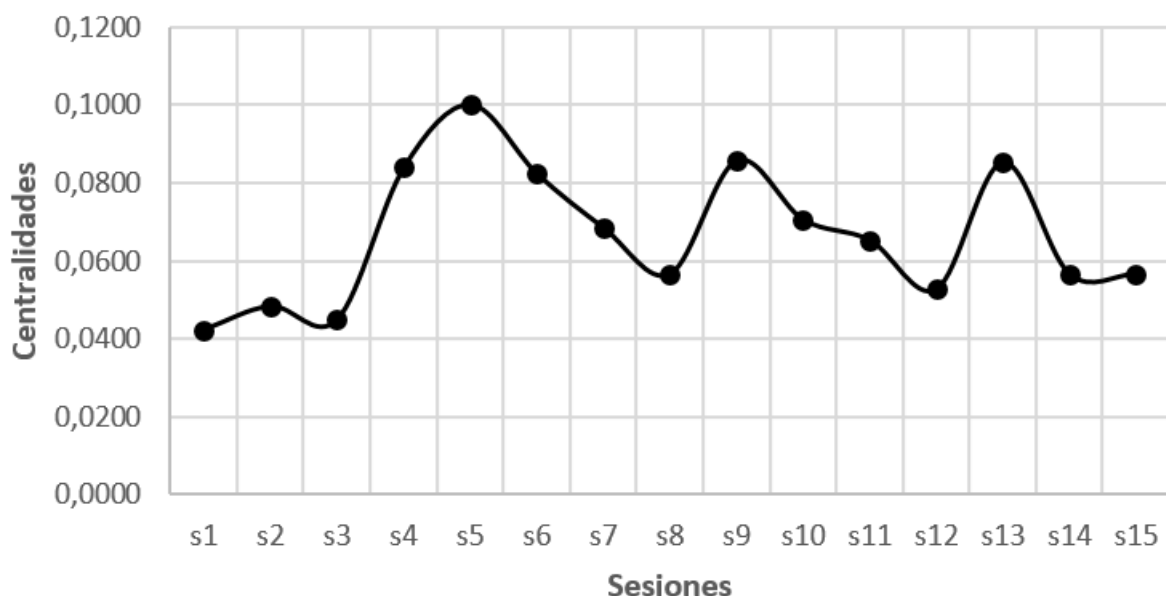


Figura 12. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante2GA. Fuente propia.

En la gráfica se observa cómo fue la interacción del estudiante en el curso, donde tuvo una menor, media o alta interacción. El estudiante2GA, inicia a interactuar lentamente con el curso, después incrementa su interacción llegando a su punto más alto en la sesión 5, empieza a disminuir y aumentar su interacción repitiendo este ciclo. La interacción del estudiante es variada a lo largo del curso, navegando de manera aleatoria. El patrón de navegación EF-ISF por sesión de mayor centralidad se observa en la gráfica, siendo este el punto máximo, es decir, s5 con su valor 0,1002.

Estudiante3GA

El estudiante3GA, al interactuar con el curso realizó en total, 43 sesiones y 673 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 30.

Tabla 30. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante3GA. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
11:49:17 pm	13/02/2017	s1	0,0181	27
07:02:53 pm	14/02/2017	s2	0,0169	5
02:05:52 am	15/02/2017	s3	0,0250	21
02:45:28 am	17/02/2017	s4	0,0169	3
11:34:22 pm	20/02/2017	s5	0,0282	13

05:40:12 pm	21/02/2017	s6	0,0254	15
01:02:40 am	24/02/2017	s7	0,0250	32
03:37:34 pm	24/02/2017	s8	0,0175	4
02:22:32 am	25/02/2017	s9	0,0212	6
12:07:05 am	28/02/2017	s10	0,0348	13
02:43:43 am	28/02/2017	s11	0,0326	14
05:29:58 pm	28/02/2017	s12	0,0261	8
06:49:09 pm	28/02/2017	s13	0,0253	8
10:20:45 pm	28/02/2017	s14	0,0198	36
12:46:17 am	1/03/2017	s15	0,0165	30
12:42:32 am	7/03/2017	s16	0,0410	15
05:22:38 pm	7/03/2017	s17	0,0370	15
08:25:25 pm	7/03/2017	s18	0,0202	14
10:35:04 pm	7/03/2017	s19	0,0217	40
01:26:28 am	8/03/2017	s20	0,0177	16
12:08:18 am	14/03/2017	s21	0,0235	8
05:31:59 pm	14/03/2017	s22	0,0190	2
07:01:51 pm	14/03/2017	s23	0,0190	30
10:56:09 pm	14/03/2017	s24	0,0239	4
05:21:07 pm	21/03/2017	s25	0,0301	8
08:36:06 pm	21/03/2017	s26	0,0187	18
10:22:12 pm	21/03/2017	s27	0,0187	21
02:07:25 am	22/03/2017	s28	0,0172	1
12:14:05 am	4/04/2017	s29	0,0180	5
05:37:05 pm	4/04/2017	s30	0,0145	2
11:21:03 pm	5/04/2017	s31	0,0180	41
02:12:04 pm	11/04/2017	s32	0,0287	7
08:08:28 pm	11/04/2017	s33	0,0282	15
04:18:29 pm	12/04/2017	s34	0,0183	50
12:44:03 am	18/04/2017	s35	0,0183	10
05:32:39 pm	18/04/2017	s36	0,0249	6
07:27:52 pm	18/04/2017	s37	0,0406	9
05:25:21 pm	25/04/2017	s38	0,0206	46
05:40:07 pm	2/05/2017	s39	0,0206	42
02:36:54 am	8/05/2017	s40	0,0338	2
03:08:27 am	9/05/2017	s41	0,0169	1
05:42:05 pm	9/05/2017	s42	0,0206	3
06:47:32 pm	9/05/2017	s43	0,0206	7
Total Sesiones		43	Total Registros	673

En la Tabla 30, en el campo centralidad el 0,0410 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s16 es la sesión

a la que le corresponde este valor, en la Tabla 31 se muestra en detalle la sesión realizada por el estudiante3GA.

Tabla 31. Sesión s16, del estudiante3GA. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
12:42:32 am	7/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Videos	NULL
12:42:54 am	7/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Taller	NULL
12:43:03 am	7/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Simuladores	NULL
12:43:09 am	7/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Página WEB	NULL
12:43:26 am	7/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Foros semana 03	NULL
12:43:33 am	7/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Taller	NULL
12:43:39 am	7/03/2017	Contenidos	U1T4: La estructura del sistema solar	Videos	NULL
12:44:03 am	7/03/2017	Contenidos	U1T4: La estructura del sistema solar	Diapositivas	NULL
12:44:12 am	7/03/2017	Contenidos	U1T4: La estructura del sistema solar	Videos complementarios	NULL
12:44:22 am	7/03/2017	Contenidos	U1T4: La estructura del sistema solar	Simuladores	NULL
12:44:37 am	7/03/2017	Contenidos	U1T4: La estructura del sistema solar	Recursos Web	NULL
12:44:43 am	7/03/2017	Contenidos	U1T4: La estructura del sistema solar	Foros semana 04	NULL
12:45:37 am	7/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Diapositivas	NULL
1:11:34 am	7/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Videos complementarios	NULL
1:11:55 am	7/03/2017	pause_video	U1T2: Los movimientos del planeta tierra	Videos complementarios	Rotación de la Tierra

Por el valor cuantitativo 0,0410, se realiza una interpretación cualitativa en la sesión 16, con la ponderación basada en la frecuencia relativa de un evento y la centralidad. De la Tabla 31, se obtienen los patrones de navegación del estudiante3GA, ver Tabla 32.

Tabla 32. Patrones de navegación del estudiante3GA. Fuente propia.

Hora	Patrones de navegación
12:42:32 am	U1T3: Principios de astronomía de observación -> Videos.
	U1T3: Principios de astronomía de observación (2) -> Taller (2).
	U1T3: Principios de astronomía de observación -> Simuladores.
	U1T3: Principios de astronomía de observación -> Página WEB.
	U1T3: Principios de astronomía de observación -> Foros semana 03.
	U1T4: La estructura del sistema solar -> Videos.
	U1T4: La estructura del sistema solar -> Diapositivas.
1:11:55 am	U1T4: La estructura del sistema solar -> Videos complementarios.
	U1T4: La estructura del sistema solar -> Simuladores.
1:11:55 am	U1T4: La estructura del sistema solar -> Recursos Web.
	U1T4: La estructura del sistema solar -> Foros semana 04.

U1T2: Los movimientos del planeta tierra -> Diapositivas.
U1T2: Los movimientos del planeta tierra -> Videos complementarios.
U1T2: Los movimientos del planeta tierra -> Videos complementarios -> Rotación de la Tierra -> pause_video.

El estudiante3GA, tiene 14 patrones de navegación, por lo general ingresa al curso en horas de la madrugada, tiene una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje para el desarrollo del curso, como: videos, videos complementarios, simulacros, foros, talleres, diapositivas. Por lo general, cuando ve videos realiza pausas en los videos que está observando. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante3GA, esta se ilustra en la Figura 13.

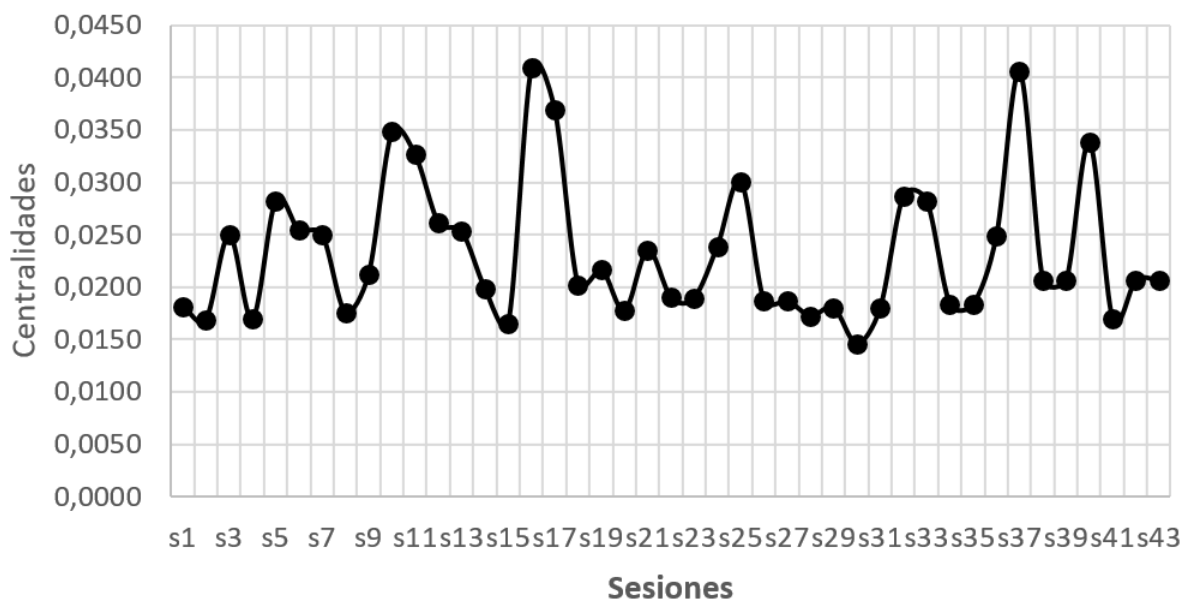


Figura 13. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante3GA. Fuente propia.

En la gráfica se observa cómo fue la interacción del estudiante en el curso. La interacción con el curso es decreciente hasta la sesión 2, muy similar con otras sesiones, incrementa su participación con el curso teniendo interacciones altas y bajas, llegando a su punto más alto en la sesión 16, empieza a disminuir y aumentar su interacción repitiendo este proceso hasta la sesión 37, un punto máximo, empieza a disminuir hasta tener casi una navegación constante en las dos últimas sesiones. De lo anterior se puede concluir, el estudiante3GA tiene una navegación aleatoria a lo largo del curso. En la sesión 30, su interacción con el curso fue menor, en comparación con las otras sesiones. Las sesiones s10, s16, s37 y s40, fue donde

el estudiante tuvo una mayor interacción con el curso. Por otra parte, en la gráfica se identifica, el patrón de navegación EF-ISF por sesión de mayor centralidad del estudiante3GA, al observar en la gráfica el punto máximo, es decir, s16 con su valor 0,0410.

5.1.2 AdministradorGA

El administradorGA, al interactuar con el curso realizo en total, 8 sesiones y 37 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 33.

Tabla 33. Centralidad de los patrones de navegación EF-ISF por sesión del administradorGA. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
05:25:59 am	13/02/2017	s1	0,1073	2
02:48:38 am	3/03/2017	s2	0,1516	1
03:51:38 am	17/04/2017	s3	0,1339	27
06:00:51 pm	18/04/2017	s4	0,1161	1
06:04:36 pm	18/04/2017	s5	0,1161	2
05:51:31 pm	12/05/2017	s6	0,1250	1
10:44:09 am	27/05/2017	s7	0,1250	1
09:46:21 pm	18/07/2017	s8	0,1250	2
Total sesiones		8	Total Registros	37

En la Tabla 33, en el campo centralidad el 0,1516 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s2 es la sesión a la que le corresponde este valor, en la Tabla 34 se muestra en detalle la sesión realizada por el administradorGA.

Tabla 34. Sesión s2, del administradorGA. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
2:48:38 am	3/03/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Examen en Línea	NULL

Por el valor cuantitativo 0,1516, se realiza una interpretación cualitativa en la sesión 2, con la ponderación basada en la frecuencia relativa de un evento y la centralidad. De la Tabla 34, se infiere que el administradorGA ingresa por lo general al curso en horas de la madrugada, se obtiene un solo el patrón de navegación: ingresa a una sección y navega hacia los exámenes en línea, el administrador por lo general ingresa a la plataforma a revisar los exámenes en línea, ya sea para prepararlos o estar pendiente de ellos cuando los estudiantes los estén presentando. Con la

información de la tabla no se puede afirmar si el administradorGA tiene una navegación lineal o aleatoria.

Ahora bien, con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del administradorGA, esta se ilustra en la Figura 14.

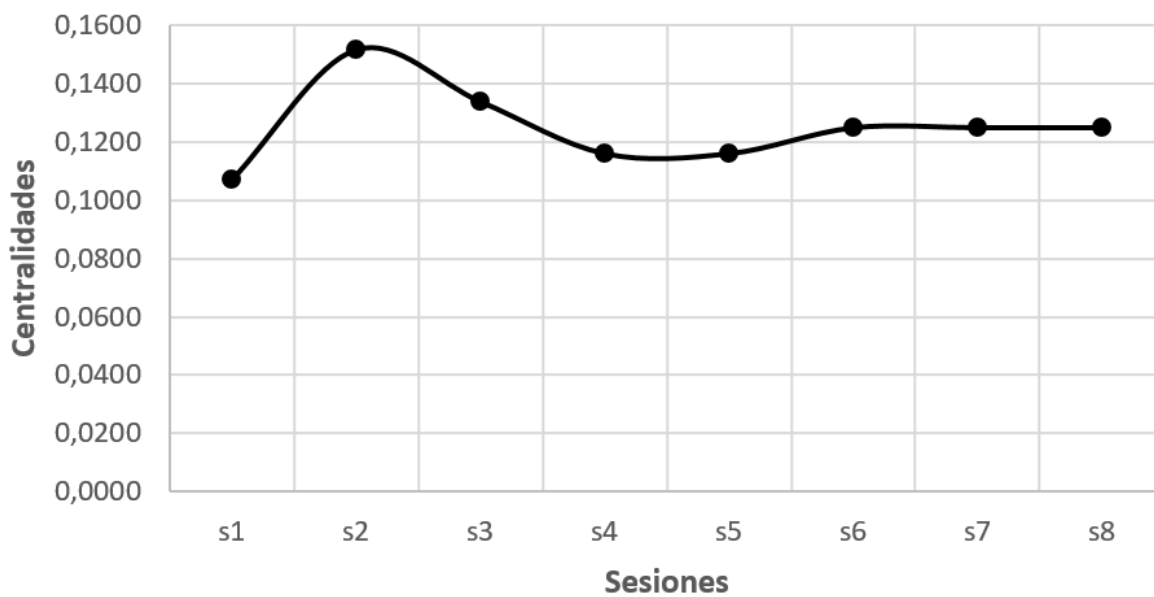


Figura 14. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del administradorGA. Fuente propia.

En la gráfica de la Figura 14 se observa cómo fue la interacción del administradorGA con el curso, al inicio del curso tiene una interacción activa, luego empezó a disminuir su uso; pero en una menor intensidad llegando incluso a tener un comportamiento constante. El punto más alto lo alcanza en la sesión 2, siendo este el patrón de navegación EF-ISF por sesión de mayor centralidad, es decir, s2 con su valor 0,1516; además, la gráfica de patrón de navegación permite afirmar que el administradorGA tiene una navegación aleatoria.

5.1.3 ProfesorGA

El profesorGA, al interactuar con el curso realizó en total, 64 sesiones y 197 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 35.

Tabla 35. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGA.
Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros	Hora	Fecha	Sesión Id	Centralidad	Registros
11:46:20 am	7/02/2017	s1	0,0129	1	04:29:24 pm	13/03/2017	s33	0,0128	1
11:52:06 am	7/02/2017	s2	0,0133	4	02:13:19 pm	14/03/2017	s34	0,0169	3
02:56:45 pm	7/02/2017	s3	0,0141	1	05:32:56 pm	14/03/2017	s35	0,0138	1
02:42:41 pm	8/02/2017	s4	0,0191	5	02:31:59 pm	16/03/2017	s36	0,0119	1
07:56:19 pm	8/02/2017	s5	0,0204	2	12:02:26 pm	20/03/2017	s37	0,0211	5
02:01:33 am	9/02/2017	s6	0,0131	1	04:04:00 pm	23/03/2017	s38	0,0131	1
02:24:47 pm	9/02/2017	s7	0,0131	1	11:29:08 pm	23/03/2017	s39	0,0131	1
07:27:13 pm	9/02/2017	s8	0,0129	1	01:39:38 am	24/03/2017	s40	0,0122	1
01:58:52 pm	10/02/2017	s9	0,0131	1	09:08:59 pm	31/03/2017	s41	0,0112	5
07:15:10 pm	11/02/2017	s10	0,0131	1	05:41:34 pm	4/04/2017	s42	0,0136	1
12:32:14 am	13/02/2017	s11	0,0141	1	02:34:52 pm	5/04/2017	s43	0,0200	26
07:16:48 pm	13/02/2017	s12	0,0226	4	12:05:29 pm	6/04/2017	s44	0,0090	8
04:36:08 pm	14/02/2017	s13	0,0141	1	04:32:09 pm	6/04/2017	s45	0,0109	2
05:08:20 pm	14/02/2017	s14	0,0314	6	01:15:00 am	7/04/2017	s46	0,0112	6
12:57:32 am	17/02/2017	s15	0,0133	2	01:58:55 am	13/04/2017	s47	0,0136	1
02:59:46 pm	17/02/2017	s16	0,0131	1	02:37:03 pm	17/04/2017	s48	0,0129	2
07:48:12 pm	17/02/2017	s17	0,0224	3	10:10:13 pm	20/04/2017	s49	0,0074	3
08:32:08 pm	19/02/2017	s18	0,0231	4	11:05:52 am	22/04/2017	s50	0,0162	1
02:16:06 pm	20/02/2017	s19	0,0146	1	09:37:17 am	24/04/2017	s51	0,0129	1
02:27:33 pm	20/02/2017	s20	0,0182	2	12:52:47 pm	25/04/2017	s52	0,0102	1
01:45:42 pm	22/02/2017	s21	0,0197	2	05:34:43 pm	25/04/2017	s53	0,0162	1
01:27:32 am	24/02/2017	s22	0,0133	1	09:32:13 pm	28/04/2017	s54	0,0225	5
03:03:47 pm	25/02/2017	s23	0,0131	3	05:35:43 pm	2/05/2017	s55	0,0190	4
02:35:29 pm	27/02/2017	s24	0,0122	1	06:12:55 pm	3/05/2017	s56	0,0122	1
06:25:10 pm	28/02/2017	s25	0,0260	2	08:16:19 am	4/05/2017	s57	0,0194	2
02:54:21 pm	2/03/2017	s26	0,0300	7	08:19:00 am	4/05/2017	s58	0,0124	12
01:33:48 am	3/03/2017	s27	0,0228	5	11:02:12 am	5/05/2017	s59	0,0124	11
12:05:39 pm	7/03/2017	s28	0,0118	5	02:31:34 pm	5/05/2017	s60	0,0099	1
05:42:41 pm	7/03/2017	s29	0,0238	2	03:11:34 pm	5/05/2017	s61	0,0136	1
08:46:02 pm	8/03/2017	s30	0,0133	2	01:08:02 pm	8/05/2017	s62	0,0129	1
08:43:15 pm	9/03/2017	s31	0,0135	6	08:57:05 am	9/05/2017	s63	0,0220	3
01:26:11 am	10/03/2017	s32	0,0238	2	05:22:05 pm	9/05/2017	s64	0,0080	3
Total sesiones							64	Total Registros	197

En la Tabla 35, en el campo centralidad el 0,0314 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s14 es la sesión

a la que le corresponde este valor, en la Tabla 36 se muestra en detalle la sesión realizada por el profesorGA.

Tabla 36. Sesión s14, del profesorGA. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
5:08:20 pm	14/02/2017	edx.forum.response.created	Sintitulo	NULL	NULL
7:53:59 pm	14/02/2017	Contenidos	Introducción	Comunicaciones SIMCA	NULL
8:05:17 pm	14/02/2017	Contenidos	Introducción	Tutoriales	NULL
8:09:33 pm	14/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Videos	NULL
10:50:11 pm	14/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
10:50:35 pm	14/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Recursos Web	NULL

En la Tabla 36 se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, de la cual se obtienen los patrones de navegación del profesorGA, ver Tabla 37.

Tabla 37. Patrones de navegación del ProfesorGA. Fuente propia.

Hora	Patrones de navegación
5:08:20 pm - 10:50:35 pm	edx.forum.response.created -> Sintitulo.
	Introducción -> Comunicaciones SIMCA.
	Introducción -> Tutoriales.
	U1T1: Ubicación de nuestro planeta en el universo -> Videos.
	U1T1: Ubicación de nuestro planeta en el universo -> Simuladores.
	U1T1: Ubicación de nuestro planeta en el universo -> Recursos Web.

El ProfesorGA, tiene seis patrones de navegación, por lo general ingresa al curso en horas de la tarde empezando la noche, tiene una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje, para publicación o consulta en la plataforma; utiliza diferente material de aprendizaje como: videos, tutoriales, foros, simuladores, comunicaciones SIMCA y recursos web. Contesta a publicaciones o consultas realizadas en los foros del curso, por los estudiantes, monitor y administrador.

Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGA, esta se ilustra en la Figura 15.

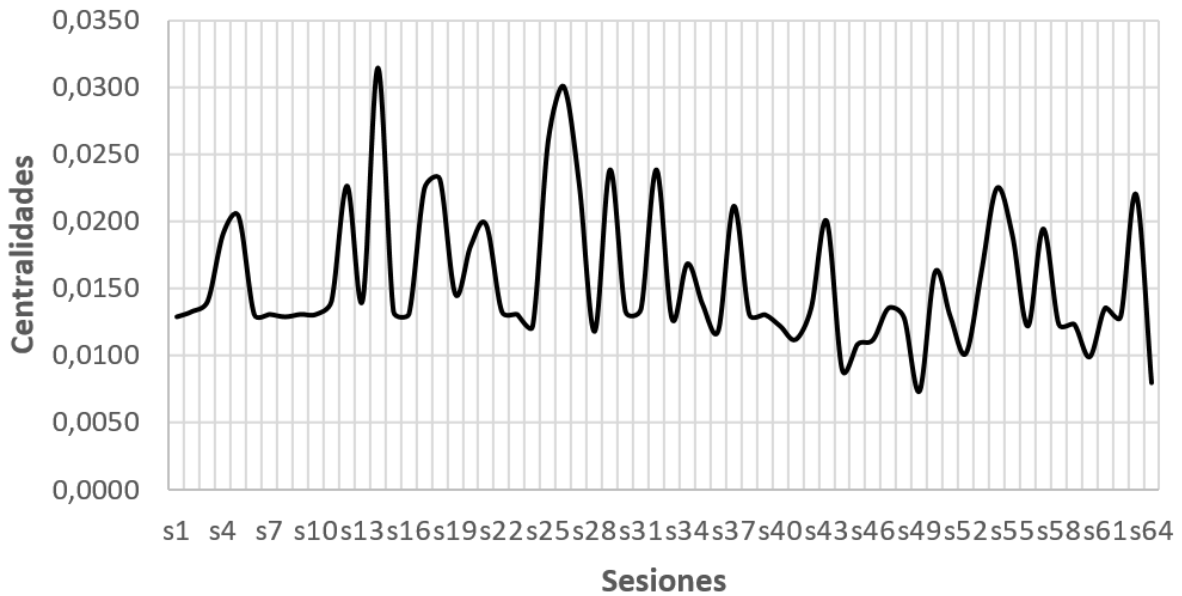


Figura 15. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGA. Fuente propia.

En la gráfica se observa cómo fue la interacción del profesorGA en el curso, la interacción fue activa; sin embargo, entre las sesiones s43 y s49 dejó de utilizar la plataforma de manera considerable, en la última sesión s64 tuvo una fuerte caída con la interacción en el curso. Su punto máximo, es la sesión s14 el cual coincide con la centralidad 0,0314 del patrón de navegación EF-ISF por sesión de mayor centralidad, además el profesorGA presenta una navegación aleatoria en el curso, lo que coincide con el análisis de la sesión s14.

5.1.4 Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5

En este apartado se realizan análisis de algunos usuarios del archivo Excel, Astronomía Cotidiana 2017-I Grupo A U1T1 a U1T5, que contiene interacciones de los usuarios de la unidad uno temática uno a la unidad uno temática cinco, para más detalle ver la sección 3.2.1.2. El análisis al administradorGA no se realiza, porque en estas unidades temáticas solo ha realizado dos sesiones, ver sección 3.2.1.2.

5.1.4.1 EstudianteGAT1T5

El estudianteGAT1T5 con número de sesiones alto, al interactuar con las unidades temáticas del curso realizó en total, 42 sesiones y 312 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 38.

Tabla 38. Centralidad de los patrones de navegación EF-ISF por sesión del estudianteGAT1T5. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros	Hora	Fecha	Sesión Id	Centralidad	Registros
05:00:27 pm	13/02/2017	s1	0,0301	9	06:21:23 pm	6/03/2017	s22	0,0236	5
05:02:45 pm	14/02/2017	s2	0,0259	8	03:05:46 pm	7/03/2017	s23	0,0270	6
06:10:04 pm	14/02/2017	s3	0,0248	16	03:34:17 pm	7/03/2017	s24	0,0243	5
01:44:22 am	15/02/2017	s4	0,0241	2	05:01:46 pm	7/03/2017	s25	0,0203	7
04:57:53 pm	20/02/2017	s5	0,0196	1	05:31:14 pm	7/03/2017	s26	0,0228	4
06:03:00 pm	20/02/2017	s6	0,0196	1	12:05:12 am	13/03/2017	s27	0,0388	4
02:29:43 pm	22/02/2017	s7	0,0151	1	05:42:21 pm	15/03/2017	s28	0,0160	1
03:25:24 pm	22/02/2017	s8	0,0196	1	11:26:12 pm	15/03/2017	s29	0,0255	10
06:55:43 pm	23/02/2017	s9	0,0290	1	01:06:46 am	16/03/2017	s30	0,0232	19
05:41:57 pm	24/02/2017	s10	0,0350	18	08:09:20 pm	16/03/2017	s31	0,0232	24
02:00:24 am	25/02/2017	s11	0,0306	24	12:28:28 am	17/03/2017	s32	0,0232	7
06:35:42 pm	26/02/2017	s12	0,0285	4	01:31:08 am	17/03/2017	s33	0,0167	2
08:11:54 pm	26/02/2017	s13	0,0263	3	03:08:01 pm	20/03/2017	s34	0,0232	6
12:54:18 am	2/03/2017	s14	0,0295	19	09:35:06 pm	22/03/2017	s35	0,0303	20
08:44:23 pm	2/03/2017	s15	0,0234	11	02:01:30 am	23/03/2017	s36	0,0222	29
01:12:39 am	3/03/2017	s16	0,0247	5	10:33:53 pm	23/03/2017	s37	0,0222	12
01:32:00 am	3/03/2017	s17	0,0267	9	12:43:39 am	24/03/2017	s38	0,0337	5
01:51:34 am	3/03/2017	s18	0,0218	1	01:31:31 am	24/03/2017	s39	0,0169	2
02:15:44 am	3/03/2017	s19	0,0218	1	05:26:20 pm	24/03/2017	s40	0,0160	1
08:35:00 pm	3/03/2017	s20	0,0219	4	02:24:37 pm	31/03/2017	s41	0,0192	2
01:46:10 am	5/03/2017	s21	0,0193	1	09:17:51 pm	2/04/2017	s42	0,0143	1
Total Sesiones							42	Total Registros	312

En la Tabla 38, en el campo centralidad el 0,0388 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s27 es la sesión a la que le corresponde este valor, en la Tabla 39 se muestra en detalle la sesión realizada por el estudianteGAT1T5.

Tabla 39. Sesión s27, del estudianteGAT1T5. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
12:05:12 am	13/03/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Foros semana 01	NULL
12:05:55 am	13/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Foros semana 02	NULL
12:06:57 am	13/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Foros semana 03	NULL
12:07:20 am	13/03/2017	Contenidos	U1T4: La estructura del sistema solar	Foros semana 04	NULL

En la Tabla 39 se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, de la cual se obtienen los patrones de navegación del estudianteGAT1T5, ver Tabla 40.

Tabla 40. Patrones de navegación del estudianteGAT1T5. Fuente propia.

Hora	Patrones de navegación
12:05:12 am -	U1T1: Ubicación de nuestro planeta en el universo -> Foros semana 01.
	U1T2: Los movimientos del planeta tierra -> Foros semana 02.
	U1T3: Principios de astronomía de observación -> Foros semana 03.
12:07:20 am	U1T4: La estructura del sistema solar -> Foros semana 04.

El estudianteGAT1T5, tiene cuatro patrones de navegación, por lo general ingresa al curso en horas de la madrugada, tiene una navegación aleatoria, porque navega en diferentes secciones y unidades; interactúa con los foros de diferentes unidades temáticas. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudianteGAT1T5, esta se ilustra en la Figura 16.

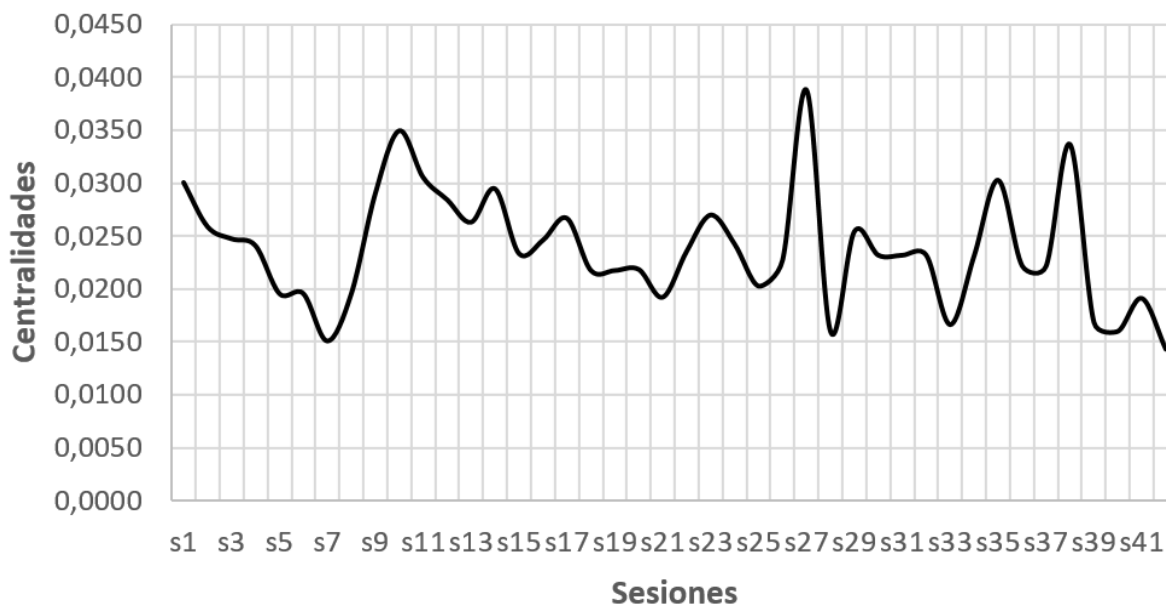


Figura 16. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudianteGAT1T5. Fuente propia.

La gráfica permite observar cómo fue la interacción del estudianteGAT1T5 en las unidades temáticas del curso, inicia con una interacción activa; pero esta va disminuyendo considerablemente hasta la sesión 7, s7, la interacción del

estudianteGAT1T5 va alternado en aumentos y disminución, una navegación aleatoria, luego llega al punto máximo, s27, el patrón de navegación EF-ISF por sesión de mayor centralidad del estudianteGAT1T5, debido a sus interacciones con el foro en las unidades temáticas; pero vuelve nuevamente a disminuir su interacción con el curso de una manera notable.

5.1.4.2 ProfesorGAT1T5

El profesorGAT1T5, al interactuar con las unidades temáticas del curso realizo en total, 28 sesiones y 87 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 41.

Tabla 41. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGAT1T5. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
02:56:45 pm	7/02/2017	s1	0,0350	1
02:42:41 pm	8/02/2017	s2	0,0306	1
12:32:14 am	13/02/2017	s3	0,0350	1
07:18:15 pm	13/02/2017	s4	0,0202	1
04:36:08 pm	14/02/2017	s5	0,0350	1
08:09:33 pm	14/02/2017	s6	0,0381	3
07:48:12 pm	17/02/2017	s7	0,0508	2
08:32:08 pm	19/02/2017	s8	0,0508	2
02:16:06 pm	20/02/2017	s9	0,0403	1
02:27:33 pm	20/02/2017	s10	0,0309	1
06:27:34 pm	28/02/2017	s11	0,0317	1
02:54:21 pm	2/03/2017	s12	0,0586	5
01:33:48 am	3/03/2017	s13	0,0370	4
12:05:39 pm	7/03/2017	s14	0,0329	5
05:42:41 pm	7/03/2017	s15	0,0290	1
08:43:15 pm	9/03/2017	s16	0,0434	5
01:26:11 am	10/03/2017	s17	0,0290	1
04:29:24 pm	13/03/2017	s18	0,0311	1
02:13:19 pm	14/03/2017	s19	0,0330	3
05:32:56 pm	14/03/2017	s20	0,0348	1
02:31:59 pm	16/03/2017	s21	0,0245	1
12:02:26 pm	20/03/2017	s22	0,0469	4
04:04:00 pm	23/03/2017	s23	0,0358	1
11:29:08 pm	23/03/2017	s24	0,0358	1
09:08:59 pm	31/03/2017	s25	0,0320	5
02:34:52 pm	5/04/2017	s26	0,0497	25
12:05:29 pm	6/04/2017	s27	0,0206	8

02:37:03 pm	17/04/2017	s28	0,0276	1
Total Sesiones		28	Total Registros	87

En la Tabla 41, en el campo centralidad el 0,0586 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s12 es la sesión a la que le corresponde este valor, en la Tabla 42 se muestra en detalle la sesión realizada por el profesorGAT1T5.

Tabla 42. Sesión, s12, del profesorGAT1T5. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
2:56:37 pm	2/03/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
2:54:21 pm	2/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Recursos Web	NULL
2:57:14 pm	2/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Simuladores	NULL
2:57:20 pm	2/03/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Recursos Web	NULL
2:57:35 pm	2/03/2017	Contenidos	U1T3: Principios de astronomía de observación	Simuladores	NULL

En la Tabla 42 se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, de la cual se obtienen los patrones de navegación del profesorGAT1T5, ver Tabla 43.

Tabla 43. Patrones de navegación del profesorGAT1T5. Fuente propia.

Hora	Patrones de navegación
2:56:37 pm	U1T1: Ubicación de nuestro planeta en el universo -> Simuladores.
-	U1T2: Los movimientos del planeta tierra (2) -> Recursos Web (2).
2:57:35 pm	U1T2: Los movimientos del planeta tierra -> Simuladores.
	U1T3: Principios de astronomía de observación -> Simuladores.

El profesorGAT1T5, tiene cuatro patrones de navegación, dos menos que en el curso completo, los materiales, simuladores y recursos web, es con los que desarrolla los patrones de navegación de las unidades temáticas, varía en las secciones; por lo general ingresa al curso en horas de la tarde, tiene una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje; observa, utiliza y publica el contenido disponible en las unidades temáticas, generalmente, usa materiales de aprendizaje como:

simuladores y recurso web. Con los resultados de la matriz estacionaria se realiza, la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGAT1T5, esta se ilustra en la Figura 17.

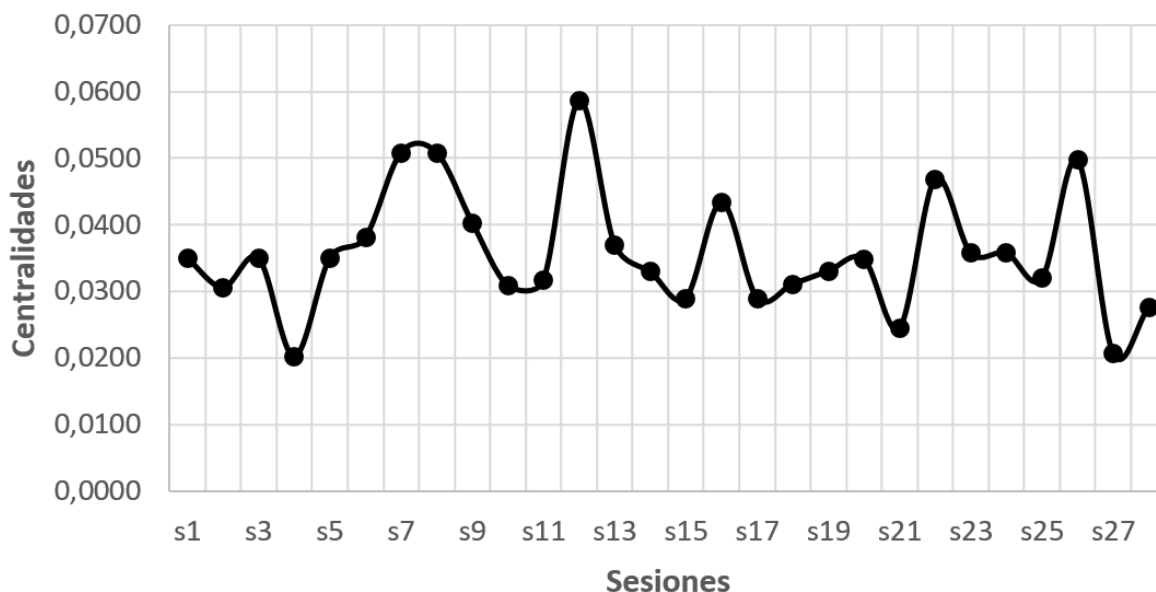


Figura 17. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGAT1T5. Fuente propia.

La gráfica permite observar cómo fue la interacción del profesorGAT1T5 en las unidades temáticas del curso, realizó una navegación aleatoria, el docente tuvo una interacción de manera regular con incrementos y decrementos. Inicia la interacción del curso con decrementos agudizándose en la sesión 4; pero vuelve nuevamente a tener una participación activa en el curso. En la sesión 12 con centralidad 0,0586 se encuentra, el patrón de navegación EF-ISF por sesión de mayor centralidad.

5.2 Astronomía Cotidiana 2017-I Grupo B

En esta sección se realiza un análisis de algunos usuarios del archivo Excel, Astronomía Cotidiana 2017-I Grupo B, el cual contiene interacciones de los usuarios con el grupo B, ver sección 3.2.2.1. Los resultados obtenidos por el LexRank con Umbral, se mostrarán en tablas con campos que se describen en la Tabla 44.

Tabla 44. Descripción de campos de resultados del algoritmo. Fuente propia.

Nombre	Descripción
Hora	Hora en la que el usuario realizo la actividad.
Fecha	Fecha en la que el usuario realizo la actividad.

Sesión Id	Identificador de la sesión del usuario.
Centralidad	Centralidad del patrón de navegación EF-ISF por sesión j , calculado por el algoritmo LexRank con Umbral.
Registros	Número de registros realizados por el usuario en la sesión j

Los resultados de la sesión j , a la que corresponde el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario, se mostrara en una tabla, con algunos campos del archivo Excel que se analizan, los campos son: hora, fecha, actividad, complemento1, complemento2 y complemento3; la descripción detallada de cada campo se puede consultar en la sección 3.1.

5.2.1 Estudiantes

Estudiante1GB

El estudiante1GB, al interactuar con el curso realizo en total, 6 sesiones y 205 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 45.

Tabla 45. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante1GB. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
11:50:00 pm	9/03/2017	s1	0,1431	2
01:41:24 am	14/03/2017	s2	0,1667	2
01:37:56 pm	4/04/2017	s3	0,1549	175
05:25:27 am	7/04/2017	s4	0,1549	17
06:05:51 am	7/04/2017	s5	0,2021	1
07:39:46 pm	9/04/2017	s6	0,1785	8
Total Sesiones		6	Total Registros	205

En la Tabla 45, en el campo centralidad el 0,2021 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s5 es la sesión a la que le corresponde este valor, en la Tabla 46 se muestra en detalle la sesión realizada por el estudiante1GB.

Tabla 46. Sesión s5, del estudiante1GB. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
6:05:51 a. m.	7/04/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Videos	NULL

Por el valor cuantitativo 0,2021, se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad a la sesión 5. El estudiante1GB por lo general ingresa al curso en horas de la mañana, ingresa a una sección y luego a videos del curso, desarrolla un solo patrón de navegación; en cuanto a la navegación, no se puede inferir a partir de la sesión por falta de información. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante1GB, esta se ilustra en la Figura 18.

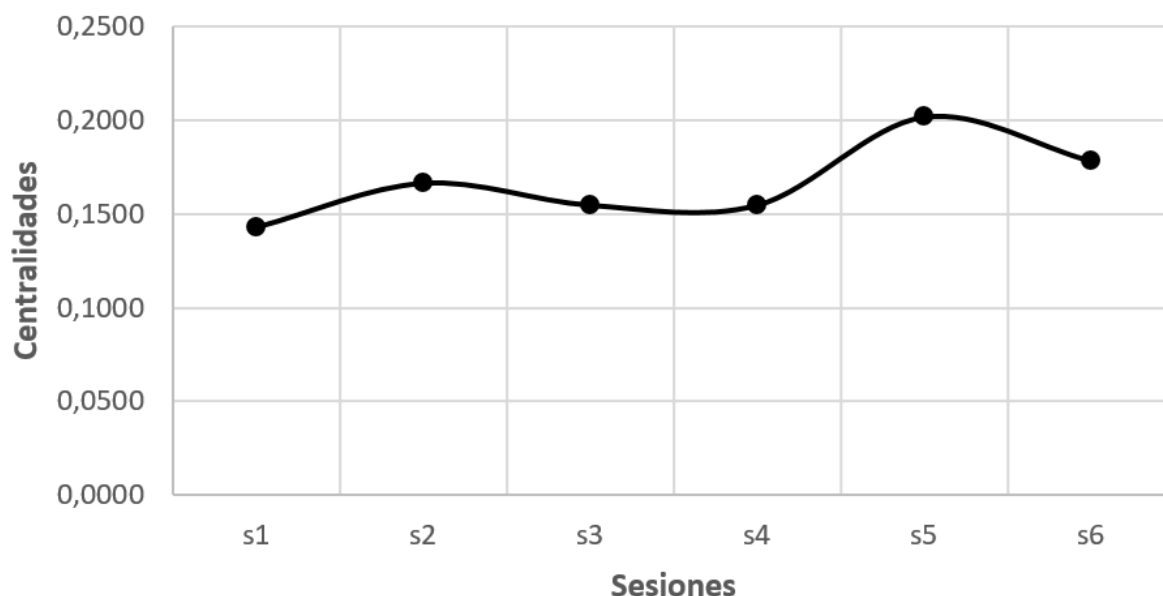


Figura 18. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante1GB. Fuente propia.

La gráfica permite observar cómo fue la interacción del estudiante1GB en el curso, la interacción con el curso va creciendo y decreciendo desde la sesión inicial hasta sesión 5, pero a partir de ahí empieza a descender, con lo cual se puede inferir que el estudiante presenta una navegación aleatoria. Por otra parte, en la gráfica se identifica el patrón de navegación EF-ISF por sesión de mayor centralidad, al observar en la gráfica el punto máximo, es decir, s5 con su valor 0,2021.

Estudiante2GB

El estudiante2GB, al interactuar con el curso realizó en total, 16 sesiones y 455 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 47.

Tabla 47. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante2GB. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
09:36:31 pm	21/02/2017	s1	0,0480	2
12:54:17 am	24/02/2017	s2	0,0900	16
11:48:48 pm	27/02/2017	s3	0,0519	3
12:11:35 am	3/03/2017	s4	0,0595	9
01:37:28 am	3/03/2017	s5	0,0874	11
07:44:27 pm	9/03/2017	s6	0,0736	6
11:42:13 pm	9/03/2017	s7	0,0595	87
01:51:30 am	17/03/2017	s8	0,0785	6
02:31:00 am	24/03/2017	s9	0,0448	4
01:39:58 am	7/04/2017	s10	0,0492	8
11:28:37 pm	20/04/2017	s11	0,0763	54
02:06:28 am	21/04/2017	s12	0,0559	1
01:12:04 am	28/04/2017	s13	0,0630	29
12:29:26 am	5/05/2017	s14	0,0554	9
10:54:55 pm	9/05/2017	s15	0,0546	201
11:18:14 pm	9/05/2017	s16	0,0524	9
Total Sesiones		16	Total Registros	455

En la Tabla 47, en el campo centralidad el 0,0900 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s2 es la sesión a la que le corresponde este valor, en la Tabla 48 se muestra en detalle la sesión realizada por el estudiante2GB.

Tabla 48 . Sesión, s2, del estudiante2GB. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
12:54:17 am	24/02/2017	Contenidos	Introducción	Comunicaciones SIMCA	NULL
12:54:23 am	24/02/2017	Contenidos	Introducción	Examen en línea - simulacro	NULL
1:18:39 am	24/02/2017	Contenidos	Introducción	Examen en línea - simulacro	NULL
1:19:10 am	24/02/2017	problem_check	Introducción	Examen en línea - simulacro	NULL
1:22:15 am	24/02/2017	problem_check	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
1:24:22 am	24/02/2017	problem_check	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
1:25:23 am	24/02/2017	problem_check	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
1:28:39 am	24/02/2017	problem_check	U1T1: Ubicación de nuestro planeta en el universo	Recursos Web	NULL
1:34:57 am	24/02/2017	problem_check	U1T2: Los movimientos del planeta tierra	Simuladores	NULL
1:39:40 am	24/02/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Foros semana 02	NULL
1:40:16 am	24/02/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Simuladores	NULL
1:40:24 am	24/02/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Recursos Web	NULL

1:40:29 am	24/02/2017	Contenidos	discussion	NULL	NULL
1:41:05 am	24/02/2017	Contenidos	discussion	NULL	NULL
1:42:41 am	24/02/2017	Contenidos	discussion	NULL	NULL
1:48:55 am	24/02/2017	Contenidos	discussion	NULL	NULL

En la Tabla 48 se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, de la cual se obtienen los patrones de navegación del estudiante2GB, ver Tabla 49.

Tabla 49. Patrones de navegación del estudiante2GB. Fuente propia.

Hora	Patrones de navegación
12:54:17 am - 1:48:55 am	Introducción -> Comunicaciones SIMCA.
	Introducción (2) -> Examen en línea (2) – simulacro (2).
	Introducción -> Examen en línea – simulacro -> problem_check.
	U1T1: Ubicación de nuestro planeta en el universo (3) -> Simuladores (3) -> problem_check (3).
	U1T1: Ubicación de nuestro planeta en el universo -> Recursos Web -> problem_check.
	U1T2: Los movimientos del planeta tierra -> Simuladores -> problem_check.
	U1T2: Los movimientos del planeta tierra -> Foros semana 02.
	U1T2: Los movimientos del planeta tierra -> Simuladores.
	U1T2: Los movimientos del planeta tierra -> Recursos Web.
	discussion (4).

El estudiante2GB, tiene diez patrones de navegación, de los patrones sobre salen, segundo, cuarto y último patrón de navegación, porque la información es más relevante para estudiante2GB en el curso. El estudiante2GB, por lo general ingresa al curso en horas de la madrugada, tiene una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje. El estudiante presenta deshonestidad académica, porque está buscando respuestas correctas e incorrectas de los exámenes de simulacro en línea, para ello, lee el examen luego busca en el material de aprendizaje la respuesta y posteriormente envía el examen en línea de simulacro.

El comportamiento deshonesto que presenta el estudiante2GB, es similar al presentado por los estudiantes 1 y 2 del grupo A, con en esta información se realiza la siguiente recomendación al docente para el diseño del curso: tener presente la forma como se está realizando el examen de simulacro en línea y los exámenes en línea, porque los estudiantes están presentando deshonestidades académicas. El docente debe replantearse la forma en como está llevando a cabo este tipo de evaluaciones.

Por otra parte, el estudiante2GB, realiza consultas en el foro para realizar preguntas o estar pendiente de las publicaciones, utiliza material de aprendizaje como: simuladores, recursos web, foros y entre otros. Con los resultados de la matriz

estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante2GB, esta se ilustra en la Figura 19.

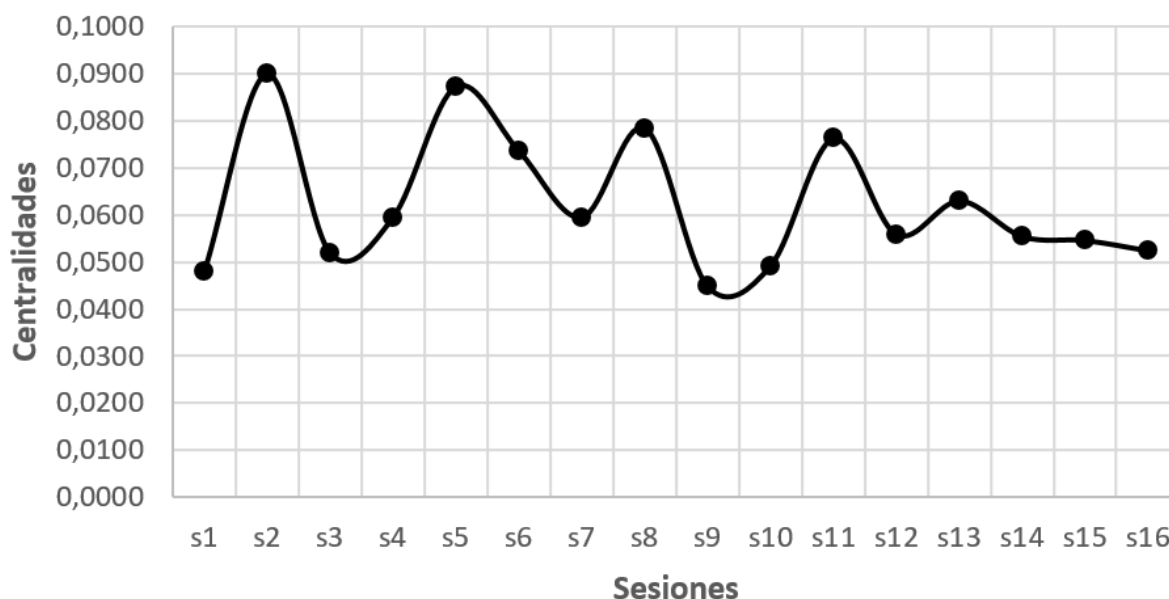


Figura 19. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante2GB. Fuente propia.

La gráfica permite observar cómo fue la interacción del estudiante con el curso, el estudiante2GB, tiene una participación muy activa hasta la sesión 11, desde ahí su interacción con el curso empieza a disminuir. Por otra parte, con la gráfica se puede identificar el patrón de navegación EF-ISF por sesión de mayor centralidad al observar en la gráfica el punto, s2 con su valor 0,0900. La grafica del patrón, también permite concluir que la navegación realizada por el estudiante2GB a lo largo del curso fue una navegación aleatoria.

Estudiante3GB

El estudiante3GB, al interactuar con el curso realizo en total, 44 sesiones y 1148 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 50.

Tabla 50. Centralidad de los patrones de navegación EF-ISF por sesión del estudiante3GB. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros	Hora	Fecha	Sesión Id	Centralidad	Registros
10:17:42 pm	9/02/2017	s1	0,0197	3	06:23:16 pm	14/03/2017	s23	0,0157	2
06:30:33 am	12/02/2017	s2	0,0215	4	03:32:53 pm	21/03/2017	s24	0,0188	155
03:13:50 pm	12/02/2017	s3	0,0140	2	08:31:14 pm	22/03/2017	s25	0,0205	4

11:17:43 pm	13/02/2017	s4	0,0360	41	11:44:49 pm	22/03/2017	s26	0,0287	1	
04:50:10 pm	14/02/2017	s5	0,0196	1	01:28:34 am	24/03/2017	s27	0,0157	11	
04:51:51 pm	14/02/2017	s6	0,0164	9	05:45:34 am	4/04/2017	s28	0,0147	5	
05:43:52 pm	14/02/2017	s7	0,0291	3	05:21:08 pm	4/04/2017	s29	0,0232	65	
05:52:51 pm	14/02/2017	s8	0,0251	6	06:22:08 pm	4/04/2017	s30	0,0205	2	
07:01:11 pm	14/02/2017	s9	0,0185	1	01:40:39 am	7/04/2017	s31	0,0225	23	
05:42:14 am	15/02/2017	s10	0,0215	5	02:15:44 am	18/04/2017	s32	0,0169	4	
10:34:31 pm	16/02/2017	s11	0,0276	6	08:53:48 pm	19/04/2017	s33	0,0251	79	
12:45:13 am	21/02/2017	s12	0,0283	3	01:35:59 am	21/04/2017	s34	0,0199	3	
02:22:47 am	24/02/2017	s13	0,0219	4	06:31:01 pm	24/04/2017	s35	0,0197	1	
10:52:26 pm	27/02/2017	s14	0,0287	6	04:03:43 am	26/04/2017	s36	0,0182	1	
05:44:11 pm	28/02/2017	s15	0,0443	25	07:36:01 pm	26/04/2017	s37	0,0141	12	
11:26:28 pm	2/03/2017	s16	0,0291	6	11:54:47 pm	27/04/2017	s38	0,0285	243	
01:32:11 am	3/03/2017	s17	0,0274	5	09:37:25 pm	28/04/2017	s39	0,0201	5	
08:59:15 pm	4/03/2017	s18	0,0275	7	12:04:13 am	3/05/2017	s40	0,0182	1	
09:28:14 pm	7/03/2017	s19	0,0287	9	12:07:08 am	5/05/2017	s41	0,0170	197	
09:45:18 pm	8/03/2017	s20	0,0224	8	02:34:20 am	5/05/2017	s42	0,0272	1	
10:52:14 pm	9/03/2017	s21	0,0284	160	06:20:20 pm	6/05/2017	s43	0,0202	1	
01:31:15 am	10/03/2017	s22	0,0216	4	08:39:54 pm	8/05/2017	s44	0,0172	14	
Total Sesiones								44	Total Registros	1148

En la Tabla 50, en el campo centralidad el 0,0443 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s15 es la sesión a la que le corresponde este valor, en la Tabla 51 se muestra en detalle la sesión realizada por el estudiante3GB.

Tabla 51. Sesión s15, del estudiante3GB. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
5:44:11 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Diapositivas	NULL
5:45:59 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Videos	NULL
5:46:55 pm	28/02/2017	play_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V02 La ubicación de nuestro planeta en el universo
5:47:05 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V02 La ubicación de nuestro planeta en el universo
5:47:30 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V02 La ubicación de nuestro planeta en el universo
5:47:42 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V02 La ubicación de nuestro planeta en el universo
5:48:08 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Videos complementarios	NULL
5:48:37 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Recursos Web	NULL
5:49:00 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
5:49:07 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Videos	NULL
5:49:42 pm	28/02/2017	play_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V04 La ubicación de nuestro planeta en el universo

5:49:52 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V04 La ubicación de nuestro planeta en el universo
5:49:58 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V04 La ubicación de nuestro planeta en el universo
5:50:09 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V04 La ubicación de nuestro planeta en el universo
5:50:19 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V04 La ubicación de nuestro planeta en el universo
5:50:41 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V04 La ubicación de nuestro planeta en el universo
6:12:07 pm	28/02/2017	play_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V01 La ubicación de nuestro planeta en el universo
6:12:08 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos	C01-V04 La ubicación de nuestro planeta en el universo
6:12:14 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Diapositivas	NULL
6:18:55 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Videos complementarios	NULL
6:19:38 pm	28/02/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Videos complementarios	NULL
6:20:00 pm	28/02/2017	play_video	U1T1: Ubicación de nuestro planeta en el universo	Videos complementarios	The Known Universe by AMNH
6:20:05 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos complementarios	The Known Universe by AMNH
6:20:15 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos complementarios	The Known Universe by AMNH
6:20:57 pm	28/02/2017	pause_video	U1T1: Ubicación de nuestro planeta en el universo	Videos complementarios	The Known Universe by AMNH

En la Tabla 51 se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, de la cual se obtienen los patrones de navegación del estudiante3GB, ver Tabla 52.

Tabla 52. Patrones de navegación del estudiante3GB. Fuente propia.

Hora	Patrones de navegación
5:44:11 pm - 6:20:57 pm	U1T1: Ubicación de nuestro planeta en el universo -> Diapositivas.
	U1T1: Ubicación de nuestro planeta en el universo -> Videos.
	U1T1: Ubicación de nuestro planeta en el universo -> Videos -> C01-V02 La ubicación de nuestro planeta en el universo -> play_video.
	U1T1: Ubicación de nuestro planeta en el universo (3) -> Videos (3) -> C01-V02 La ubicación de nuestro planeta en el universo (3) -> pause_video (3).
	U1T1: Ubicación de nuestro planeta en el universo -> Videos complementarios.
	U1T1: Ubicación de nuestro planeta en el universo -> Recursos Web.
	U1T1: Ubicación de nuestro planeta en el universo -> Simuladores.
	U1T1: Ubicación de nuestro planeta en el universo -> Videos.
	U1T1: Ubicación de nuestro planeta en el universo -> Videos -> C01-V04 La ubicación de nuestro planeta en el universo -> play_video.
	U1T1: Ubicación de nuestro planeta en el universo (6) -> Videos (6) -> C01-V04 La ubicación de nuestro planeta en el universo (6) -> pause_video (6).
	U1T1: Ubicación de nuestro planeta en el universo -> Videos -> C01-V01 La ubicación de nuestro planeta en el universo -> play_video.
	U1T1: Ubicación de nuestro planeta en el universo -> Diapositivas.
	U1T1: Ubicación de nuestro planeta en el universo (2) -> Videos complementarios (2).
	U1T1: Ubicación de nuestro planeta en el universo -> Videos complementarios -> The Known Universe by AMNH -> play_video.
	U1T1: Ubicación de nuestro planeta en el universo (3) -> Videos complementarios (3) -> The Known Universe by AMNH (3) -> pause_video (3).

El estudiante3GB, tiene quince patrones de navegación, de los que sobre salen, cuarto, diez, trece y último de los patrones de navegación, porque la información es más relevante para estudiante3GB en el curso. El estudiante3GB por lo general

ingresa al curso en horas de la tarde, presenta una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje como: diapositivas, videos, videos complementarios, recursos web y simuladores. Refuerza el aprendizaje en el curso revisando material complementario, en este caso video. Con los resultados de la matriz estacionaria, se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante3GB, esta se ilustra en la Figura 20.

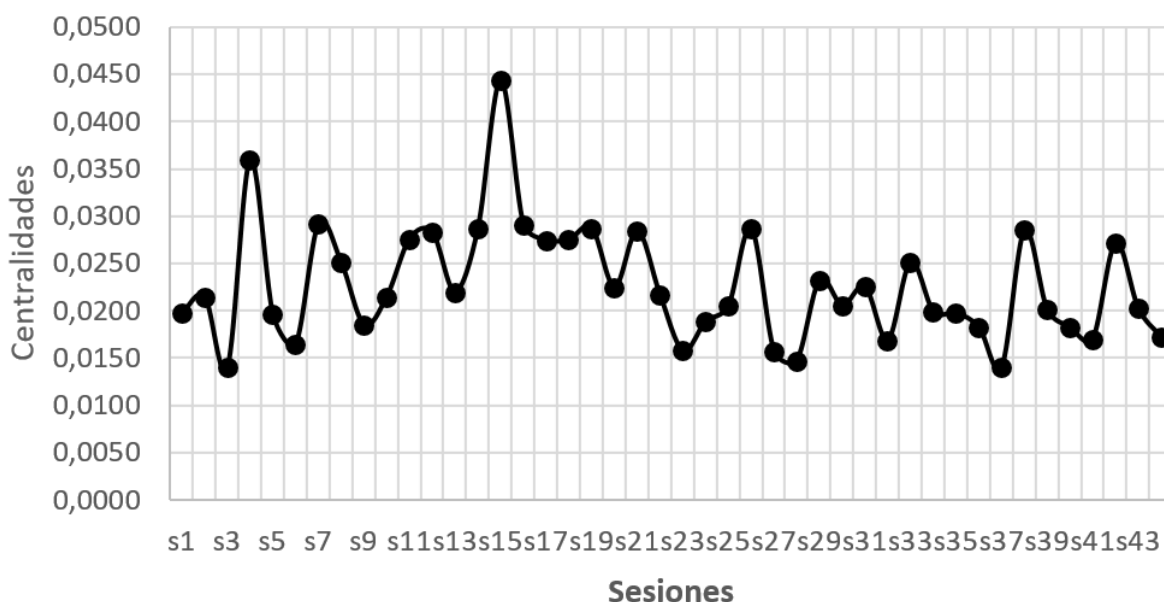


Figura 20. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudiante3GB. Fuente propia.

La gráfica permite observar cómo fue la interacción del estudiante3GB en el curso, inicia la interacción con crecimiento en la participación; pero descende su participación de una manera notable en la sesión 3, este nivel de participación lo alcanza varias veces a lo largo del curso; tiene una participación regular en todo el curso, su participación es notable en el patrón de navegación EF-ISF por sesión de mayor centralidad, al observar en la gráfica el punto máximo, es decir, s15 con su valor 0,0443. Además, en la gráfica del patrón se evidencia una navegación aleatoria por parte del estudiante3GB, el cual coincide con el análisis de la sesión 15.

5.2.2 Administrador AdministradorAGB

El administradorAGB, al interactuar con el curso realizó en total, 11 sesiones y 33 registros; a la información se le realiza un pre procesamiento, luego es procesada

por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 53.

Tabla 53. Centralidad de los patrones de navegación EF-ISF por sesión del administradorAGB. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
05:07:21 pm	8/02/2017	s1	0,0845	10
05:28:11 am	13/02/2017	s2	0,0845	4
11:34:05 pm	14/02/2017	s3	0,0845	2
04:25:46 am	31/03/2017	s4	0,0832	1
11:27:43 pm	7/04/2017	s5	0,1154	2
03:51:14 am	17/04/2017	s6	0,0909	5
06:09:56 pm	18/04/2017	s7	0,0961	2
06:14:55 pm	2/05/2017	s8	0,0832	3
05:37:21 pm	26/05/2017	s9	0,0961	1
10:44:36 am	27/05/2017	s10	0,0909	1
08:42:05 pm	18/07/2017	s11	0,0909	2
Total Sesiones		11	Total Registros	33

En la Tabla 53, en el campo centralidad el 0,1154 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s5 es la sesión a la que le corresponde este valor, en la Tabla 54 se muestra en detalle la sesión realizada por el administradorAGB.

Tabla 54. Sesión s5, del administradorAGB. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
11:27:43 pm	7/04/2017	Contenidos	U2T2: Modelos del Universo	Examen en Línea	NULL
12:24:56 am	8/04/2017	Contenidos	Introducción	Plan Curricular	NULL

Al valor cuantitativo 0,1154 de s5, se le realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad. El administrador tiene dos patrones de navegación, frecuentemente ingresa al curso

en altas horas de la noche y en la madrugada, por lo general ingresa a la plataforma a revisar los exámenes en línea, ya sea para prepararlos o estar pendiente de ellos cuando los estudiantes los estén presentando, también ha estado pendiente del plan curricular del curso, ha estado visitando varias veces la sección introducción. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del administradorAGB, esta se ilustra en la Figura 21.

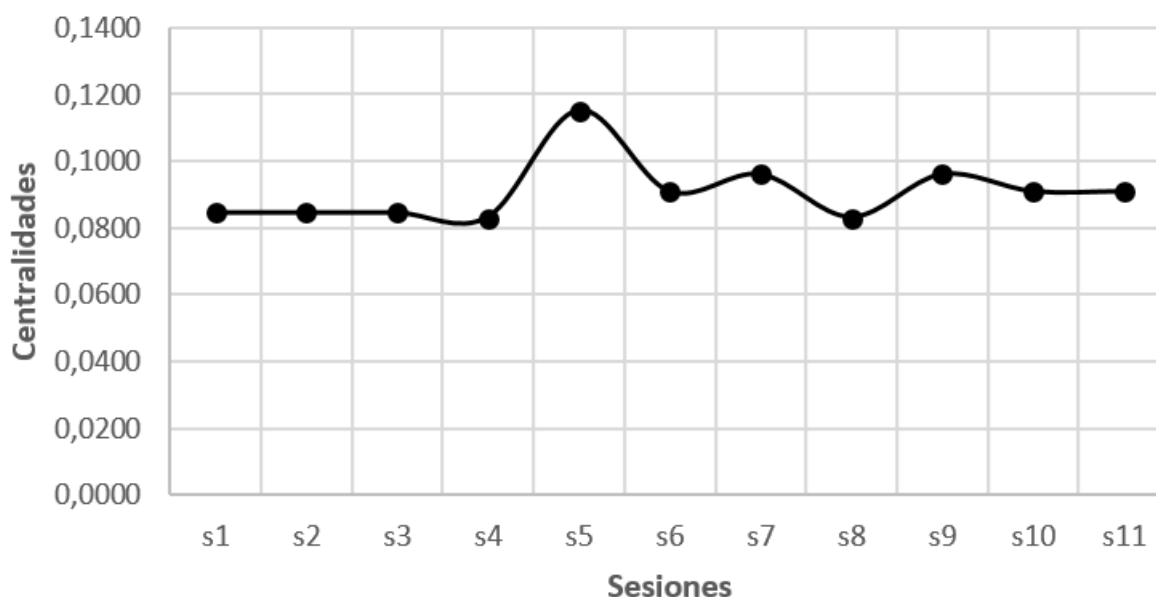


Figura 21. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del administradorAGB. Fuente propia.

La gráfica permite observar cómo fue la interacción del administradorAGB en el curso, el administradorAGB, tiene una navegación constante de s1 hasta la s3; estas interacciones son bajas con respecto a las demás interacciones del curso; de s3 a s4 decrece a un más su interacción, de s10 a s11 presenta una navegación constante. A partir de la gráfica también se puede concluir que la navegación del administradorAGB es aleatoria y el patrón de navegación EF-ISF por sesión de mayor centralidad se observa en el punto, s5 con su valor 0,1154.

Administrador B

La interpretación para el administrador B, no se realiza porque tiene solo dos sesiones, ver descripción del conjunto de datos en la sección 3.2.2.1 para mayor detalle del administrador B.

5.2.3 ProfesorGB

El profesorGB, al interactuar con el curso realizo en total, 56 sesiones y 140 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 55.

Tabla 55. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGB. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros	Hora	Fecha	Sesión Id	Centralidad	Registros
07:58:03 pm	8/02/2017	s1	0,0155	4	02:02:22 am	10/03/2017	s29	0,0122	1
02:26:20 pm	9/02/2017	s2	0,0155	1	05:28:24 pm	14/03/2017	s30	0,0170	1
07:58:21 pm	9/02/2017	s3	0,0155	2	05:31:22 pm	14/03/2017	s31	0,0088	2
01:56:24 pm	10/02/2017	s4	0,0150	3	12:13:17 pm	20/03/2017	s32	0,0320	6
07:59:42 pm	12/02/2017	s5	0,0155	2	04:13:58 pm	23/03/2017	s33	0,0184	8
12:33:55 am	13/02/2017	s6	0,0255	2	01:26:23 am	24/03/2017	s34	0,0159	2
07:44:25 pm	13/02/2017	s7	0,0155	2	02:47:19 pm	24/03/2017	s35	0,0171	1
02:11:29 pm	14/02/2017	s8	0,0155	1	07:28:02 pm	3/04/2017	s36	0,0145	1
04:38:40 pm	14/02/2017	s9	0,0380	4	05:41:26 pm	4/04/2017	s37	0,0149	1
02:56:27 am	17/02/2017	s10	0,0155	1	02:20:47 pm	5/04/2017	s38	0,0110	1
03:07:08 pm	17/02/2017	s11	0,0155	1	01:40:57 am	7/04/2017	s39	0,0149	2
02:15:06 pm	20/02/2017	s12	0,0154	2	10:37:53 am	21/04/2017	s40	0,0171	1
02:33:36 pm	20/02/2017	s13	0,0155	1	02:01:49 pm	24/04/2017	s41	0,0111	1
10:13:20 am	21/02/2017	s14	0,0316	4	08:09:37 pm	24/04/2017	s42	0,0095	2
01:48:32 pm	22/02/2017	s15	0,0274	6	05:34:55 pm	25/04/2017	s43	0,0174	1
01:41:15 pm	23/02/2017	s16	0,0110	1	05:07:25 pm	27/04/2017	s44	0,0195	1
07:16:40 pm	23/02/2017	s17	0,0155	1	05:36:17 pm	2/05/2017	s45	0,0190	1
01:28:20 am	24/02/2017	s18	0,0155	1	06:14:33 pm	3/05/2017	s46	0,0179	1
03:49:29 pm	24/02/2017	s19	0,0155	1	08:44:35 am	4/05/2017	s47	0,0190	7
03:23:26 pm	25/02/2017	s20	0,0155	1	09:05:37 am	4/05/2017	s48	0,0215	4
02:11:59 pm	27/02/2017	s21	0,0213	2	09:58:02 am	5/05/2017	s49	0,0190	2
09:20:31 pm	27/02/2017	s22	0,0155	1	11:09:46 am	5/05/2017	s50	0,0190	13
10:39:23 pm	27/02/2017	s23	0,0176	2	02:32:02 pm	5/05/2017	s51	0,0179	2
12:15:34 pm	28/02/2017	s24	0,0155	1	05:57:58 pm	5/05/2017	s52	0,0220	3
05:59:33 pm	28/02/2017	s25	0,0122	1	10:01:12 am	8/05/2017	s53	0,0154	4
01:33:20 am	3/03/2017	s26	0,0240	2	05:20:26 pm	9/05/2017	s54	0,0118	4
01:34:26 pm	3/03/2017	s27	0,0155	2	09:11:20 am	10/05/2017	s55	0,0163	2
05:42:21 pm	7/03/2017	s28	0,0247	2	06:29:23 pm	22/05/2017	s56	0,0383	11
Total Sesiones							56	Total Registros	140

En la Tabla 55, en el campo centralidad el 0,0383 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s56 es la sesión a la que le corresponde este valor, en la Tabla 56 se muestra en detalle la sesión realizada por el profesorGB.

Tabla 56. Sesión s56, del profesorGB. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
6:29:23 pm	22/05/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Examen en Línea	NULL
6:41:51 pm	22/05/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Examen en Línea	NULL
6:43:06 pm	22/05/2017	Contenidos	U1T4: La estructura del sistema solar	Examen en Línea	NULL
6:44:20 pm	22/05/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Examen en Línea	NULL
6:46:22 pm	22/05/2017	Contenidos	Examen extra U1: Fundamentos de Astronomía	Examen Extra U1	NULL
6:50:14 pm	22/05/2017	Contenidos	U2T1: Astronomía y los albores de la civilización	Examen en Línea	NULL
6:50:43 pm	22/05/2017	Contenidos	U2T1: Astronomía y los albores de la civilización	Examen en Línea	NULL
6:51:05 pm	22/05/2017	Contenidos	U2T2: Modelos del Universo	Examen en Línea	NULL
6:51:21 pm	22/05/2017	Contenidos	U2T3: Galileo Galilei	Examen en Línea	NULL
6:51:39 pm	22/05/2017	Contenidos	U2T4: Astronomía Moderna	Examen en Línea	NULL
6:52:04 pm	22/05/2017	Contenidos	U3T1: Calendarios	Examen en Línea 11	NULL

En la Tabla 56 se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, de la cual se obtienen los patrones de navegación del profesorGB, ver Tabla 57.

Tabla 57. Patrones de navegación del profesorGB. Fuente propia.

Hora	Patrones de navegación
6:29:23 pm - 6:52:04 pm	U1T2: Los movimientos del planeta tierra -> Examen en Línea.
	U1T1: Ubicación de nuestro planeta en el universo -> Examen en Línea.
	U1T4: La estructura del sistema solar -> Examen en Línea.
	U1T5: Exoplanetas y la zona de habitabilidad -> Examen en Línea.
	Examen extra U1: Fundamentos de Astronomía -> Examen Extra U1.
	U2T1: Astronomía y los albores de la civilización -> Examen en Línea.
	U2T1: Astronomía y los albores de la civilización -> Examen en Línea.
	U2T2: Modelos del Universo -> Examen en Línea.
	U2T3: Galileo Galilei -> Examen en Línea.
	U2T4: Astronomía Moderna -> Examen en Línea.
	U3T1: Calendarios -> Examen en Línea 11.

El profesorGB, tiene once patrones de navegación, navega por diferentes secciones y el material de aprendizaje más utilizado son los exámenes en línea y exámenes en línea extra. El profesorGB por lo general ingresa al curso en horas de la tarde empezando la noche, sus ingresos frecuentemente son para las evaluaciones o evaluaciones extras, tiene una navegación aleatoria, porque navega en diferentes secciones y unidades; está pendiente de los exámenes que están desarrollando los estudiantes o está preparando el contenido a publicar en la plataforma para las evaluaciones. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGB, esta se ilustra en la Figura 22.

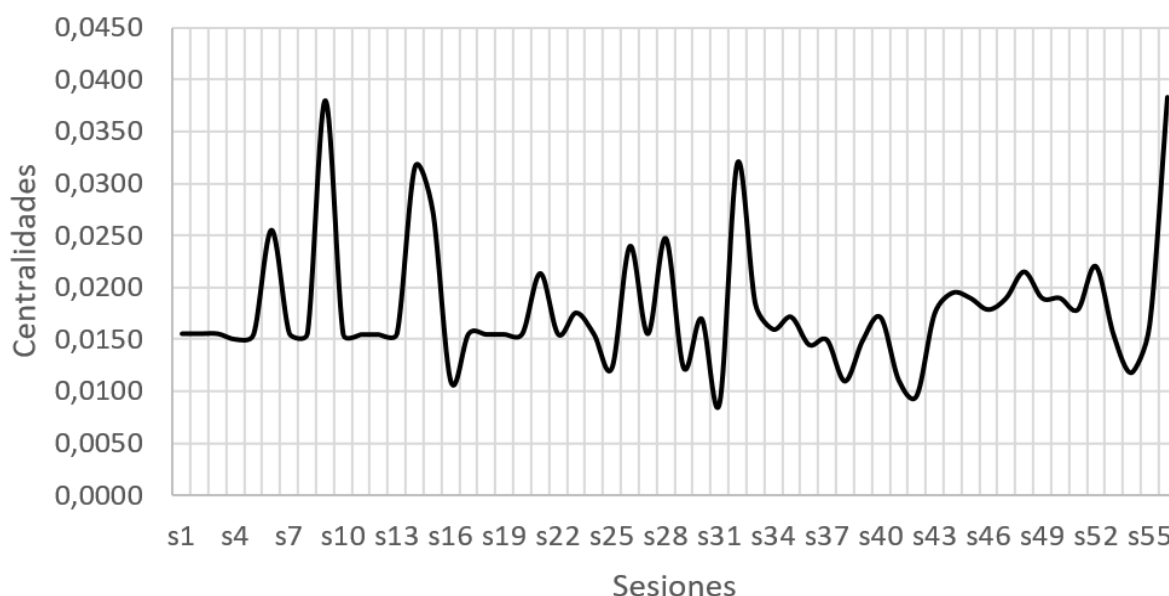


Figura 22. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGB. Fuente propia.

La gráfica permite observar cómo fue la interacción del profesorGB en el curso, en el inicio la interacción con el curso es constante, tiene ciclos de incrementos y decrementos en el transcurso del curso, con lo cual se concluye que la navegación del profesorGB es aleatoria. Al finalizar la navegación del profesorGB hay un incremento notable y el punto máximo de la función coincide con el patrón de navegación EF-ISF por sesión de mayor centralidad, es decir, s56 con su valor 0,0383.

5.2.4 MonitorGB

El monitorGB, al interactuar con el curso realizó en total, 19 sesiones y 54 registros; a la información se le realiza un pre procesamiento, luego es procesada por el

algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 58.

Tabla 58. Centralidad de los patrones de navegación EF-ISF por sesión del monitorGB. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
01:31:08 am	24/02/2017	s1	0,0526	2
01:44:17 am	3/03/2017	s2	0,0428	1
05:28:15 pm	7/03/2017	s3	0,0764	1
01:34:36 am	10/03/2017	s4	0,0764	1
01:31:40 am	17/03/2017	s5	0,0631	4
05:26:49 pm	21/03/2017	s6	0,0463	2
07:49:00 pm	22/03/2017	s7	0,0519	1
01:28:32 am	24/03/2017	s8	0,0419	1
04:55:46 pm	2/04/2017	s9	0,0768	19
06:18:14 pm	3/04/2017	s10	0,0567	7
05:08:01 pm	4/04/2017	s11	0,0428	1
01:24:28 am	7/04/2017	s12	0,0526	2
05:27:26 pm	18/04/2017	s13	0,0402	1
02:15:07 pm	21/04/2017	s14	0,0526	1
04:09:50 pm	21/04/2017	s15	0,0402	1
02:08:46 am	23/04/2017	s16	0,0411	6
05:31:18 pm	25/04/2017	s17	0,0402	1
01:44:46 am	6/05/2017	s18	0,0526	1
05:35:12 pm	9/05/2017	s19	0,0526	1
Total Sesiones		19	Total Registros	54

En la Tabla 58, en el campo centralidad el 0,0768 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s9 es la sesión a la que le corresponde este valor, en la Tabla 59 se muestra en detalle la sesión realizada por el monitorGB.

Tabla 59. Sesión s9, del monitorGB. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
4:55:46 pm	2/04/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
4:58:13 pm	2/04/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Recursos Web	NULL
5:05:39 pm	2/04/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Foros semana 01	NULL
5:07:09 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Simuladores	NULL
5:11:46 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Recursos Web	NULL
5:14:55 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Taller	NULL
5:15:16 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Foros semana 02	NULL

5:17:03 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Simuladores	NULL
5:29:32 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Página WEB	NULL
5:35:05 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Taller	NULL
5:35:48 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Foros semana 03	NULL
7:05:30 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Simuladores	NULL
7:10:56 pm	2/04/2017	Contenidos	U1T4: La estructura del sistema solar	Simuladores	NULL
7:13:29 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Simuladores	NULL
7:23:05 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Recursos Web	NULL
7:27:05 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Foros Semana 05	NULL
7:28:56 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Recursos Web	NULL
7:31:06 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Foros Semana 05	NULL
7:34:12 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Examen en Línea	NULL

En la Tabla 59 se realiza una interpretación cualitativa, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, de la cual se obtienen los patrones de navegación del monitorGB, ver Tabla 60.

Tabla 60. Patrones de navegación del monitorGB. Fuente propia.

Hora	Patrones de navegación
4:55:46 pm - 7:34:12 pm	U1T1: Ubicación de nuestro planeta en el universo -> Simuladores.
	U1T1: Ubicación de nuestro planeta en el universo -> Recursos Web.
	U1T1: Ubicación de nuestro planeta en el universo -> Foros semana 01.
	U1T2: Los movimientos del planeta tierra -> Simuladores.
	U1T2: Los movimientos del planeta tierra -> Recursos Web.
	U1T2: Los movimientos del planeta tierra -> Taller.
	U1T2: Los movimientos del planeta tierra -> Foros semana 02.
	U1T3: Principios de astronomía de observación (2) -> Simuladores (2).
	U1T3: Principios de astronomía de observación -> Página WEB.
	U1T3: Principios de astronomía de observación -> Taller.
	U1T3: Principios de astronomía de observación -> Foros semana 03.
	U1T4: La estructura del sistema solar -> Simuladores.
	U1T5: Exoplanetas y la zona de habitabilidad -> Simuladores.
	U1T5: Exoplanetas y la zona de habitabilidad (2) -> Recursos Web (2).
	U1T5: Exoplanetas y la zona de habitabilidad (2) -> Foros Semana 05 (2).
	U1T5: Exoplanetas y la zona de habitabilidad -> Examen en Línea.

El monitorGB, tiene dieciséis patrones de navegación, de los que sobre salen, octavo, catorce y quince de los patrones de navegación, porque la información es más relevante para él en el curso. Ingres a al curso en horas de la tarde empezando la noche, tiene una navegación aleatoria, ya que navega en diferentes secciones,

unidades y hace uso de diferente material de aprendizaje. El material de aprendizaje más utilizado es: simuladores, recursos Web, foros, talleres, página Web y examen en línea. Está pendiente de los foros, para responder a los estudiantes alguna consulta o a los demás usuarios; en el examen en línea, está pendiente de los estudiantes. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del monitorGB, esta se ilustra en la Figura 23.

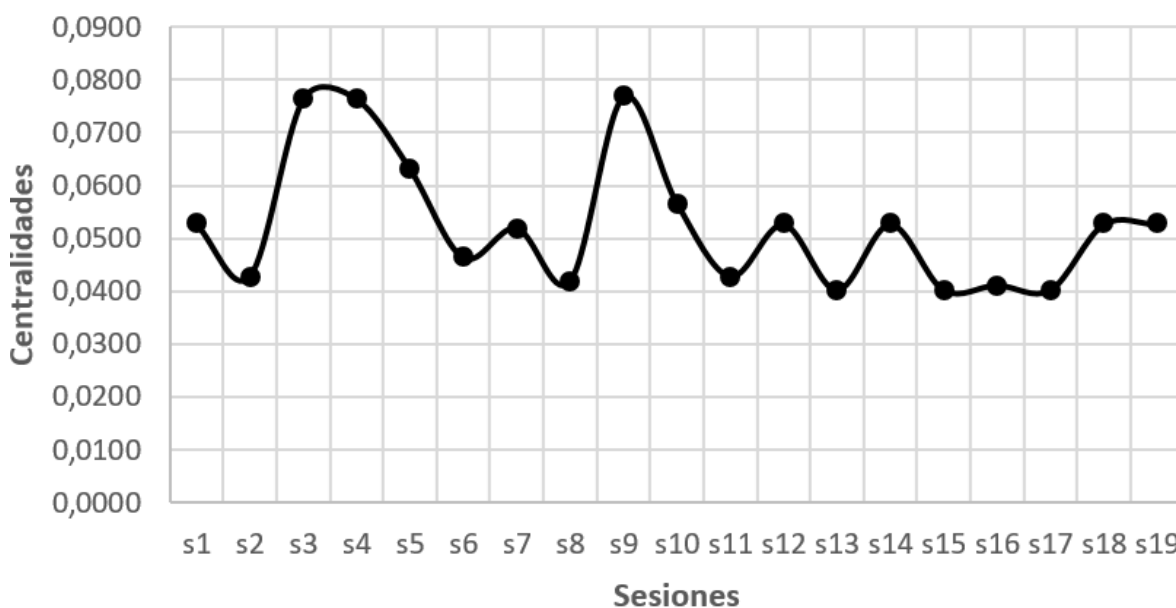


Figura 23. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del monitorGB. Fuente propia.

La gráfica permite observar cómo fue la interacción del monitorGB en el transcurso del curso. De la sesión 1 a la 2 tuvo un decrecimiento en la participación en el curso, los puntos máximos de la gráfica, indican la mayor interacción con el curso. El material de aprendizaje de mayor uso es el siguiente: simuladores, recursos Web, foros, talleres, página Web y examen en línea. En las sesiones 15, 16 y 17 su interacción es baja, no estuvo tan pendiente del curso, en la sesión 18 empieza a descender la participación. Su navegación es aleatoria y el patrón de navegación EF-ISF por sesión de mayor centralidad es el punto, s9 con su valor 0,0768.

5.2.5 Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5

En este apartado se realizan análisis a algunos usuarios del archivo Excel, Astronomía Cotidiana 2017-I Grupo B U1T1 a U1T5, que contiene interacciones de los usuarios de la unidad uno temática uno a la unidad uno temática cinco, para más detalle ver la sección 3.2.2.2.

5.2.5.1 EstudianteGBT1T5

El estudianteGBT1T5 con número de sesiones alto, al interactuar con el curso realizo en total, 32 sesiones y 740 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 61.

Tabla 61. Centralidad de los patrones de navegación EF-ISF por sesión del estudianteGBT1T5. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros	Hora	Fecha	Sesión Id	Centralidad	Registros
07:54:14 pm	14/02/2017	s1	0,0359	42	01:47:55 pm	20/03/2017	s17	0,0293	5
11:54:01 pm	14/02/2017	s2	0,0275	3	12:07:40 am	22/03/2017	s18	0,0380	32
12:15:29 am	21/02/2017	s3	0,0390	23	12:49:05 am	23/03/2017	s19	0,0280	5
01:02:46 am	22/02/2017	s4	0,0379	45	01:44:50 pm	23/03/2017	s20	0,0241	1
11:27:40 pm	23/02/2017	s5	0,0397	33	09:35:45 pm	23/03/2017	s21	0,0279	6
11:29:22 pm	27/02/2017	s6	0,0356	39	10:26:09 pm	23/03/2017	s22	0,0405	10
12:39:06 pm	28/02/2017	s7	0,0159	19	11:14:57 pm	23/03/2017	s23	0,0367	155
01:13:20 pm	1/03/2017	s8	0,0257	5	06:14:59 pm	24/03/2017	s24	0,0241	1
12:21:36 am	3/03/2017	s9	0,0387	8	10:33:46 pm	25/03/2017	s25	0,0444	3
01:34:40 am	3/03/2017	s10	0,0249	9	01:24:35 pm	4/04/2017	s26	0,0330	4
12:09:59 pm	9/03/2017	s11	0,0251	7	02:31:23 pm	5/04/2017	s27	0,0215	1
10:01:32 pm	9/03/2017	s12	0,0289	28	01:51:33 am	7/04/2017	s28	0,0368	2
11:30:36 pm	9/03/2017	s13	0,0391	42	11:24:14 pm	8/04/2017	s29	0,0233	1
08:32:41 pm	14/03/2017	s14	0,0283	145	11:37:12 pm	8/04/2017	s30	0,0233	3
12:31:50 pm	16/03/2017	s15	0,0363	10	12:34:33 am	19/04/2017	s31	0,0390	3
12:31:11 am	17/03/2017	s16	0,0283	49	02:59:46 pm	27/04/2017	s32	0,0233	1
Total sesiones							32	Total Registros	740

En la Tabla 61, en el campo centralidad el 0,0444 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s25 es la sesión a la que le corresponde este valor, en la Tabla 62 se muestra en detalle la sesión realizada por el estudianteGBT1T5.

Tabla 62. Sesión s25, del estudianteGBT1T5. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
10:33:46 pm	25/03/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Simuladores	NULL
10:33:5 pm	25/03/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Recursos Web	NULL
10:34:13 pm	25/03/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Recursos Web	NULL

Por el valor cuantitativo 0,0444 se le realiza una interpretación cualitativa a la sesión 25, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, ver Tabla 63

Tabla 63. Patrones de navegación del estudianteGBT1T5. Fuente propia.

Hora	Patrones de navegación
10:33:46 pm	U1T5: Exoplanetas y la zona de habitabilidad -> Simuladores.
-	U1T5: Exoplanetas y la zona de habitabilidad -> Recursos Web.
10:34:13 pm	U1T1: Ubicación de nuestro planeta en el universo -> Recursos Web.

El estudianteGBT1T5 tiene tres patrones de navegación, por lo general ingresa al curso en horas de la noche, simuladores y recursos web es el material de aprendizaje más utilizado, tiene una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudianteGBT1T5, esta se ilustra en la Figura 24.

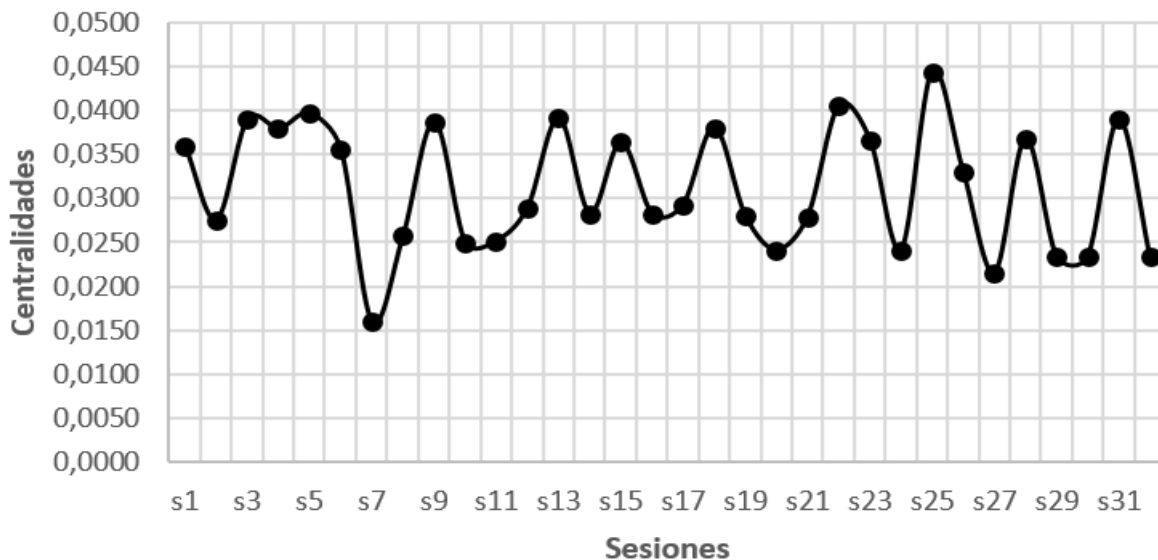


Figura 24. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del estudianteGBT1T5. Fuente propia.

La gráfica permite observar cómo fue la interacción del estudianteGBT1T5 en las unidades temáticas del curso, tiene una navegación aleatoria, decreciente y creciente a lo largo del curso, con su punto mínimo de interacción en el curso en la sesión 7, s7; alcanza el patrón de navegación EF-ISF por sesión de mayor centralidad del estudianteGBT1T5 en la sesión 25 y centralidad 0,0444.

5.2.5.2 Administrador

El análisis al administradorAGB no se realiza, porque en estas unidades temáticas solo aparece con una sesión, su interacción con el curso fue baja y coincide con la interpretación de baja participación que se evidencio en la gráfica del patrón de navegación EF-ISF por sesión del administradorAGB en el curso completo, ver sección 5.2.2.

5.2.5.3 ProfesorGBT1T5

El profesorGBT1T5, al interactuar con el curso realizo en total, 10 sesiones y 20 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 64.

Tabla 64. Centralidad de los patrones de navegación EF-ISF por sesión del profesorGBT1T5. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
04:38:40 pm	14/02/2017	s1	0,0974	1
02:15:06 pm	20/02/2017	s2	0,0851	2
01:33:20 am	3/03/2017	s3	0,0981	1
05:42:21 pm	7/03/2017	s4	0,0946	1
05:28:24 pm	14/03/2017	s5	0,0946	1
05:31:22 pm	14/03/2017	s6	0,0847	2
12:13:17 pm	20/03/2017	s7	0,1328	3
04:15:53 pm	23/03/2017	s8	0,0804	4
01:26:23 am	24/03/2017	s9	0,0740	1
06:29:23 pm	22/05/2017	s10	0,1583	4
Total Sesiones		10	Total Registros	20

En la Tabla 64, en el campo centralidad el 0,1583 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s10 es la sesión a la que le corresponde este valor, en la Tabla 65 se muestra en detalle la sesión realizada por el profesorGBT1T5.

Tabla 65. Sesión s10, del profesorGBT1T5. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
6:29:23 pm	22/05/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Examen en Línea	NULL
6:41:51 pm	22/05/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Examen en Línea	NULL
6:43:06 pm	22/05/2017	Contenidos	U1T4: La estructura del sistema solar	Examen en Línea	NULL

6:44:20 pm	22/05/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Examen en Línea	NULL
------------	------------	------------	--	-----------------	------

Por el valor cuantitativo 0,1583 se realiza una interpretación cualitativa a la sesión 10, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, ver Tabla 66.

Tabla 66. Patrones de navegación del profesorGBT1T5. Fuente propia.

Hora	Patrones de navegación
6:29:23 pm	U1T2: Los movimientos del planeta tierra -> Examen en Línea.
-	U1T1: Ubicación de nuestro planeta en el universo -> Examen en Línea.
6:44:20 pm	U1T4: La estructura del sistema solar -> Examen en Línea.
	U1T5: Exoplanetas y la zona de habitabilidad -> Examen en Línea.

El profesorGBT1T5, tiene cuatro patrones de navegación, por lo general ingresa al curso en horas de la tarde, tiene una navegación aleatoria, porque navega en diferentes secciones y unidades; frecuente mente, está pendiente de las evaluaciones de los estudiantes ya sea para evaluarlos o preparando el examen en la plataforma. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGBT1T5, esta se ilustra en la Figura 25.

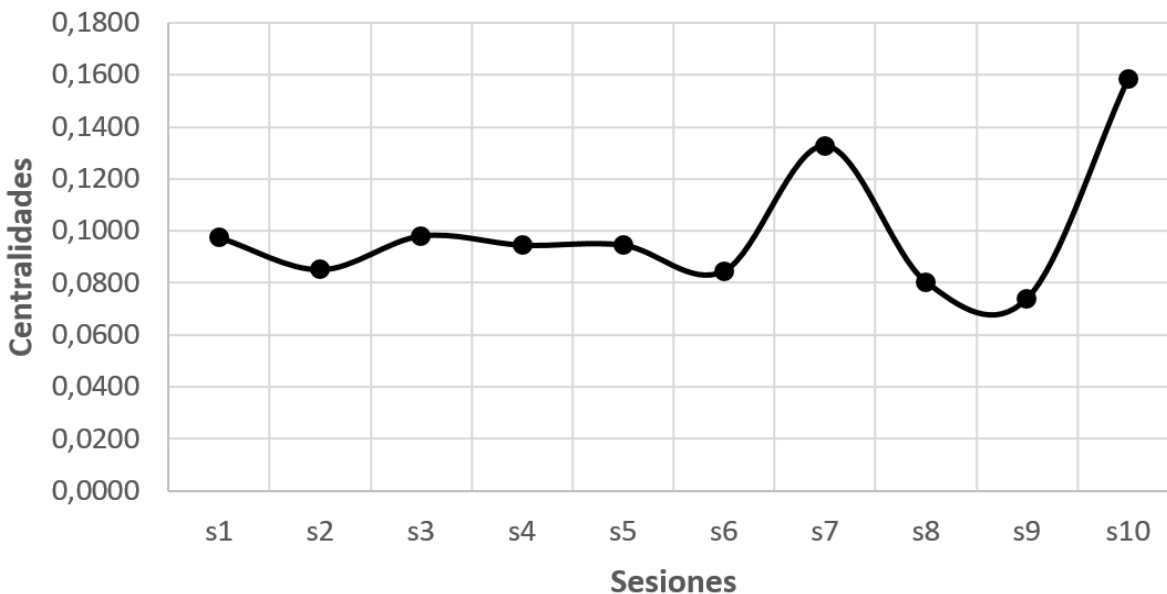


Figura 25. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del profesorGBT1T5. Fuente propia.

La gráfica permite observar cómo fue la interacción del profesorGBT1T5 en las unidades temáticas del curso, el profesorGBT1T5 tiene una navegación aleatoria. El patrón de navegación EF-ISF por sesión de mayor centralidad, se puede observar en la gráfica en el máximo, es decir, s10 con su valor 0,1583.

5.2.5.4 MonitorGBT1T5

El monitorGBT1T5, al interactuar con el curso realizó en total, 11 sesiones y 43 registros; a la información se le realiza un pre procesamiento, luego es procesada por el algoritmo LexRank con Umbral, generando como salida las centralidades de los patrones de navegación EF-ISF por sesión, estas se muestran en la Tabla 67.

Tabla 67. Centralidad de los patrones de navegación EF-ISF por sesión del monitorGBT1T5. Fuente propia.

Hora	Fecha	Sesión Id	Centralidad	Registros
01:44:17 am	3/03/2017	s1	0,0834	1
05:28:15 pm	7/03/2017	s2	0,0834	1
01:34:36 am	10/03/2017	s3	0,0834	1
01:31:40 am	17/03/2017	s4	0,1202	4
05:26:49 pm	21/03/2017	s5	0,0762	2
07:49:00 pm	22/03/2017	s6	0,0944	1
01:28:32 am	24/03/2017	s7	0,0762	1
04:55:46 pm	2/04/2017	s8	0,1331	19
06:18:14 pm	3/04/2017	s9	0,1073	7
05:08:01 pm	4/04/2017	s10	0,0834	1
02:08:46 am	23/04/2017	s11	0,0590	5
Total Sesiones		11	Total Registros	43

Al revisar la Tabla 67, en el campo centralidad el 0,1331 es el número mayor, representa el patrón de navegación EF-ISF por sesión de mayor centralidad, s8 es la sesión a la que le corresponde este valor, en la Tabla 68 se muestra en detalle la sesión realizada por el monitorGBT1T5.

Tabla 68. Sesión s8, del monitorGBT1T5. Adaptado (Jaramillo & Mario, 2017).

Hora	Fecha	Actividad	Complemento1	Complemento2	Complemento3
4:55:46 pm	2/04/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Simuladores	NULL
4:58:13 pm	2/04/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Recursos Web	NULL
5:05:39 pm	2/04/2017	Contenidos	U1T1: Ubicación de nuestro planeta en el universo	Foros semana 01	NULL
5:07:09 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Simuladores	NULL
5:11:46 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Recursos Web	NULL
5:14:55 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Taller	NULL
5:15:16 pm	2/04/2017	Contenidos	U1T2: Los movimientos del planeta tierra	Foros semana 02	NULL

5:17:03 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Simuladores	NULL
5:29:32 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Página WEB	NULL
5:35:05 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Taller	NULL
5:35:48 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Foros semana 03	NULL
7:05:30 pm	2/04/2017	Contenidos	U1T3: Principios de astronomía de observación	Simuladores	NULL
7:10:56 pm	2/04/2017	Contenidos	U1T4: La estructura del sistema solar	Simuladores	NULL
7:13:29 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Simuladores	NULL
7:23:05 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Recursos Web	NULL
7:27:05 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Foros Semana 05	NULL
7:28:56 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Recursos Web	NULL
7:31:06 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Foros Semana 05	NULL
7:34:12 pm	2/04/2017	Contenidos	U1T5: Exoplanetas y la zona de habitabilidad	Examen en Línea	NULL

Por el valor cuantitativo 0,1331 se realiza una interpretación cualitativa a la sesión 8, con la ponderación basada en la frecuencia relativa de un evento y la centralidad, ver Tabla 69.

Tabla 69. Patrones de navegación del monitorGBT1T5. Fuente propia.

Hora	Patrones de navegación
4:55:46 pm - 7:34:12 pm	U1T1: Ubicación de nuestro planeta en el universo -> Simuladores.
	U1T1: Ubicación de nuestro planeta en el universo -> Recursos Web.
	U1T1: Ubicación de nuestro planeta en el universo -> Foros semana 01.
	U1T2: Los movimientos del planeta tierra -> Simuladores.
	U1T2: Los movimientos del planeta tierra -> Recursos Web.
	U1T2: Los movimientos del planeta tierra -> Taller.
	U1T2: Los movimientos del planeta tierra -> Foros semana 02.
	U1T3: Principios de astronomía de observación (2) -> Simuladores (2).
	U1T3: Principios de astronomía de observación -> Página WEB.
	U1T3: Principios de astronomía de observación -> Taller.
	U1T3: Principios de astronomía de observación -> Foros semana 03.
	U1T4: La estructura del sistema solar -> Simuladores.
	U1T5: Exoplanetas y la zona de habitabilidad -> Simuladores.
	U1T5: Exoplanetas y la zona de habitabilidad (2) -> Recursos Web (2).
	U1T5: Exoplanetas y la zona de habitabilidad (2) -> Foros Semana 05 (2).
	U1T5: Exoplanetas y la zona de habitabilidad -> Examen en Línea.

El monitorGBT1T5, tiene dieciséis patrones de navegación, de los que sobre salen, octavo, catorce y quince de los patrones de navegación, porque la información es más relevante para él en el curso, por lo general ingresa al curso en horas de la tarde y empezando la noche, tiene una navegación aleatoria, porque navega en diferentes secciones, unidades y hace uso de diferente material de aprendizaje; generalmente ingresa a los foros para dar respuesta o proponer algún tema, está

pendiente de las evaluaciones que presentan los estudiantes, el material de aprendizaje más utilizado es el siguiente: simuladores, recursos Web, foros, talleres, página Web y examen en línea. Con los resultados de la matriz estacionaria se realiza la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del monitorGBT1T5, esta se ilustra en la Figura 26.

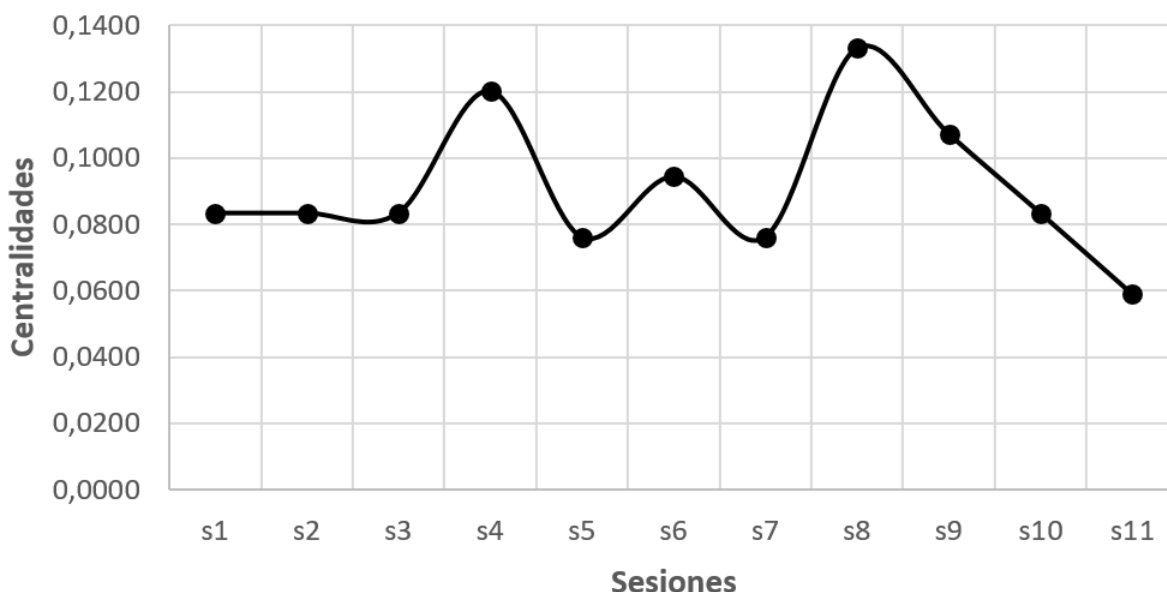


Figura 26. Gráfica de la centralidad del patrón de navegación EF-ISF por sesión del monitorGBT1T5. Fuente propia.

La gráfica permite observar cómo fue la interacción del monitorGBT1T5 en las unidades temáticas del curso, tiene una navegación aleatoria. Al inicio del curso tiene una navegación constante, en los puntos máximos es donde el monitorGBT1T5 tiene mayor interacción con el curso; desde la sesión s8 hasta s11, se observa que el monitor poco a poco deja de estar pendiente de la plataforma, esto coincide con la interpretación realizada en grafica del patrón en el curso completo, en las sesiones s15, s16 y s17. Por otra parte, en la gráfica también se identifica el patrón de navegación EF-ISF por sesión de mayor centralidad del monitorGBT1T5, al observar en la gráfica el punto máximo, es decir, s8 con su valor 0,1331.

5.3 Conclusiones

Del análisis realizado a los usuarios, con la sesión a la que le corresponde el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario y la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del usuario, se puede concluir lo siguiente:

Los usuarios navegan por diferentes secciones, unidades y material de aprendizaje, con lo cual se puede inferir que todos los usuarios tienen una navegación aleatoria, en la interacción realizada en el curso o en las unidades temáticas del curso. Esto se puede observar en la sesión a la que le corresponde el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario o en la gráfica de patrón. También se identificó en la gráfica, el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario, siendo este un punto máximo de la gráfica del patrón, ver Figura 27.

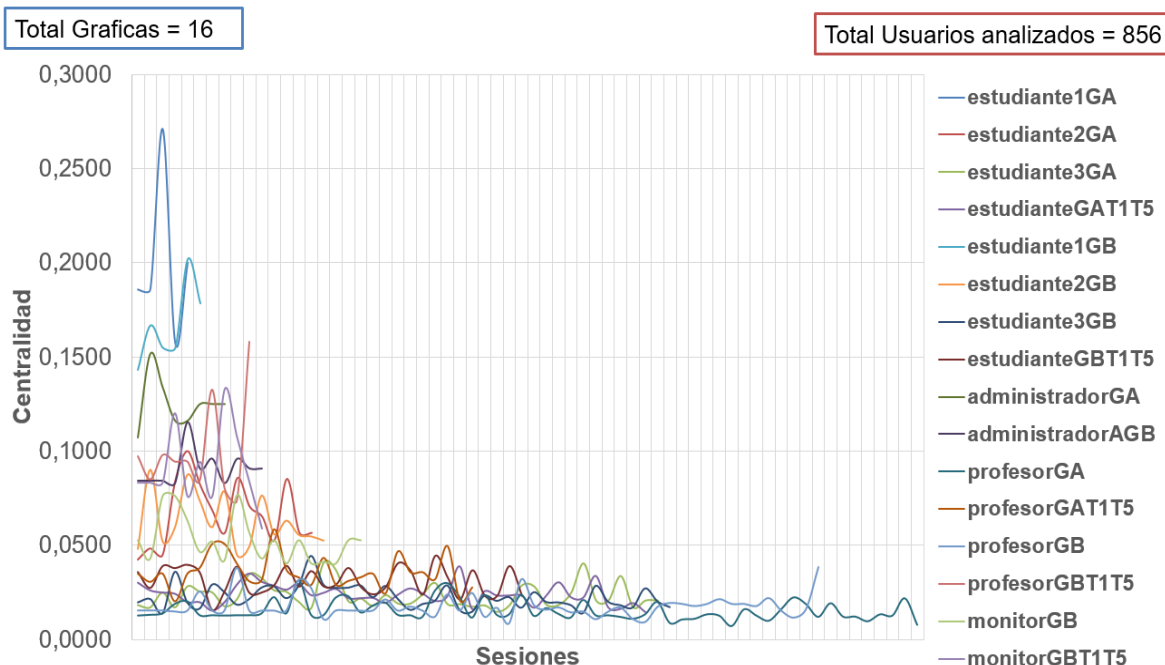


Figura 27. Graficas de centralidad de los patrones de navegación EF-ISF de los usuarios analizados. Fuente propia.

En la Figura 27, se observan las gráficas de todos usuarios analizados, las gráficas de los patrones se van agrupando por numero de decisiones; de 5 a 6, intervalo 0.1500 a 0.3000; de 7 a 14, intervalo 0.1000 a 0.1500; de 15 a 31, intervalo 0.0500 a 0.1000; de 32 en adelante el intervalo 0 a 0.0500.

En el presente trabajo de investigación se analizaron 856 usuarios, con lo cual se obtuvieron 856 patrones de navegación EF-ISF por sesión de mayor centralidad. De los cuales se analizaron dieciséis usuarios, estos se describen a continuación.

De los tres estudiantes del grupo A, el estudiante 1 y 2 ingresan al curso a revisar los simulacros de exámenes para obtener las repuestas correctas o incorrectas y posteriormente presentar el examen en línea, además, al realizar este comportamiento antes de presentar el examen en línea, sin cumplir las actividades

de estudio planteadas por el docente, se infiere una deshonestidad académica. El estudiante 1 por lo general ingresa al curso en horas de la tarde y el estudiante 2 comúnmente ingresa al curso en horas de la madrugada. Entre tanto, el estudiante 3 normalmente ingresa al curso en horas de la madrugada, realiza la mayoría de actividades programadas por el profesor en el curso e incluso el uso de materiales de aprendizaje extra para complementar el desarrollo del curso. El estudiante del grupo A de las unidades temáticas, usualmente ingresa al curso en horas de la madrugada.

De los tres estudiantes del grupo B, se puede decir que el estudiante 1 por lo general ingresa al curso en horas de la mañana, utiliza en mayor medida video, tiene una interacción con el curso activa en el inicio; pero luego empieza a descender; entre tanto, los estudiantes 2 y 3 tienen una participación activa en el curso. El estudiante 2, frecuentemente ingresa al curso en horas de la madrugada, revisa los simulacros de exámenes para obtener las respuestas correctas o incorrectas para luego presentar el examen, a pesar de que el estudiante tiene una interacción activa en el curso utilizando diferente material de aprendizaje, este comportamiento es similar al presentado por los estudiantes 1 y 2 del Grupo A, con lo cual estudiante 2 del Grupo B también presenta una deshonestidad académica; esto contrasta con el estudiante 3, comúnmente ingresa al curso en horas de la tarde, tiene una participación activa y similar en todo el curso utilizando diferente material de aprendizaje e incluso material adicional. El estudiante del grupo B de las unidades temáticas, por lo regular ingresa al curso en horas de la noche, tiene una interacción activa y similar con el curso.

Los estudiantes del Grupo A y B analizados, ingresan en diferentes horarios en el transcurso del día para realizar sus actividades académicas. El material de aprendizaje más utilizado son videos, videos complementarios, simuladores, foros, recursos web, simulacros de exámenes y exámenes, además, se identificó que algunos usuarios desarrollan un solo patrón de navegación y otros usuarios desarrollan varios patrones de navegación en la interacción con el curso.

Con base en la deshonestidad académica presentada por los estudiantes 1 y 2 del Grupo A y el estudiante 2 del Grupo B, se realiza la siguiente recomendación al docente para el diseño del curso: tener presente la forma como se está realizando el examen de simulacro en línea y los exámenes en línea, porque los estudiantes están presentando deshonestidades académicas. El docente debe replantear la forma en como está llevando a cabo este tipo de evaluaciones.

El administrador del grupo A y B, por lo general ingresa al curso a estar pendiente de las diferentes evaluaciones que presentan los estudiantes en línea. El ingreso

del administrador a la plataforma, por lo general es en horas de la madrugada en ambos grupos.

El profesor en el Grupo A y en las unidades temáticas de este grupo, publica o prepara el material de aprendizaje para el curso; entre tanto, en el Grupo B y en las unidades temáticas de este grupo, está pendiente de los diferentes exámenes en línea y los exámenes en línea extras, que presentan los estudiantes en el curso. En ambos Grupos el profesor ingresa por lo general en horas de la tarde empezando la noche y en las unidades temáticas ingresa frecuente mente en horas de la tarde.

El monitor en el Grupo B y en las unidades temáticas, ingresa al curso en horas de la tarde empezando la noche, está pendiente de los foros, para responder a los estudiantes alguna consulta o enviar algún mensaje a los estudiantes o a los demás usuarios, también ha estado pendiente de las evaluaciones presentadas por los estudiantes; el material de aprendizaje más utilizado es: simuladores, recursos Web, foros, talleres, página Web y examen en línea. En las unidades temáticas, en las últimas sesiones su interacción con el curso fue menor y va mermando de manera importante, esto coincide con la interpretación realizada en grafica del patrón en el curso del Grupo B, en las sesiones s15, s16 y s17, porque el monitor en esas fechas refleja una baja interacción con el curso, con lo que se puede concluir, el monitor en ese tiempo no estuvo tan pendiente de la plataforma o ingreso poco.

CAPITULO 6

En este capítulo se describen las conclusiones del presente trabajo de investigación, además de recomendar trabajos futuros.

6 CONCLUSIONES Y TRABAJO FUTURO

6.1 Conclusiones

En el presente trabajo de investigación, se identificó los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios (estudiantes, administradores, profesores y monitor), en cuatro conjuntos de datos de un MPOC, soportado en una instancia de la plataforma de Open edX, a través del algoritmo LexRank con Umbral. Los conjuntos de datos contienen algunas interacciones de los usuarios con el curso Astronomía Cotidiana, ofertado en la Universidad del Cauca en el primer periodo académico de 2017.

Se adaptó la propuesta, Descubrimiento de patrones de navegación en Open edX – una aproximación arquitectónica, con lo cual se propone, una aproximación arquitectónica para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios, en conjuntos de datos, que contienen la interacción de los usuarios con cursos MPOC de la plataforma Open edX.

Se realizó la caracterización de los patrones de navegación EF-ISF por sesión de los usuarios, a cuatro conjuntos de datos de un MPOC, que contiene algunas interacciones de los usuarios, del curso Astronomía Cotidiana 2017-I, con limpieza y adecuación de la información, el modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento.

Se adaptó el algoritmo LexRank con Umbral que pertenece a los métodos de grafos, de generación automática de resúmenes de múltiples documentos, para la identificación de las centralidades de los patrones de navegación EF-ISF por sesión de los usuarios, en cuatro conjuntos de datos, de un MPOC soportado en una instancia de la plataforma Open edX.

Con los resultados de la matriz estacionaria, se propuso la forma de construir la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del usuario, para ilustrar la interacción del usuario a lo largo del curso. En la gráfica se identifica el patrón de navegación EF-ISF por sesión de mayor centralidad del usuario. La grafica es continua y dependiendo de la interacción del usuario con el curso, puede ser: creciente, decreciente, constante, con presencia de máximos y mínimos.

Del análisis realizado a los usuarios seleccionados en los conjuntos de datos, del curso Astronomía Cotidiana, con la sesión a la que le corresponde, el patrón de

navegación EF-ISF por sesión de mayor centralidad del usuario y la gráfica de la centralidad del patrón de navegación EF-ISF por sesión del usuario, se identificó, todos los usuarios tienen una navegación aleatoria; algunos usuarios desarrollan un solo patrón de navegación y otros usuarios desarrollan varios patrones de navegación en la interacción con el curso. Los estudiantes, tienen diferentes horarios para estudiar, el material de aprendizaje que más utilizan es: videos, videos complementarios, simuladores, foros, recursos web, simulacros de exámenes en línea y exámenes en línea. Hay algunos estudiantes que revisan los simulacros de exámenes o los exámenes en línea para obtener las repuestas correctas o incorrectas y posteriormente presentar el examen, al realizar este comportamiento antes de presentar el examen en línea, sin cumplir las actividades de estudio planteadas por el docente, se infiere un comportamiento deshonesto en las actividades académicas desarrolladas por el estudiante.

Con base en la deshonestidad académica presentada por los estudiantes, se realiza una recomendación al docente para el diseño del curso: tener presente la forma como se está realizando el examen de simulacro en línea y los exámenes en línea, porque los estudiantes están presentando deshonestidades académicas. El docente debe replantear la forma en como está llevando a cabo este tipo de evaluaciones.

Por otra parte, el administrador, por lo general ingresa en horas de la madrugada a los cursos, está pendiente de los exámenes en línea de los estudiantes; el profesor, prepara el material del curso, está pendiente de los exámenes en línea y los exámenes en línea extras de los estudiantes, su hora de ingreso por lo regular es en horas de la tarde empezando la noche; el monitor, comúnmente ingresa al curso en horas de la tarde empezando la noche, está pendiente de los exámenes en línea y los foros para dar respuesta a los estudiantes o demás usuarios.

6.2 Trabajo Futuro

Implementar la aproximación arquitectónica para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios, con algoritmos de método de agrupamiento, luego realizar interpretaciones específicas con el algoritmo LexRank con Umbral en el usuario(s) analizar.

Implementar la aproximación arquitectónica para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios, con otros algoritmos de diferentes metodologías como: algebraico, meta heurístico, consenso y entre otros, luego comparar los resultados obtenidos con el LexRank con Umbral.

Evaluar la presente propuesta de investigación, con conjuntos de datos que contengan mayor información de interacción de los usuarios con MOOC, soportado en una instancia de la plataforma de Open edX.

Trabajar en patrones argumentativos para realizar el análisis cualitativo realizado a los patrones de navegación EF-ISF por sesión de los usuarios de mayor centralidad y la gráfica de la centralidad del patrón de navegación EF-ISF.

REFERENCIAS

- Alguliev, R. M., Aliguliyev, R. M., Hajirahimova, M. S., & Mehdiyev, C. A. (2011). MCMR: Maximum coverage and minimum redundant text summarization model. *Expert Systems with Applications*, 38(12), 14514-14522. doi: <http://dx.doi.org/10.1016/j.eswa.2011.05.033>
- Alguliyev, R. M., Aliguliyev, R. M., & Isazade, N. R. (2015). An unsupervised approach to generating generic summaries of documents. *Applied Soft Computing*, 34, 236-250.
- Anacona, F., Cobos, C., & Mendoza, M. (2015). *Algoritmo para generación automática de resúmenes extractivos genéricos de múltiples documentos basado en consensos*. (Monografía), Universidad del Cauca. (01)
- Anacona, F., Solarte, M., & González, G. (2018). Modelo de Espacio Vectorial con ponderación basada en frecuencia relativa de eventos de navegación en una instancia de Open edX para caracterización del estudiantado. *Segunda Conferencia Interanacional MOOC-MAKER Medellín 2018*, 2224, 87-95.
- Anacona, F., Solarte, M., & Ramírez, G. (2017). Descubrimiento de patrones de navegación en Open edX—una aproximación arquitectónica. *Ingeniería e Innovación*, 5(1).
- Anacona, F. A., Solarte, M., & González, G. R. (2018). LexRank con Umbral para identificar patrones de navegación EF-ISF en xMOOC de Open edX. *Memorias SINNEM18*, 1, 206.
- Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). *Engaging with massive online courses*. Paper presented at the Proceedings of the 23rd international conference on World wide web.
- Campbell, J., Gibbs, A. L., Najafi, H., & Severinski, C. (2014). A comparison of learner intent and behaviour in live and archived MOOCs. *The International Review of Research in Open and Distributed Learning*, 15(5).
- Central, C. (2018). Class Central. Retrieved 23-01-2019, 2020, from <https://www.classcentral.com/report/mooc-stats-2018/>
- Das, D., & Martins, A. F. T. (2007). A Survey on Automatic Text Summarization. 1-31.
- edX. (2016). edX Research Guide. Retrieved 29-12-2018, 2018, from https://edx.readthedocs.io/projects/devdata/en/stable/internal_data_formats/tracking_logs.html
- edX. (2017). About Open edX. Retrieved 10-12-2018, 2018, from <https://open.edx.org/about-open-edx>
- edX. (2018). Open edX Architecture. Retrieved 24-12-2018, 2018, from <https://edx.readthedocs.io/projects/edx-developer-guide/en/latest/architecture.html>
- Erkan, G., & Radev, D. R. (2004). LexRank: graph-based lexical centrality as salience in text summarization. *J. Artif. Int. Res.*, 22(1), 457-479.
- Fasihuddin, H., Skinner, G., & Athauda, R. (2015). Towards adaptive open learning environments: Evaluating the precision of identifying learning styles by tracking learners' behaviours. *Education and Information Technologies*, 1-19.
- Fox, A. (2013). From moocs to spocs. *Communications of the ACM*, 56(12), 38-40.

- Gillani, N., & Eynon, R. (2014). Communication patterns in massively open online courses. *The Internet and Higher Education*, 23, 18-26.
- Glance, D. G., Barrett, P. H. R., & Hugh, R. (2014). *Attrition patterns amongst participant groups in Massive Open Online Courses*. Paper presented at the ASCILITE Conference, Dunedin, New Zealand. Retrieved from <http://ascilite2014.otago.ac.nz/files/fullpapers/16-Glance.pdf>.
- Gong, Y., & Liu, X. (2001). *Generic text summarization using relevance measure and latent semantic analysis*. Paper presented at the Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, New Orleans, Louisiana, USA.
- Grossman, S. I. (2001). *Álgebra Lineal*: Editorial Offset, S.A. de C.V, Durazno No. 1 esq. Ejido, Col. Las Peritas, Tepepan Xochimilco, C.P. 16010 México D.F.
- Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC design principles: A pedagogical approach from the learner's perspective. *eLearning Papers*(33).
- Guo, P. J., & Reinecke, K. (2014). *Demographic differences in how students navigate through MOOCs*. Paper presented at the Proceedings of the first ACM conference on Learning@ scale conference.
- Guo, W. (2014). *From SPOC to MPOC--The Effective Practice of Peking University Online Teacher Training*. Paper presented at the Educational Innovation through Technology (EITT), 2014 International Conference of.
- Hernández, S. (2013). SCOPEO INFORME N°2: MOOC: Estado de la situación actual, posibilidades, retos y futuro.
- Jaramillo, D., & Mario, S. (2017). *Seguimiento automático de actividades de aprendizaje en cursos en línea abiertos y masivos*. (Monografía). (01)
- Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). *Understanding in-video dropouts and interaction peaks in online lecture videos*. Paper presented at the Proceedings of the first ACM conference on Learning @ scale conference, Atlanta, Georgia, USA.
- Kizilcec, R. F. (2013). *Collaborative Learning in Geographically Distributed and In-person Groups*. Paper presented at the Proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED).
- Kizilcec, R. F., Piech, C., & Schneider, E. (2013). *Deconstructing disengagement: analyzing learner subpopulations in massive open online courses*. Paper presented at the Proceedings of the third international conference on learning analytics and knowledge.
- Korhonen, A., & Multisilta, J. (2017). Learning analytics *New Ways to Teach and Learn in China and Finland: Crossing Boundaries with Technology* (pp. 301-310).
- Lin, F.-R., & Kao, C.-M. (2018). Mental effort detection using EEG data in E-learning contexts. *Computers & Education*, 122, 63-79.
- Liyanagunawardena, T., Williams, S., & Adams, A. (2013). The impact and reach of MOOCs: a developing countries' perspective. *eLearning Papers*(33).
- Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distributed Learning*, 14(3), 202-227.
- Mackness, J., Mak, S., & Williams, R. (2010). The ideals and reality of participating in a MOOC.

- Maldonado, J. J., Palta, R., Vázquez, J., Bermeo, J. L., Pérez-Sanagustín, M., & Muñoz-Gama, J. (2016). *Exploring differences in how learners navigate in MOOCs based on self-regulated learning and learning styles: A process mining approach*. Paper presented at the 2016 XLII Latin American Computing Conference (CLEI).
- Maldonado, J. J., Pérez-Sanagustín, M., Bermeo, J. L., Muñoz, L., Pacheco, G., & Espinoza, I. (2017). *Flipping the classroom with MOOCs. A pilot study exploring differences between self-regulated learners*. Paper presented at the 12th Latin American Conference on Learning Objects and Technologies, LACLO 2017.
- Manning, C. D., Raghavan, P., & Schütze, H. (2009). *An introduction to information retrieval*. Cambridge: Cambridge University Press.
- Martin, F. G. (2012). Will massive open online courses change how we teach? *Communications of the ACM*, 55(8), 26-28.
- McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). The MOOC model for digital practice.
- Mendoza, M., & Leon, E. (2015). Una revisión de la generación automática de resúmenes extractivos. *UIS Ingenierías*, 12(1), 7-27.
- Merino, Ó. (2012). Tabla de conversión de símbolos UTF-8 a ISO-8859-1 (ANSI). Retrieved 06-02, 2019, from <http://www.merinomartinez.com/2012/06/tabla-de-conversion-de-simbolos-utf-8-a-iso-8859-1-ansi/>
- Montiel Soto, R., García-Hernández, R. A., Ledeneva, Y., & Cruz Reyes, R. (2009). Comparación de Tres Modelos de Texto para la Generación Automática de Resúmenes. 303-311.
- Palta, R., & Vázquez, J. (2016). *Descubrimiento de patrones de interacción en cursos MOOC en entornos ONLINE: Un enfoque utilizando Minería de Procesos. Caso de Estudio: "Curso de la metodología DICREVOA en Open edX"*. Universidad de Cuenca.
- Rizvi, S., Rienties, B., & Rogaten, J. (2018). *Temporal dynamics of MOOC learning trajectories*. Paper presented at the Proceedings of the First International Conference on Data Science, E-learning and Information Systems.
- Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Commun. ACM*, 18(11), 613-620. doi: 10.1145/361219.361220
- Siemens, G. (2004). A learning Theory for the Digital Age. *Elearnspace Everything Learning*, 1-8.
- Singhal, A. (2001). Modern Information Retrieval: A Brief Overview. *IEEE Data Eng. Bull.*, 24(4), 35-43.
- Solarte, M., Ramírez, G. A., & Jaramillo, D. A. (2017). Hábitos de ingreso y resultados en las evaluaciones en cursos en línea masivos con reconocimiento académico. *Ingeniería e Innovación*, 5(1).
- Sonwalkar, N. (2013). *The first adaptive MOOC: A case study on pedagogy framework and scalable cloud Architecture—Part I*. Paper presented at the MOOCs Forum.
- Steinberger, J., & Ježek, K. (2009). Text Summarization: An Old Challenge and New Approaches. In A. Abraham, A.-E. Hassanien, A. Leon F. de Carvalho & V. Snášel (Eds.), *Foundations of Computational, Intelligence Volume 6* (Vol. 206, pp. 127-149): Springer Berlin Heidelberg.

- Tang, H., Xing, W., & Pei, B. (2018). Exploring the temporal dimension of forum participation in MOOCs. *Distance Education*, 39(3), 353-372.
- Wang, Y., Law, N., Hemberg, E., & O'Reilly, U.-M. (2019). *Using Detailed Access Trajectories for Learning Behavior Analysis*. Paper presented at the Proceedings of the 9th International Conference on Learning Analytics & Knowledge.
- Xu, Z., Goldwasser, D., Bederson, B. B., & Lin, J. (2014). *Visual analytics of MOOCs at maryland*. Paper presented at the Proceedings of the first ACM conference on Learning@ scale conference.
- Yadav, R. K. (2014). *Understanding Logs in edX for Monitoring Student Progress*. (Doctoral Thesis), Indian Institute of Technology, Bombay.
- Yang, D., Wen, M., Kumar, A., Xing, E. P., & Rose, C. P. (2014). Towards an integration of text and graph clustering methods as a lens for studying social interaction in MOOCs. *The International Review of Research in Open and Distributed Learning*, 15(5).
- Zhang, X., & Lin, H. (2018) Modeling and interpreting user navigation patterns in MOOCs. *Vol. 464. Lecture Notes in Electrical Engineering* (pp. 403-413).

Contribuciones a la identificación de patrones de navegación de estudiantes en
cursos de Open edX



FABIÁN ANDRÉS ANACONA ANACONA

Anexos

Director:

Ph.D. Mario Fernando Solarte Sarasty

Universidad del Cauca
Facultad de Ingeniería Electrónica y Telecomunicaciones
Departamento de Telemática
Maestría en Ingeniería Telemática
Línea de investigación: Aplicaciones y Servicios sobre Internet
Popayán, abril de 2020

FABIÁN ANDRÉS ANACONA ANACONA

Contribuciones a la identificación de patrones de navegación de estudiantes en
cursos de Open edX

Tesis presentada a la Facultad de Ingeniería
Electrónica y Telecomunicaciones de la
Universidad del Cauca, para la obtención del
Título de

Magíster en:
Ingeniería Telemática

Director:
Ph.D. Mario Fernando Solarte Sarasty

Popayán
2020

TABLA DE CONTENIDO

Lista de Tablas.....	iii
Lista de Figuras.....	iv
ANEXO A – Artículo	1
1 DESCUBRIMIENTO DE PATRONES DE NAVEGACIÓN EN OPEN EDX – UNA APROXIMACIÓN ARQUITECTÓNICA.....	1
2 INTRODUCCIÓN.....	2
3 APROXIMACIÓN ARQUITECTÓNICA.....	3
4 TRACKING.LOG.....	4
5 PRE-PROCESAMIENTO.....	5
5.1 Archivo de texto plano.....	6
5.2 Representación de Estudiantes.....	6
6 ALGORITMO LEXRANK CON UMBRAL.....	8
7 CONCLUSIONES Y TRABAJO FUTURO	10
8 AGRADECIMIENTOS.....	10
ANEXO B - Artículo.....	11
9 LEXRANK WITH THRESHOLD TO IDENTIFY EF-ISF NAVIGATION PATTERNS IN XMOOC OF OPEN EDX.....	11
10 INTRODUCTION.....	12
11 TRACKING.LOG OF THE OPEN EDX PLATFORM	12
12 STUDENT REPRESENTATION.....	13
13 LEXRANK WITH THRESHOLD ALGORITHM	13
14 IDENTIFICATION OF EF-ISF NAVIGATION PATTERNS WITH THE LEXRANK WITH THRESHOLD ALGORITHM	17
15 CONCLUSIONS AND FUTURE WORK.....	19
ACKNOWLEDGEMENTS.....	20
ANEXO C - Artículo.....	21
16 MODELO DE ESPACIO VECTORIAL CON PONDERACIÓN BASADA EN FRECUENCIA RELATIVA DE EVENTOS DE NAVEGACIÓN EN UNA INSTANCIA DE OPEN EDX PARA CARACTERIZACIÓN DEL ESTUDIANTADO	21
17 INTRODUCCIÓN.....	21

18	TRACKING.LOG DE LA PLATAFORMA OPEN EDX	23
19	CARACTERIZACIÓN DE LOS PATRONES DE NAVEGACIÓN CON EL MODELO DE ESPACIO VECTORIAL Y EF-ISF	24
20	ALGORITMOS PARA EL DESCUBRIMIENTO DE PATRONES DE NAVEGACIÓN	27
21	CONCLUSIONES Y TRABAJO FUTURO	28
	AGRADECIMIENTOS	29
	ANEXO D – Metodologías y Diagrama de Clases.....	30
22	METODOLOGÍAS USADAS EN EL PROYECTO DE INVESTIGACIÓN.....	30
23	DIAGRAMA DE CLASES	31
	REFERENCIAS.....	33

Lista de Tablas

Tabla 1. Eventos de navegación en un MOOC. Tomado de (Palta & Vásquez, 2016).	5
Tabla 2. Campos de identificación de pestañas dentro de un MOOC. Tomado de (Palta & Vásquez, 2016).....	5
Tabla 3. Parts of the tracking.log file of two events of a student.....	17
Tabla 4. log events.	18
Tabla 5. EF-ISF navigation patterns of the anonymous student.	19
Tabla 6. Descripción de campos. Tomado de (Palta & Vásquez, 2016).	24
Tabla 7. Campos del evento context. Tomado de (Palta & Vásquez, 2016).	24
Tabla 8. Sub campos adicionales del evento context. Tomado de (Palta & Vásquez, 2016).....	24
Tabla 9. Descripción de Clases. Fuente Propia.	31

Lista de Figuras

Figura 1. Aproximación arquitectónica para el descubrimiento de patrones de navegación. Fuente Propia.....	4
Figura 2. Estructura general del archivo traking.log de Open edX.	4
Figura 3. Transformación del archivo Tracking.log en el modelo de espacio vectorial. Fuente Propia.....	6
Figura 4. LexRank con Umbral para el cálculo de patrones de navegación.....	8
Figura 5. General structure of the Open edX traking.log file. Source: taken from (Fabián Anacona et al., 2018).	13
Figura 6. Estructura general del archivo traking.log de Open edX. Tomado de (Fabián Anacona et al., 2017).	23
Figura 7. Diagrama de Clases. Fuente Propia.	31

ANEXO A – Artículo

En este apartado se describe el artículo titulado “Descubrimiento de patrones de navegación en Open edX – una aproximación arquitectónica”, publicado en la Revista de Ingeniería e Innovación de la Universidad de Córdoba – Colombia, revista nacional indexada por COLCIENCIAS categoría C en la fecha de publicación, el documento también puede ser consultado en la siguiente dirección: <https://revistas.unicordoba.edu.co/index.php/rii/article/view/1103>.

1 DESCUBRIMIENTO DE PATRONES DE NAVEGACIÓN EN OPEN EDX – UNA APROXIMACIÓN ARQUITECTÓNICA

DISCOVERY OF NAVIGATION PATTERNS IN OPEN EDX – AN ARCHITECTURAL APPROACH

Fabián A. Anacona¹, Mario F. Solarte², Gustavo A. Ramírez³

Resumen: Los MOOC (Massive Open Online Courses) se han convertido en una estrategia educativa disruptiva en la última década, caracterizada por dar soporte a procesos formativos a gran escala y de forma gratuita; a pesar de su éxito igualmente tienen debilidades como la rigidez pedagógica pues generalmente las actividades de aprendizaje son las mismas sin importar el perfil y número de los estudiantes. Desde hace algunos años se aplican técnicas de adaptación; pero sin un referente claro sobre cómo realizarla. Este trabajo presenta una aproximación a una arquitectura para descubrir patrones de navegación en los estudiantes en MOOC soportados en la plataforma Open edX, a través del modelo de espacio vectorial y el algoritmo LexRank con Umbral.

Palabras Claves: MOOC, SPOC, MPOC, algoritmo, patrones de navegación, Open edX.

Abstract: Massive Open Online Courses (MOOCs) have become a disruptive educational strategy in the last decade, characterized by supporting large-scale and free training processes; in spite of their success they also have debilities such as pedagogical rigidity as generally the learning activities are the same import of the profile and number of the students. Adaptive techniques have been applied for some years; but

¹ Ingeniero de Sistemas. Estudiante de Maestría en Ingeniería Telemática. Universidad del Cauca, Popayán, Colombia

² Magíster en Ingeniería. Profesor. Universidad del Cauca. Calle 5 # 4-70. Teléfono (+57) 2-809800, extensión 2175. Correo electrónico: msolarte@unicauca.edu.co. Popayán, Colombia.

³ Doctor en Ingeniería Telemática. Profesor. Universidad del Cauca, Popayán, Colombia

without a clear reference on how to do it. This work presents an approach to an architecture to discover navigation patterns in students in MOOC supported on the Open edX platform, through the vector space model and the algorithm LexRank with Threshold.

Key words: MOOC, SPOC, MPOC, algorithm, navigation patterns, Open edX.

2 INTRODUCCIÓN

Los MOOC (Massive Open Online Courses) son una propuesta para universalizar la educación gratuita y de calidad, están basados en el Conectivismo (Siemmens, 2004). Su principal objetivo es brindar alternativas educativas soportadas en tecnologías de Internet de acceso abierto (Guàrdia, Maina, & Sangrà, 2013; Hernández, 2013; T. Liyanagunawardena, Williams, & Adams, 2013).

Uno de los primeros MOOC ofrecidos fue el de "Conectivismo y conocimiento colectivo" orientado en 2008 por Stephen Downes de la Universidad de Manitoba de Canadá (Mackness, Mak, & Williams, 2010); sin embargo, el crecimiento del movimiento MOOC se da con el curso de "Introducción a la Inteligencia Artificial" orientado por Sebastián Thrun, profesor de la Universidad de Stanford de EE.UU y Peter Norvig, Investigador de Google, en el cual se matricularon aproximadamente 160.000 estudiantes (Martin, 2012). Los MOOC adquieren fama mundial en el año 2012 de la mano de diversas plataformas como: Udacity, Coursera, edX, entre otras; en las que participan las principales universidades del mundo (Hernández, 2013). Por su parte, edX es una organización de aprendizaje en línea, fundada por la Universidad de Harvard y el Instituto Tecnológico de Massachusetts (MIT) en el año 2012, la cual cuenta con más de 90 socios a nivel mundial, entre entidades sin ánimo de lucro y las principales instituciones que encabezan el ranking de universidades del mundo QS (edX, 2013, 2016). Open edX es la versión de software libre del sistema de gestión de aprendizaje utilizado en edX.

En Colombia, algunas universidades han realizado MOOC en diferentes plataformas (Semana, 2014); entre tanto, la Universidad del Cauca usa Open edX, con nombre de instancia Selene, para ofrecer cursos en línea en ambientes de masividad.

Los MOOC se clasifican de manera general en (Guàrdia et al., 2013; Hernández, 2013): cMOOC, cuyo objetivo es la generación de conocimiento de forma colaborativa basándose en la cantidad de aportes generados por sus participantes y los xMOOC, basados en la presentación de contenidos a través de videos cortos, documentos de apoyo, foros de discusión y evaluaciones.

En el año 2013, aparece el término SPOC (Small Private Online Courses) que son cursos en línea soportados en la plataforma MOOC; pero con la diferencia de no ser abiertos a todo el público y con una cantidad menor de estudiantes (Fox, 2013). En el año 2014, aparece el término MPOC (Massive Private Online Courses) cursos en línea igualmente soportados en plataforma MOOC, con la posibilidad de ser masivos, pero sin estar abiertos a todo el público y con una cantidad menor de estudiantes (W. Guo, 2014). Tanto SPOC como MPOC son estrategias encaminadas al reconocimiento académico de las habilidades y conocimientos adquiridos a través de las estrategias de los MOOC.

Los MOOC presentan características similares como: por lo general son de corta duración (T. R. Liyanagunawardena, Adams, & Williams, 2013), atraen a miles de estudiantes de diferentes lugares del mundo (Kizilcec, 2013), algunos emiten un certificado de aprobación, el cual puede tener un valor monetario (Kizilcec, 2013; McAuley, Stewart, Siemens, & Cormier, 2010).

De igual forma, los MOOC presentan dificultades como: deficiencia pedagógica (Sonwalkar, 2013), bajas tasas de retención (Matías González & Pérez Avila, 2014) y carencia de atención personalizada (Zapata-Ros, 2015); por estas razones Sonwalkar propone el concepto de aMOOC (adaptive MOOC), el cual busca desarrollar un proceso de aprendizaje personalizado permitiendo abrir nuevos campos de investigación, por ejemplo: la identificación de patrones de comportamiento de los estudiantes y la personalización en el desarrollo de un MOOC.

Una alternativa para obtener los patrones de comportamiento, es la identificación de las interacciones de los estudiantes con los servicios de la plataforma MOOC, almacenados en el archivo tracking.log, por esta razón, en este artículo se presenta una forma de encontrar los patrones de navegación en el archivo tracking.log, de la plataforma Open edX a través de un algoritmo.

Para ello, en la sección 2 se describe la aproximación arquitectónica para el descubrimiento de patrones de navegación, en la sección 3 la estructura del archivo de registro de eventos, en la sección 4 la estrategia del pre-procesamiento requerido, en la sección 5 el algoritmo LexRank con Umbral y en la sección 6 las conclusiones y trabajo futuro.

3 APROXIMACIÓN ARQUITECTÓNICA

Partiendo de una arquitectura propuesta para identificar perfiles de aprendizaje de los estudiantes, la cual consta de los siguientes componentes: Recopilación de

datos, Pre-procesamiento, Extracción de características, Clasificación, Perfil del estudiante, Adaptación (recomendación) (Hmedna, El Mezouary, Baz, & Mammass, 2017). Basado en lo anterior, en este artículo se propone la aproximación arquitectónica mediante la cual es posible descubrir patrones de navegación en Open edX. Consta de los siguientes módulos: archivo tracking.log, pre procesamiento (archivo de texto plano y representación de estudiantes) y algoritmo LexRank con Umbral; estos se muestran en la Figura 1. Cada uno de los elementos se describe en las siguientes secciones.

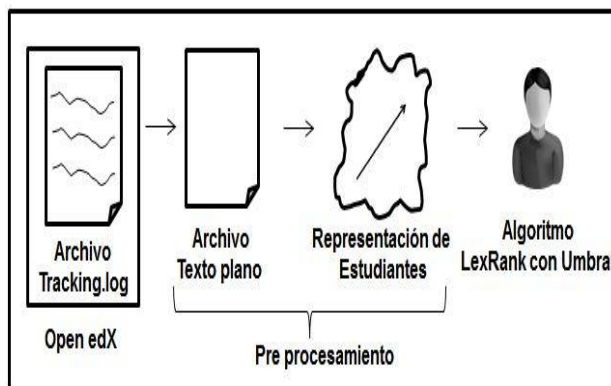


Figura 1. Aproximación arquitectónica para el descubrimiento de patrones de navegación. Fuente Propia.

4 TRACKING.LOG

El archivo Tracking.log se encarga de almacenar todos los eventos realizados sobre la plataforma Open edX en un formato JSON. Los eventos que se almacenan son generados por el servidor, navegador o dispositivo móvil, estos se conocen como emisores. Los eventos se refieren a la información de interacción que se realiza sobre el material del curso y el tablero de instrumentos en el sistema de gestión de aprendizaje de la plataforma. Para tener acceso a la información se sigue a la ruta /edx/var/log/tracking/, en este lugar se encuentra un conjunto de archivos comprimidos que contienen toda la interacción que los participantes realizan sobre la plataforma. Los eventos generalmente representan una acción de los emisores sobre la plataforma, a pesar de que dichos eventos pueden representar acciones distintas tienen un conjunto de campos comunes, cuya estructura general se presenta a continuación (Palta & Vásquez, 2016):

```

{ "username": "", "event_source": "", "name": "", "accept_language": "", "time": "", "agent":
"", "page": "", "host": "", "session": "", "referer": "", "context": { "user_id": "", "org_id": "",
"course_id": "", "path": ""}, "ip": "", "event": "", "event_type": "" }
  
```

Figura 2. Estructura general del archivo tracking.log de Open edX.

Los eventos de los usuarios que permiten conocer cómo fue la navegación de los participantes sobre un curso MOOC, se detallan en las Tablas 1 y 2 (Palta & Vásquez, 2016).

Dichos eventos están asociados a la navegación de los estudiantes por los distintos servicios de Open edX que también depende de la metodología escogida por el diseñador y orientador de un curso: Ingresos a la plataforma, acceso a videos y otros documentos, participación en Foros de Discusión (lectura, creación de hilos, respuestas), desarrollo de exámenes y auto evaluaciones, entre otros.

Tabla 1. Eventos de navegación en un MOOC. Tomado de (Palta & Vásquez, 2016).

EVENTO	DESCRIPCIÓN
page_close	Este evento se registra desde una acción JavaScript.
seq_goto	Se emite cuando un usuario salta entre las unidades en una secuencia.
seq_next	Se emite cuando un usuario navega a la siguiente unidad en una secuencia.
seq_prev	Se emite cuando un usuario navega a la unidad anterior en una secuencia

Los eventos seq_goto, seq_next y seq_prev, tienen algunos campos que se describen a continuación.

Tabla 2. Campos de identificación de pestañas dentro de un MOOC. Tomado de (Palta & Vásquez, 2016).

CAMPO	TIPO	DESCRIPCIÓN
id	Number	Identifica el id edX de la secuencia.
new	Number	Para seq_goto, el índice de la unidad que se saltó al valor especificado. Para seq_next y seq_prev, el índice de la unidad navega al valor especificado.
old	Number	Identifica la pestaña origen del evento.

5 PRE-PROCESAMIENTO

El Pre-procesamiento es el primer paso que se realiza en los datos sin procesar, esta etapa tiene como objetivo (Hmedna et al., 2017):

- Limpiar los datos recolados de información que no aporten.
- Transformar los datos en un formato limpio que pueda ser utilizado por la aplicación.
- Preparar los datos para el análisis por un algoritmo(s), que pertenezca a método de Grafo, Algebraico y Meta Heurístico, entre otros (Fabián Anacona, Cobos, & Mendoza, 2015).

En las siguientes sub secciones de describirá cómo transformar el archivo tracking.log a un modelo de espacio vectorial, con lo cual se prepara los datos del archivo para que sean procesados por un algoritmo, como se ilustra en la Figura 3.

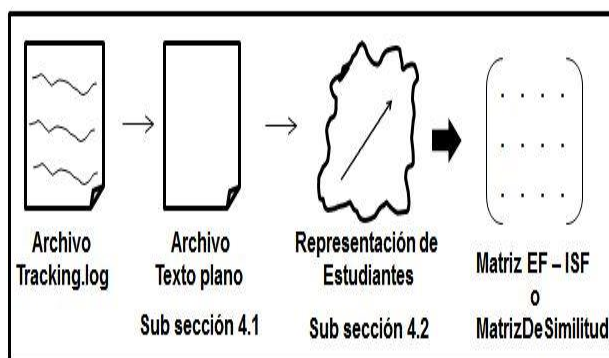


Figura 3. Transformación del archivo Tracking.log en el modelo de espacio vectorial.
Fuente Propia.

5.1 Archivo de texto plano

El archivo Tracking.log se debe leer y procesar para obtener un archivo de texto plano, en este proceso de limpieza y adecuación de los datos se debe eliminar la estructura de formato JSON, manteniendo los eventos de navegación de los estudiantes.

5.2 Representación de Estudiantes

En el documento se propone trabajar con el Modelo de Espacio Vectorial (Salton, Wong, & Yang, 1975), porque permite encontrar la relación de los eventos de navegación de los estudiantes en un curso MOOC.

Modelo de Espacio Vectorial

Un vector estudiante \vec{s}_j identifica en qué grado se satisface cada una de las m características del conjunto de estudiantes. Las características (componentes) del vector son un valor numérico. El concepto de característica puede definir la ocurrencia de determinados eventos (Manning, Raghavan, & Schütze, 2009; Zazo Rodríguez, Figuerola Paniagua, Berrocal, & Gómez Díaz, 2002).

Se considera un espacio vectorial de un estudiante \vec{s}_j , el cual es representado por uno o más eventos e_i ; los eventos pueden ser ponderados (w_{ij}) o no ponderados de acuerdo a su importancia, esto se representa por un m -vector de la siguiente forma (Salton et al., 1975):

$$\vec{s}_{mj} = (w_{1j}, w_{2j}, \dots, w_{ij}, \dots, w_{mj}) \quad (1)$$

Donde w_{ij} representa el peso del evento i -ésimo del estudiante j .

Cada evento e_i se convierte en una dimensión independiente en el espacio dimensional del vector, la mayoría de los vectores operan en los cuadrantes positivos, dado que a ningún evento se le asigna un valor negativo (Singhal, 2001).

Ponderación basada en la frecuencia relativa de un evento

Para la ponderación de los eventos, se usa la Frecuencia Relativa de un Evento (EF-ISF), la cual combina las definiciones de frecuencia del evento ($EF = f_{ij}$) y frecuencia inversa del estudiante ($ISF = \log \frac{N}{n_j}$) para asignar un peso a cada evento del estudiante \vec{s}_j . Se otorga mayor relevancia a los eventos menos frecuentes en el conjunto de estudiantes; pero a su vez más frecuentes en cada estudiante, como se observa en la siguiente ecuación donde w_{ij} representa el peso del evento i –ésimo del estudiante \vec{s}_j (Manning et al., 2009):

$$w_{ij} = (f_{ij}) \left(\log \frac{N}{n_j} \right) \quad (2)$$

Medidas de Similitud

Para el cálculo de similitud entre los estudiantes, estos se deben representar como vectores en el modelo de espacio vectorial (Alguliev, Aliguliyev, Hajirahimova, & Mehdiyev, 2011). Teniendo los estudiantes como vectores es posible calcular su semejanza, con medidas de similitud, como la de coseno u otras (Salton et al., 1975). Con similitud de cosenos sería: sean \vec{s}_i, \vec{s}_j dos m -vectores diferentes del vector cero, donde el peso de los eventos se calcula con EF-ISF (Ver ecuación (2)). Entonces el ángulo φ o la similitud entre \vec{s}_i y \vec{s}_j se define en el intervalo $[0, 1]$ de acuerdo a la ecuación (Alguliev et al., 2011; Singhal, 2001):

$$\text{simcos}(\vec{s}_i, \vec{s}_j) = \frac{\sum_{k=1}^m w_{ki} w_{kj}}{\sqrt{\sum_{k=1}^m w_{ki}^2 \sum_{k=1}^m w_{kj}^2}} \quad (3)$$

Representación del Traking.log por medio de Matrices

La representación en el espacio multidimensional del conjunto de vectores de estudiantes y la similitud de los mismos, se puede hacer con la Matriz de Eventos por Estudiantes y la Matriz de similitud de Cosenos.

- **Matriz de Eventos por Estudiantes**

La *Matriz EF – ISF* _{$m \times n$} es una matriz dispersa de pesos dispuestos en m eventos (filas) y n estudiantes (columnas), el elemento ij de la Matriz es denotado por w_{ij} , que corresponde al peso del evento i en el estudiante j , este

peso es calculado de acuerdo a la ecuación (2) (Gong & Liu, 2001; Grossman, 2001).

- **Matriz de Similitud de Cosenos**

La *MatrizDeSimilitud* _{$n \times n$} de cosenos es una matriz cuadrada, donde n es el número de estudiantes, el elemento ij de la matriz denotado por a_{ij} , es la similitud entre la estudiante \vec{s}_i y \vec{s}_j que aparece en la fila i y la columna j , los elementos de diagonal principal son iguales a 1, ya que es el cálculo de cada estudiante con el mismo, \vec{s}_i con \vec{s}_i cuando $i = j$ (ecuacion (3)) (Grossman, 2001).

6 ALGORITMO LEXRANK CON UMBRAL

Al representar los datos del archivo Tracking.log en el modelo de espacio vectorial, estos pueden ser procesados por un algoritmo con algún propósito particular. Se ha considerado el algoritmo LexRank con Umbral para intentar descubrir patrones de comportamiento en la navegación de los estudiantes a través de la arquitectura de información de un curso en Open edX. Dicho algoritmo pertenece a los Métodos de Grafos, siendo de los más referenciados en generación automática de resúmenes de múltiples documentos; además, al realizar pruebas con los conjuntos de documentos DUC2005⁴, DUC2006 y evaluación automática con las medidas ROUGE-1⁵, ROUGE-2 y ROUGE-SU4, con los siguientes algoritmos: Máxima Cobertura Mínima Redundancia (MCMR), método meta heurístico; Análisis Semítico Latente (LSA), método algebraico y LexRank con Umbral, método de grafo; se evidenció que el LexRank con un Umbral de 0.1 obtuvo mejores resultados (Anaconda Anaconda, Cobos Lozada, & Mendoza Becerra, 2015). La Figura 4 muestra el proceso para encontrar el patrón de navegación a través del LexRank con Umbral.

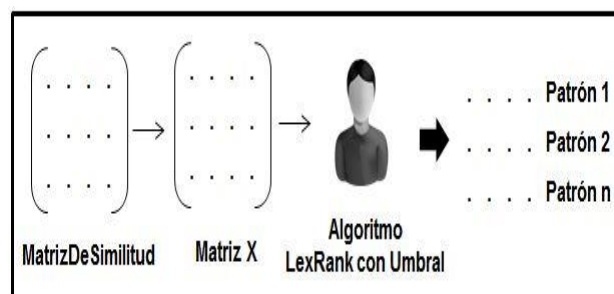


Figura 4. LexRank con Umbral para el cálculo de patrones de navegación.

Erkan y Radev plantean que un grupo de documentos puede verse como una red de oraciones relacionadas; en la presente investigación se plantea que un

⁴ La Conferencia de Entendimiento del Documento (Document Understanding Conference), ofrece un conjunto de documentos para realizar pruebas.

⁵ ROUGE (Recall-Oriented Understudy for Gisting Evaluation), es un estándar para la evaluación automática de resúmenes de textos

estudiante en diferente tiempo con sus respectivos eventos (ver ecuación (1)), puede verse como una red relacionada. Ahora bien, si un estudiante es muy similar a las otras interacciones del mismo estudiante, se puede considerar como el central o sobresaliente (Erkan & Radev, 2004).

La similitud de eventos de los estudiantes es representada en una matriz de adyacencia (matriz de similitud de cosenos), ésta se transforma en una matriz X de transición de una cadena de Markov que sea irreducible y aperiódica; luego se encuentra la matriz estacionaria que representa la solución, es decir, los patrones de navegación.

A continuación se presenta el pseudo-código del LexRank con Umbral⁶ para la identificación de patrones de navegación, el Algoritmo 1 resume cómo calcular las puntuaciones LexRank con Umbral para un conjunto de estudiantes.

```

Entrada: Arreglo S de n estudiantes, umbral t, valor de amortiguamiento dampingFactor
Salida: Arreglo L con las puntuaciones definidos por LexRank para cada estudiante
Arreglo MatrizDeSimilitud[n][n];
Arreglo L[n];
01 Para i=1 hasta n haga
02     suma=0
03     Para j=1 hasta n haga
04         MatrizDeSimilitud[i][j] = idf-modificado-coseno(S[i],S[j]);
05         Si MatrizDeSimilitud[i][j] > t haga
06             MatrizDeSimilitud[i][j] = 1;
07             suma++;
08         Si No
09             MatrizDeSimilitud[i][j] = 0;
10         Fin Si
11     Fin Para
15 Fin Para
16 Para i=1 hasta n haga
17     Para j=1 hasta n haga
18         MatrizDeSimilitud[i][j] = MatrizDeSimilitud[i][j] / suma;
19     Fin Para
20 Fin Para
21 Para i=1 hasta n haga
22     Para j=1 hasta n haga
23         MatrizDeSimilitud[i][j] = (dampingFactor/n) + (1- dampingFactor)* MatrizDeSimilitud[i][j];
24     Fin Para
25 Fin Para
26 L = MetodoDePotencia(MatrizDeSimilitud, n, ε ); // Ver Algoritmo 2Método de Potencia.
27 retorne L;

```

Algoritmo 1: Cálculo de puntuación de LexRank con Umbral. Adaptado de (Erkan & Radev, 2004).

El Método de Potencia describe como se calcula la matriz estacionaria de una Cadena de Markov, mostrado en el Algoritmo 2.

⁶ El umbral, clasifica los estudiantes, de los cuales se obtiene el sobresaliente.

Entrada: Una matriz M estocástica, irreducible y aperiódica

Entrada: Tolerancia de error ϵ

Salida: Vector propio p

01 $p_o = \frac{1}{N} \mathbf{1}$

02 $t = 0$

03 repita

04 $t = t + 1$

05 $p_t = M^T p_{t-1}$

06 $\delta = \|p_t - p_{t-1}\|$

07 hasta que $\delta < \epsilon$;

08 retorne p_t

Algoritmo 2: Método de Potencia. Adaptado de (Erkan & Radev, 2004).

7 CONCLUSIONES Y TRABAJO FUTURO

Ante la oportunidad de usar la gran cantidad de datos producidos y almacenados en las plataformas de cursos en línea masivos, es posible la aplicación del Modelo de Espacio Vectorial y Algoritmo(s), en entornos educativos encaminados a comprender los hábitos y comportamiento de los estudiantes para mejorar su proceso formativo.

Al representar los eventos de navegación del archivo Traking.log de Open edX, en el modelo de espacio vectorial y posteriormente procesarlos con el algoritmo LexRank con Umbral, permite descubrir patrones de navegación; sin embargo la implementación no se ha desarrollado.

Como trabajo a futuro se plantea la implementación de la propuesta y su aplicación a los cursos en línea con reconocimiento académico ofrecidos en la Universidad del Cauca, Popayán, Colombia. Posteriormente se puede articular el descubrimiento de patrones de navegación en Open edX con el de los perfiles de los estudiantes (demográficos, estilos de aprendizaje, entre otros) para dar soporte a procesos de adaptación y personalización efectivas del proceso educativo dentro de una plataforma MOOC.

8 AGRADECIMIENTOS

Los autores agradecen el apoyo recibido por el proyecto MOOC-Maker (561533-EPP-1-2015-1-ESEPPKA2-CBHE-JP) financiado por la Comisión Europea a través del Programa Erasmus+ para la realización y divulgación de los resultados expuestos en el presente artículo.

ANEXO B - Artículo

Ponencia en el Seminario Internacional de Innovaciones Educativas y MOOC SINNEM18, realizada en la Universidad del Cauca, Popayán, Colombia, el día viernes 5 de octubre de 2018; el título de la ponencia fue “LexRank con Umbral para identificar patrones de navegación EF-ISF en xMOOC de Open edX”. El artículo se encuentra en las actas del evento en la siguiente dirección: <https://www.unicauca.edu.co/moocmaker/sinnem/memorias.html>, el artículo está en proceso de evaluación en la revista de Ingenierías de la Universidad de Medellín, Colombia.

9 LEXRANK WITH THRESHOLD TO IDENTIFY EF-ISF NAVIGATION PATTERNS IN XMOOC OF OPEN EDX

Abstract: The use of the Open edX platform to offer xMOOC courses by different Universities in the world, has led to a growth in the participation of students in the courses, thus generating a set of navigation patterns when interacting with the xMOOC which are recorded in the tracking.log file. So far there is no study that identifies EF-ISF navigation patterns of the set of patterns, for this reason in this document propose the use of the LexRank with Threshold algorithm for the identification of the EF-ISF navigation patterns.

Keywords: LexRank with Threshold, navigation, EF-ISF navigation pattern, tracking.log, xMOOC.

LexRank con Umbral para identificar patrones de navegación EF-ISF en xMOOC de Open edX

Resumen: El uso de la plataforma Open edX para ofrecer cursos xMOOC por diferentes Universidades en el mundo, ha llevado a un crecimiento en la participación de los estudiantes en los cursos, generando de esta forma un conjunto de patrones de navegación al interactuar con los xMOOC los cuales se registran en el archivo tracking.log. Hasta el momento no hay un estudio que identifique patrones de navegación EF-ISF del conjunto de patrones, por esta razón en este documento se propone el uso del algoritmo LexRank con Umbral para la identificación de los patrones de navegación EF-ISF.

Palabras Claves: LexRank con Umbral, navegación, patrón de navegación EF-ISF, tracking.log, xMOOC.

10 INTRODUCTION

The Massive Open Online Courses (MOOC) are a proposal to universalize the education, to offer it in a free way and with quality to people in any place in the world (DeBoer & Breslow, 2016; Solarte, Ramírez, & Jaramillo, 2017).

On the Open edX platform, xMOOCs are offered; The platform has a static structure and in it the learning material developed to a large extent by the teacher(s) in charge of the course is published. In the xMOOCs, there are students who navigate freely without necessarily following what is suggested by the content creators or the structure of the platform, there are other students who prefer to navigate, as imposed by a teacher or the online learning environment. Student navigations generate a set of navigations that create navigation patterns, which are recorded in a text file called tracking.log, for the Open edX platform (Fabián Anacona, Solarte, & González, 2018; Fabián Anacona, Solarte, & Ramírez, 2017; P. J. Guo & Reinecke, 2014; Yadav, 2014).

The xMOOCs offered by different universities in the world, has led to a growth in the participation of students in the courses (Alario-Hoyos et al.), generated an increase in the navigation patterns registered in the tracking.log file. so far, there is no study that identifies the navigation patterns according to the weighting of the relative frequency of an event (EF-ISF) (Fabián Anacona et al., 2018; Fabián Anacona et al., 2017; Burbano, Anacona, Solarte, & Ramírez, 2016). For this reason in this document proposes the use of the LexRank with Threshold algorithm to identify the EF-ISF navigation patterns registered in the tracking.log file of the xMOOC of the Open edX platform.

This article is organized as follows: section 1, describes in a general way the structure of the tracking.log file; section 2, representation of students with the vector space model and the EF-ISF; Section 3, describes the LexRank with Threshold algorithm; section 4, an example for identification of EF-ISF navigation patterns and section 5, conclusions and future work.

11 TRACKING.LOG OF THE OPEN EDX PLATFORM

The tracking.log is a JavaScript Object Notation (JSON) file that saves a record each time an activity or event is performed by the student(s), in the xMOOC course of the Open edX platform. The file has a structure general for all events, which is shown in Figura 5 (Fabián Anacona et al., 2018; Fabián Anacona et al., 2017; Yadav, 2014).

```
{ "username": "", "event_source": "", "name": "", "accept_language": "", "time": "", "agent":
  "", "page": "", "host": "", "session": "", "referer": "", "context": { "user_id": "", "org_id": "",
    "course_id": "", "path": ""}, "ip": "", "event": "", "event_type": "" }
```

Figura 5. General structure of the Open edX tracking.log file. Source: taken from (Fabián Anacona et al., 2018).

12 STUDENT REPRESENTATION

In this document the Vector Space Model is used, allowing to find the relation of the navigation events of the students in an xMOOC course, registered in the tracking.log file of the Open edX platform. In order to give a value to the navigation events, the weighting based on the relative frequency of an event (EF-ISF) is used, where w_{ij} represents the weight of the event i th of the student \vec{s}_j , as can be seen in equation (1) (Alguliev et al., 2011; Fabián Anacona et al., 2015; Fabián Anacona et al., 2018; Fabián Anacona et al., 2017).

$$w_{ij} = (f_{ij}) \left(\log \frac{N}{n_j} \right) \quad (1)$$

The calculation of similarity between the students \vec{s}_i and \vec{s}_j is performed with the similarity of cosines, as shown in equation (2) (Alguliev et al., 2011; Fabián Anacona et al., 2015; Fabián Anacona et al., 2017).

$$\text{simcos}(\vec{s}_i, \vec{s}_j) = \frac{\sum_{k=1}^m w_{ki} w_{kj}}{\sqrt{\sum_{k=1}^m w_{ki}^2 \sum_{k=1}^m w_{kj}^2}} \quad (2)$$

The representation in the multidimensional space of the set of student vectors is made with the Matrix of Events by Students ($Matrix\ EF - ISF_{m \times n}$) and the similarity of the same ones, is made with the Matrix of similarity of Cosines ($Matrix\ Of\ Similarity_{n \times n}$); the elements ij of the Matrixes $\in \mathbb{R}$ (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017).

13 LEXRANK WITH THRESHOLD ALGORITHM

The LexRank with Threshold algorithm is used to automatically generate summaries of one or multiple documents (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017). In this document it is proposed to make use of the Algorithm for the identification of EF-ISF navigation patterns.

The LexRank with Threshold is based on the concept of prestige in social networks. A social network is a map of relationships between entities (students, organizations)

that interact. Social networks are commonly represented in the form of graphs, where the nodes represent the entities and the links represent the relationships between the nodes (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

A set of students can be seen as a network of related students; some are more similar to each other, while others may share little information with the rest of the students. If a student is very similar to the other students, this can be considered as the most central or representative. That is why there are two key points to support this definition of centrality, first, how to define the similarity between two students and second, how to calculate the global centrality of a student given his similarity with other students (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

To define similarity, initially the set of students is represented in the vector space model and for the weighing of events EF-ISF from equation (1) is used. The similarity between two students is defined by the cosine similarity shown in equation (2); then the set of students is represented as a graph through an adjacency matrix (MatrixOfSimilarity), each value of the matrix corresponds to the cosine similarity between the students (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

Subsequently, for grade centrality, the student node must take into account the votes of each node and where those votes come from. This can be considered if each node has a centrality value distributed between the node itself and its neighbors; as shown in equation (3) (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

$$p(u) = \sum_{v \in adj[u]} \frac{p(v)}{\deg(v)} \quad (3)$$

Where $p(u)$ is the centrality of the u node, $adj[u]$ is the set of nodes that are adjacent to u , and $\deg(v)$ is the degree of the v node. This equation can be written in matrix notation as follows, see equation (4) (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

$$p = B^T p, \text{ which equals: } p^T B = p^T \quad (4)$$

Where matrix B is obtained from the adjacency matrix of the similarity graph by dividing each element by the sum of the corresponding row, see equation (5) (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

$$B(i, j) = \frac{A(i, j)}{\sum_k A(i, k)} \quad (5)$$

The sum of a row is equal to the degree of the corresponding node and each student is at least similar to himself, for this reason, the sums of the rows are different from zero. The equation (4) establishes that p^T is the own left vector of the matrix B with the own value corresponding to 1, to guarantee the existence of an own vector that can be identified and calculated in a unique way, the following needs to be taken into account (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

A stochastic matrix X , is the transition matrix of a Markov chain, an element $X(i, j)$ specifies the probability of transition from a state i to a state j . By probability axioms, all rows of a stochastic matrix must add 1. $x^n(i, j)$, is the probability of state i to reach state j in n transitions. A Markov string with stochastic matrix X converges to a stationary distribution yes (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

$$\lim_{n \rightarrow \infty} X^n = 1^T r \quad (6)$$

Where $1 = (1, 1, \dots, 1)$, and the vector r is called the stationary distribution of the Markov chain; each element of the vector r gives the asymptotic probability of ending in the corresponding long-term state, regardless of the starting state. A Markov string is irreducible if any state is accessible from any other state, that is, for all i, j there is an n such that $X^n(i, j) \neq 0$. A Markov string is aperiodic if for all i , $\gcd \{n : X^n(i, i) > 0\} = 1$. By Perron-Frobenius theorem, an irreducible and aperiodic Markov chain converges to a single stationary distribution (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

Since the similarity matrix B in equation (4) satisfies the properties of a stochastic matrix, it can be treated as a Markov chain. The centrality vector P corresponds to the stationary distribution of B . However, it must be ensured that the similarity matrix is irreducible and aperiodic. To solve this, a low probability is saved for jumping to any node in the graph, which makes the graph irreducible and aperiodic. If you assign a uniform probability to jump to any node in the graph, you get the following modified version of equation (3), which is known as the LexRank with Threshold algorithm, see equation (7) (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

$$p(u) = \frac{d}{N} + (1 - d) \sum_{v \in \text{adj}[u]} \frac{p(v)}{\text{deg}(v)} \quad (7)$$

Where N is the total number of nodes in the graph and d is a "damping factor", which is usually chosen in the range $[0.1, 0.2]$. Equation (7) can be written in matrix form as shown in equation (8) (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

$$p = [dU + (1 - d)B]^T p \quad (8)$$

Where U is a square matrix of $N \times N$ with all elements equal to $1/N$. The transition kernel $[dU + (1 - d)B]$ of the resulting Markov chain is a mixture of two U and B kernels. A random walker in this Markov chain chooses one of the adjacent states of the current state with probability $1 - d$, or jumps to any state in the graph, including the current state, with probability d (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

Below is the pseudocode of the LexRank with Threshold algorithm for the identification of EF-ISF navigation patterns in a set of students, as described in Algorithm 1 (Fabián Anacona et al., 2015; Fabián Anacona et al., 2017; Erkan & Radev, 2004).

```

InPut: MatrixOfSimilarity [n][n] of n students, threshold t, value dampingFactor
OutPut: Matrix L[n] with the student's EF-ISF navigation pattern scores
01 for i=1 to n do
02     Degree=0
03     for j=1 to n do
04         MatrixOfSimilarity[i][j] = ef-isf-simcos(S[i],S[j]);
05         if MatrixOfSimilarity[i][j] > t then
06             MatrixOfSimilarity [i][j] = 1;
07             Degree++;
08         else
09             MatrixOfSimilarity[i][j] = 0;
10         endif
11     endfor
15 endfor
16 for i=1 to n do
17     for j=1 to n do
18         MatrixOfSimilarity[i][j] = MatrixOfSimilarity[i][j] / Degree;
19     endfor
20 endfor
21 for i=1 to n do
22     for j=1 to n do
23         MatrixOfSimilarity[i][j] = (dampingFactor/n) + (1- dampingFactor)*
MatrixOfSimilarity[i][j];
24     endfor
25 endfor

```

```

26 L = MethodOfPower(MatrixOfSimilarity, n, ε ); // View Algorithm 2 MethodOfPower
27 return L

```

Algorithm 1. score calculation LexRank with Threshold. Source: adapted from (Erkan & Radev, 2004).

The Power Method describes how to calculate the Stationary Matrix of a Markov Chain, this is shown in Algorithm 2.

InPut: A stochastic, irreducible and aperiodic M matrix students, Tolerance Error ϵ

OutPut: eigen vector p

```

01  $p_0 = \frac{1}{N} \mathbf{1}$ 
02  $t = 0$ 
03 repeat
04      $t = t + 1$ 
05      $p_t = M^T p_{t-1}$ 
06      $\delta = ||p_t - p_{t-1}||$ 
07 until  $\delta < \epsilon$ ;
08 return  $p_t$ 

```

Algorithm 2. Power Method. Source: adapted from (Erkan & Radev, 2004).

14 IDENTIFICATION OF EF-ISF NAVIGATION PATTERNS WITH THE LEXRANK WITH THRESHOLD ALGORITHM

This section shows an example of the process of identifying EF-ISF navigation patterns with the LexRank with Threshold algorithm. For this, two events are used generated by a student of the course of Everyday Astronomy of Group B, from the year 2017 first academic period of the University of Cauca (Fabián Anacona et al., 2018).

For the example the following parameters are used: two events of a student, threshold = 0.9, damping factor = 0.15 and tolerance error = 1. The events are obtained from the tracking.log file of the Selene platform, from them some fields will be taken and the student will be called anonymous, this is shown in Tabla 3.

Tabla 3. Parts of the tracking.log file of two events of a student.

<p><i>First event</i> <i>username:</i> anónimo <i>name:</i> pause_video <i>time:</i> 2017-05-19T03:23:26.966429+00:00 <i>referer:</i> http://selene.unicauca.edu.co/courses/course-v1:Unicauca+AstronomiaCotidianaGrupoB+2017-1/</p>
--

```
courseware/9ee2d4e6ba4f4c8cb5a1aea3b66220a8/83d11edf15c446a5be18be0014144
fcb/
event:{"code":"0xlv1RoSXNk","id":"721d0f414ed44cdea917313b29848d48","curre
ntTime":0.793626168800354}
```

```
Second event
username: anónimo
name: load_video
time: 2017-05-19T03:21:00.319791+00:00
referer:http://selene.unicauca.edu.co/courses/course-
v1:Unicauca+AstronomiaCotidianaGrupoB+2017-l/
courseware/9ee2d4e6ba4f4c8cb5a1aea3b66220a8/83d11edf15c446a5be18be0014144
fcb/
event:{"code":"P2uUPX2y8Ks","id":"067290f9c4374d5da91cf317e5ccb863"}
```

Source: taken from (Fabián Anacona et al., 2018).

From the fragments of the record, seven fields are taken which will be called events, as shown in Tabla 4 (Fabián Anacona et al., 2018).

Tabla 4. log events.

```
event1 = pause_video,
event2 = AstronomiaCotidianaGrupoB
event3 = 9ee2d4e6ba4f4c8cb5a1aea3b66220a8
event4 = 83d11edf15c446a5be18be0014144fcb
event5 = 0xlv1RoSXNk
event6 = load_video
event7 = P2uUPX2y8Ks
```

Source: taken from (Fabián Anacona et al., 2018).

Applying equation (1), the following EF-ISF Matrix is obtained, see Matrix 1 (Fabián Anacona et al., 2018).

$$\begin{pmatrix} 0.3 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0.3 & 0 \\ 0 & 0.3 \\ 0 & 0.3 \end{pmatrix}$$

Matrix 1. EF-ISF Matrix. Source: taken from (Fabián Anacona et al., 2018).

With the EF-ISF Matrix and the equation (2), the Cosine similarity Matrix is calculated; applying the threshold = 0.9 and distributing the centrality of the student, the following Stochastic Matrix is created, see Matrix 2.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Matrix 2. Stochastic Matrix. Source: own source.

The Stochastic Matrix is transformed into the Matrix X, which is irreducible and aperiodic with damping factor = 0.15, see Matrix 3.

$$\begin{pmatrix} 0.925 & 0.075 \\ 0.075 & 0.925 \end{pmatrix}$$

Matrix 3. Matrix X. Source: own source.

Then calculate the Stationary Matrix with tolerance error = 1, this is shown in Matrix 4.

$$\begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}$$

Matrix 4. Stationary Matrix. Source: own source.

Based on the results of the Stationary Matrix, it is evident that the student's two EF-ISF navigation patterns have the same value, from which one can conclude, both have equal importance. The EF-ISF navigation patterns are now shown, which were determined with the LexRank with Threshold algorithm, see Tabla 5.

Tabla 5. EF-ISF navigation patterns of the anonymous student.

<i>EF-ISF Navigation1 Patterns</i>
AstronomiaCotidianaGrupoB -> 9ee2d4e6ba4f4c8cb5a1aea3b66220a8 -> 83d11edf15c446a5be18be0014144fcb -> pause_video -> 0xlv1RoSXNk
<i>EF-ISF Navigation2 Patterns</i>
AstronomiaCotidianaGrupoB -> 9ee2d4e6ba4f4c8cb5a1aea3b66220a8 -> 83d11edf15c446a5be18be0014144fcb -> load_video -> P2uUPX2y8Ks

Source: own source.

15 CONCLUSIONS AND FUTURE WORK

With the LexRank with Threshold Algorithm you can determine the EF-ISF navigation patterns of the set of student navigation patterns, recorded in the xMOOC tracking.log file of the Open edX platform.

The largest number $\in \mathbb{R}$ in the Stationary Matrix identifies a single EF-ISF navigation pattern, which represents the set of EF-ISF navigation patterns, from the set of

student navigation patterns, recorded in the xMOOC tracking.log file of the Open edX platform.

In the Stationary Matrix, the EF-ISF navigation patterns are represented with a value of $\in \mathbb{R}$, this value indicates the importance of each EF-ISF navigation pattern, in the set of EF-ISF navigation patterns.

As future work is proposed the implementation of LexRank with Threshold Algorithm, for the identification of navigation patterns EF-ISF, the set of navigation patterns of students, recorded in the xMOOC tracking.log file of the Open edX platform.

ACKNOWLEDGEMENTS

The authors are grateful for the support received by the project MOOC-Maker Construction of Management Capacities of MOOCs in Higher Education (561533-EPP-1-2015-1-ESEPPKA2-CBHE-JP) funded by the European Commission through the Erasmus+ Programme for the implementation and dissemination of the results set out in this article.

I would also like to thank the VRI 49694 MOOCMENTES project "Capacity Building for MOOC Management for Vocational Training, Rural Development and New Generations of Rural Students in Improving their Transit to Higher Education", co-financed within the framework of rural partnerships by the Ministry of National Education of Colombia.

ANEXO C - Artículo

ponencia en la Segunda Conferencia Internacional MOOC MAKER 2018, realiza el día jueves 11 de octubre de 2018, en el Hotel TRYP de la ciudad de Medellín, Colombia; el título de la ponencia fue “Modelo de Espacio Vectorial con ponderación basada en frecuencia relativa de eventos de navegación en una instancia de Open edX para caracterización del estudiantado”. El artículo de la ponencia se encuentra en las actas del evento y puede ser consultado en la siguiente dirección: <http://ceur-ws.org/Vol-2224/10.pdf>.

16 MODELO DE ESPACIO VECTORIAL CON PONDERACIÓN BASADA EN FRECUENCIA RELATIVA DE EVENTOS DE NAVEGACIÓN EN UNA INSTANCIA DE OPEN EDX PARA CARACTERIZACIÓN DEL ESTUDIANTADO

Fabián Andrés Anacona¹, Mario Solarte¹ and Gustavo Ramírez González¹

Universidad del Cauca, Popayán, Colombia
(fanacona, msolarte, gramirez)@unicauca.edu.co

Resumen. El uso de los Cursos en Línea Abiertos y Masivos –MOOC, está creciendo en los últimos años, debido a que diversas Universidades en el mundo los están ofreciendo, esto ha permitido generar una gran cantidad de datos por las actividades desarrolladas de los estudiantes en los cursos, estas actividades generan patrones que son registrados en el archivo tracking.log para xMOOC de la plataforma Open edX. Hasta el momento no hay un referente claro en la forma de procesar los patrones de navegación de este archivo. Por tal razón en este documento se presenta una propuesta, para caracterizar los patrones de navegación de los estudiantes en cursos xMOOC de la plataforma Open edX, a través del modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento.

Palabras claves: xMOOC, tracking.log, ponderación, patrón de navegación, evento, caracterización.

17 INTRODUCCIÓN

Los Cursos en Línea Abiertos y Masivos (MOOC, Massive Open Online Courses) son una propuesta basada en el Conectivismo, para universalizar la educación y ofrecer educación gratuita y de calidad a las personas desde cualquier lugar en el mundo (Siemens, 2004). Su principal objetivo es brindar alternativas educativas

soportadas en tecnologías de Internet de acceso abierto (Guàrdia et al., 2013; Hernández, 2013; T. Liyanagunawardena et al., 2013), además, los estudiantes inscritos a estos cursos virtuales tienen la oportunidad de aprender con los docentes de diferentes Universidades (Sonwalkar, 2013). Los MOOC presentan ciertas características similares como: empleo de material de aprendizaje (mapas de conocimiento, videoconferencias, textos, entre otros) (Fasihuddin, Skinner, & Athauda, 2015), por lo general son de corta duración (T. R. Liyanagunawardena et al., 2013), atraen a miles de estudiantes de diferentes lugares del mundo (Kizilcec, 2013) y algunos emiten un certificado de aprobación, el cual puede tener un valor monetario (Kizilcec, 2013; McAuley et al., 2010).

En la plataforma Open edX, se ofrecen xMOOCs, en este curso es el profesor quien crea la mayor parte del material de aprendizaje, el cual generalmente es organizado para el curso; la plataforma cuenta con una estructura lineal de arriba hacia abajo, en secciones, subsecciones y unidades, mostrando a través de la plataforma al estudiante, la organización del curso que plantea el docente. Por otra parte, los xMOOCs son criticados por replicar la enseñanza tradicional, basada en conferencias o clases magistrales y no tener presente las diferentes estrategias de aprendizaje, como la preferencia del estudiante por una estructura más o menos lineal en la organización del material (P. J. Guo & Reinecke, 2014; Jaramillo-Morillo, Sarasty, González, & Pérez-Sanagustín, 2017).

En el contexto del aprendizaje en línea, existen estudiantes que se sienten seguros en entornos de aprendizaje no lineales, es decir en la definición de sus propios caminos de aprendizaje, lo que indica que navegan libremente sin seguir necesariamente lo sugerido por los creadores de contenido o la estructura de la plataforma. Hay otros estudiantes, que prefieren seguir un camino de aprendizaje definido externamente, como lo impone un profesor o el entorno de aprendizaje en línea (P. J. Guo & Reinecke, 2014; Jaramillo-Morillo et al., 2017), las navegaciones de los estudiantes generan patrones de navegación (Fabián Anacona et al., 2017) los cuales son registrados en un archivo de texto llamado tracking.log, para la plataforma Open edX (Yadav, 2014).

Ahora bien, el auge de los MOOC ha implicado indirectamente un crecimiento de los datos en educación y no hay un referente claro en la creación de un conjunto de datos (Burbano et al., 2016), el cual permita la identificación de patrones de navegación acorde a la ponderación basada en la frecuencia relativa de un evento (EF-ISF) (Fabián Anacona et al., 2017), por esta razón, en este documento se presenta una propuesta para caracterizar los patrones de navegación de los estudiantes, registrados en el archivo de texto tracking.log de cursos xMOOC de la plataforma Open edX, a través del Modelo de Espacio Vectorial y EF-ISF.

Este artículo está organizado de la siguiente manera: la sección 2, describe la estructura general del archivo tracking.log; la sección 3, caracterización del conjunto de datos y un ejemplo; la sección 4, describe los pasos generales de algoritmos para el descubrimiento de patrones de navegación y la sección 5, conclusiones y trabajo futuro.

18 TRACKING.LOG DE LA PLATAFORMA OPEN EDX

Log es un registro en un archivo o medio de almacenamiento; cada vez que se realiza alguna actividad en el Sistema de Gestión de Aprendizaje (Learning Management System - LMS), se crea un historial o registro en un archivo de texto llamado tracking.log. Las actividades pueden ser las siguientes (Yadav, 2014):

1. Hacer clic en un enlace.
2. Ver o descargar un video.
3. Leer o descargar documentos, pdf y ppt, entre otros.

Existen varios beneficios de almacenar registros para el Sistema de Gestión de Aprendizaje, entre los que se encuentran: monitoreo del progreso del estudiante, verificar el uso del material de aprendizaje en el LMS por parte del estudiante; comprobar el uso de LMS por los estudiantes, como han utilizado los estudiantes el material de aprendizaje y el foro en LMS y consultas personalizadas, responder a preguntas a partir de la información almacenada en el archivo de texto (Yadav, 2014).

La entrada de registro, es una estructura de datos que almacena pocos campos, según el tipo de entrada o tipo de actividad, en el que la actividad también se conoce como evento. La entrada de registro es clave - valor (key - value) para los campos y sus valores, el cual se registra en formato JSON (Yadav, 2014).

Los eventos que se guardan son generados por el servidor, navegador o dispositivo móvil, el archivo tracking.log en la plataforma Open edX se encuentra en la ruta /edx/var/log/tracking/; a pesar de que dichos eventos pueden representar acciones distintas tienen un conjunto de campos comunes, cuya estructura general se presenta a continuación (Fabián Anacona et al., 2017):

```
{ "username": "", "event_source": "", "name": "", "accept_language": "", "time": "", "agent": "", "page": "", "host":  
  "", "session": "", "referer": "", "context": { "user_id": "", "org_id": "", "course_id": "", "path": ""}, "ip": "",  
  "event": "", "event_type": "" }
```

Figura 6. Estructura general del archivo tracking.log de Open edX. Tomado de (Fabián Anacona et al., 2017).

En la siguiente tabla se describe cada uno de los eventos.

Tabla 6. Descripción de campos. Tomado de (Palta & Vásquez, 2016).

Campo	Tipo	Descripción
accept_languaje	string	Determina el lenguaje por defecto del navegador.
agent	string	Reconoce al principal agente de usuario que inicia el evento.
context	object	Tiene un conjunto de sub campos que describen información del evento, los sub campos pueden ser comunes para todos los eventos como se muestra en la Tabla 2; sin embargo, pueden contener campos adicionales, ver Tabla 3.
event	object	Reconoce los campos específicos de cada evento iniciado, los campos pueden variar dependiendo del tipo de evento, los cuales pueden ser: de navegación y de inscripción, entre otros.
event_source	string	Determina el origen que genero el evento, los cuales pueden ser: "browser", "mobile", "server" y "task".
event_type	string	Identifica el tipo de evento registrado acorde al campo de "event_source".
host	string	Registra el sitio visitado por el usuario.
ip	string	La dirección "ip" del usuario cuando genera el evento.
name	string	Registra el tipo de evento lanzado.
page	string	Registra la URL de la página visitada al momento de generar el evento.
referer	string	Registra la cabecera HTTP del cual llego el evento.
session	string	Registra una cadena de 32 caracteres de la sesión del usuario.

Tabla 7. Campos del evento context. Tomado de (Palta & Vásquez, 2016).

Campo	Tipo	Descripción
course_id	string	Registra el curso en el que se generó el evento.
org_id	string	Registra la organización que oferta el curso.
path	string	identifica la URL que genera el evento.
user_id	number	Registra el usuario que género el evento.

Tabla 8. Sub campos adicionales del evento context. Tomado de (Palta & Vásquez, 2016).

Campo	Tipo	Descripción
course_user_tags	object	Registra la clave y el valor de la tabla user_api_usercoursetag para el usuario en el curso.
module	object	Registra información de los componentes que se involucran en el evento generado por el servidor.

19 CARACTERIZACIÓN DE LOS PATRONES DE NAVEGACIÓN CON EL MODELO DE ESPACIO VECTORIAL Y EF-ISF

Teniendo en un archivo de texto plano los eventos de navegación de los estudiantes, se procede a construir la Matriz de ponderación basada en la frecuencia relativa de un evento (Matriz EF-ISF), a través del Modelo de Espacio

Vectorial, el cual permite encontrar la relación de los eventos de navegación de los estudiantes en un curso MOOC. Para dar un valor a los eventos de navegación se utiliza, la ponderación basada en la frecuencia relativa de un evento (EF-ISF), con lo que se otorga mayor relevancia a los eventos menos frecuentes en el conjunto de estudiantes; pero a su vez más frecuentes en cada estudiante, como se observa en la siguiente ecuación (Fabián Anacona et al., 2017):

$$w_{ij} = (f_{ij}) \left(\log \frac{N}{n_j} \right) \quad (1)$$

donde w_{ij} representa el peso del evento i -ésimo del estudiante (\vec{s}_j) (Fabián Anacona et al., 2017).

La Matriz EF-ISF, es una matriz dispersa de pesos dispuestos en m eventos (filas) y n estudiantes (columnas), en el que elemento ij de la Matriz es denotado por w_{ij} , que corresponde al peso del evento i en el estudiante j , este peso es calculado de acuerdo a la ecuación (1) (Fabián Anacona et al., 2017), la cual se muestra a continuación (Anacona Anacona et al., 2015):

$$Matriz\ EF - ISF = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1j} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2j} & \dots & w_{2n} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ w_{i1} & w_{i2} & \dots & w_{ij} & \dots & w_{in} \\ \vdots & \vdots & \dots & \vdots & \dots & \vdots \\ w_{m1} & w_{m2} & \dots & w_{mj} & \dots & w_{mn} \end{pmatrix} \quad (2)$$

Al procesar los eventos de navegación con la ecuación (1), en la Matriz EF-ISF, se guarda la caracterización de los patrones de navegación de los estudiantes, acorde a la ponderación basada en la frecuencia relativa de un evento.

Ahora se muestra un ejemplo del proceso de construcción de la Matriz EF-ISF, a partir de dos eventos generados por un estudiante del curso de Astronomía Cotidiana del Grupo B, del año 2017 primer periodo académico de la Universidad del Cauca.

Los eventos son obtenidos del archivo tracking.log de la plataforma Selene, de ellos se tomarán algunos campos y al estudiante se lo llamara anónimo, esto se muestra a continuación:

Primer evento

username: anónimo

name: pause_video

time: 2017-05-19T03:23:26.966429+00:00
referer:http://selene.unicauca.edu.co/courses/course-
v1:Unicauca+AstronomiaCotidianaGrupoB+2017-
/courseware/9ee2d4e6ba4f4c8cb5a1aea3b66220a8/
83d11edf15c446a5be18be0014144fcb/
event:{\"code\": \"0xlv1RoSXNk\", \"id\": \"721d0f414ed44cdea917313b29848d48\",
\"currentTime\": 0.793626168800354}

Segundo evento

username: anónimo
name: load_video
time: 2017-05-19T03:21:00.319791+00:00
referer:http://selene.unicauca.edu.co/courses/course-
v1:Unicauca+AstronomiaCotidianaGrupoB+2017-
/courseware/9ee2d4e6ba4f4c8cb5a1aea3b66220a8/
83d11edf15c446a5be18be0014144fcb/
event:{\"code\": \"P2uUPX2y8Ks\", \"id\": \"067290f9c4374d5da91cf317e5ccb863\"}

De los fragmentos del registro, se toman siete (7) campos los cuales se denominarán eventos, como se muestra a continuación:

evento1 = pause_video,
evento2 = AstronomiaCotidianaGrupoB
evento3 = 9ee2d4e6ba4f4c8cb5a1aea3b66220a8
evento4 = 83d11edf15c446a5be18be0014144fcb
evento5 = 0xlv1RoSXNk
evento6 = load_video
evento7 = P2uUPX2y8Ks

Aplicando la ecuación (1), se obtiene la siguiente Matriz EF-ISF

$$\text{Matriz EF - ISF} = \begin{pmatrix} 0.3 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0.3 & 0 \\ 0 & 0.3 \\ 0 & 0.3 \end{pmatrix}$$

Con base en los resultados de la Matriz EF-ISF, se evidencia que el primer evento del estudiante anónimo difiere del segundo, en los eventos 1 y 5; el segundo evento del estudiante anónimo se diferencia del primero, en los eventos 6 y 7; pero ambos

tienen valores similares de 0.3 en la Matriz, de esta forma el patrón de navegación del estudiante, acorde a la ponderación basada en la frecuencia relativa de un evento, puede ser cualquiera de las dos navegaciones siguientes:

Navegacion1:

AstronomiaCotidianaGrupoB -> 9ee2d4e6ba4f4c8cb5a1aea3b66220a8 -> 83d11edf15c446a5be18be0014144fcb -> pause_video -> 0xlv1RoSXNk

Navegacion2:

AstronomiaCotidianaGrupoB -> 9ee2d4e6ba4f4c8cb5a1aea3b66220a8 -> 83d11edf15c446a5be18be0014144fcb -> load_video -> P2uUPX2y8Ks

20 ALGORITMOS PARA EL DESCUBRIMIENTO DE PATRONES DE NAVEGACIÓN

Para el descubrimiento de patrones de navegación en un conjunto de estudiantes, acorde a la ponderación basada en la frecuencia relativa de un evento (EF-ISF), se pueden utilizar algoritmos como: Análisis Semántico Latente (LSA), LexRank con Umbral, Máxima Cobertura y Mínima Redundancia (MCMR) y Moderador, entre otros (Anaconda Anaconda et al., 2015).

A continuación, se muestran los pasos generales del algoritmo LexRank con Umbral (Anaconda Anaconda et al., 2015):

1. Representar el conjunto de estudiantes en la Matriz EF-ISF y construir la Matriz de similitudes.
2. Aplicar el umbral a la Matriz de Similitudes.
3. Distribuir la centralidad del estudiante con los que tiene alguna relación, creando la Matriz Estocástica.
4. Transformar la Matriz Estocástica a irreducible y aperiódica.
5. Encontrar la Matriz Estacionaria.

Ahora se muestran los pasos generales del algoritmo Análisis Semántico Latente (LSA) (Anaconda Anaconda et al., 2015):

1. Representar el conjunto de estudiantes en la Matriz EF-ISF.
2. Aplicar Descomposición de Valores Singulares (SVD) en la Matriz EF-ISF.
3. Escoger los estudiantes que más aporten al tópico de la matriz U o V^T sin repetir estudiantes.
4. Elegir los estudiantes que tengan mayor similitud con el vector centroide.

Luego se muestran los pasos generales del algoritmo Máxima Cobertura y Mínima Redundancia (MCMR) (Anaconda Anaconda et al., 2015):

1. Representar el conjunto de estudiantes en la Matriz EF-ISF y construir la Matriz de similitudes.
2. Utilizar el algoritmo de Optimización por Enjambre de Partículas (PSO) para elegir los mejores estudiantes.
3. Evaluar la función objetivo con los estudiantes elegidos.
4. Repetir los pasos 2 y 3 hasta el número de iteraciones.
5. Escoger los estudiantes que tienen el mayor valor de la función objetivo.

Posteriormente, se muestran los pasos generales del algoritmo Moderador, basado en la Toma de Decisiones en Grupo (GDM) con consenso suave con doble ronda retroalimentada (Anaconda Anaconda et al., 2015):

1. Representar el conjunto de estudiantes en la Matriz EF-ISF y construir la Matriz de similitudes.
2. El moderador aplica proceso de consenso en dos rondas, en cada ronda se aplica fase de agregación.
3. El moderador sigue al proceso de selección y aplica fase de explotación.
4. El moderador genera los patrones de navegación de la GDM con consenso suave.

21 CONCLUSIONES Y TRABAJO FUTURO

Se describió una alternativa para caracterizar los patrones de navegación de los estudiantes, registrados en el archivo de texto tracking.log de xMOOC de la plataforma Open edX, a través del modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento.

Con la caracterización de los patrones de navegación de los estudiantes, registrados en el archivo de texto tracking.log de xMOOC de la plataforma Open edX, a través del modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento, se pueden determinar patrones de navegación acorde a EF-ISF.

Se describieron los pasos generales de los algoritmos: Análisis Semántico Latente (LSA), LexRank con Umbral, Máxima Cobertura y Mínima Redundancia (MCMR) y Moderador; para identificar patrones de navegación de un conjunto de estudiantes acorde a EF-ISF.

Como trabajo a futuro se plantea la implementación de la cartelización de los patrones de navegación, del archivo de texto tracking.log de xMOOC de la plataforma Open edX, con el modelo de espacio vectorial y la ponderación basada en la frecuencia relativa de un evento. De igual forma, la implementación de los

algoritmos: Análisis Semántico Latente (LSA), LexRank con Umbral, Máxima Cobertura y Mínima Redundancia (MCMR) y Moderador.

AGRADECIMIENTOS

Los autores agradecen el apoyo recibido por el proyecto MOOC-Maker Construction of Management Capacities of MOOCs in Higher Education (561533-EPP-1-2015-1-ESEPPKA2-CBHE-JP) financiado por la Comisión Europea a través del Programa Erasmus+ para la realización y divulgación de los resultados expuestos en el presente artículo.

De igual manera agradecer al proyecto VRI 49694 MOOCMenTES "Construcción de capacidades para la gestión de MOOC para la formación profesional, el desarrollo rural y nuevas generaciones de estudiantes rurales en el Mejoramiento de su Tránsito a la Educación Superior", co-financiado en el marco de alianzas rurales por el Ministerio de Educación Nacional de Colombia.

ANEXO D – Metodologías y Diagrama de Clases

22 METODOLOGÍAS USADAS EN EL PROYECTO DE INVESTIGACIÓN

Para el desarrollo de la propuesta se plantea tres Metodologías: Metodología de Investigación, Patrón de Investigación Iterativa (PII) (Pratt, 2009), que cuenta con cuatro fases principales, observación (del problema), identificación (del problema), desarrollo (de la solución) y prueba (de la solución); Metodología de Documentación, Mapeo Sistemático y Revisión Sistemática; Metodología de Desarrollo, una instanciación del Proceso Unificado Ágil de Desarrollo de Software, el cual cuenta con las siguientes fases: Iniciación, Elaboración, Construcción y Transición. Además, se contempla una fase adicional de Documentación y Divulgación que es transversal a todo el proyecto.

Para el desarrollo del Primer Objetivo Específico se llevará a cabo un ciclo con las siguientes actividades por fase:

- **Observación:** Búsqueda de documentación para pre procesamiento de datos en MOOC u otras, aplicando mapeo sistemático.
- **Identificación:** Aplicando el Mapeo sistemático y Revisión sistemática se hace la clasificación de la documentación a utilizar para el pre procesamiento
- **Desarrollo:** Caracterizar los conjuntos de datos en el modelo de espacio vectorial y la ponderación basado en la frecuencia relativa de un evento EF-ISF.

Para el desarrollo del Segundo Objetivo Específico se llevará a cabo un ciclo con las siguientes actividades por fase:

- **Observación:** Búsqueda de documentación y código fuente del algoritmo LexRank con Umbral, aplicando mapeo sistemático.
- **Identificación:** Utilizando mapeo sistemático y revisión sistemática, se hace la clasificación de la documentación a utilizar y los posibles algoritmos.
- **Desarrollo:** se adapta el algoritmo LexRank con Umbral, usando UP ágil.

Para el desarrollo del Tercer Objetivo Específico se realizará un solo ciclo con la siguiente actividad en la fase de Prueba, se realiza prueba de concepto en los conjuntos de datos, cuatro archivos Excel, por medio del algoritmo LexRank con Umbral, para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios y la centralidad de los patrones de navegación EF-ISF por sesión de los usuarios.

Documentación y Divulgación: En forma paralela al desarrollo de los anteriores ciclos y actividades, se realizará una actividad permanente de documentación y ciertos hitos de divulgación de resultados; entre ellos la realización de un artículo, para un evento o revista nacional o internacional, ver Anexos A, B y C, la monografía de grado, anexos y la sustentación de la tesis ante los jurados definidos por el Consejo de Facultad de la FIET.

23 DIAGRAMA DE CLASES

En las Figura 7 se muestran las clases desarrolladas para la identificación de los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios y la centralidad de los patrones de navegación EF-ISF por sesión de los usuarios.

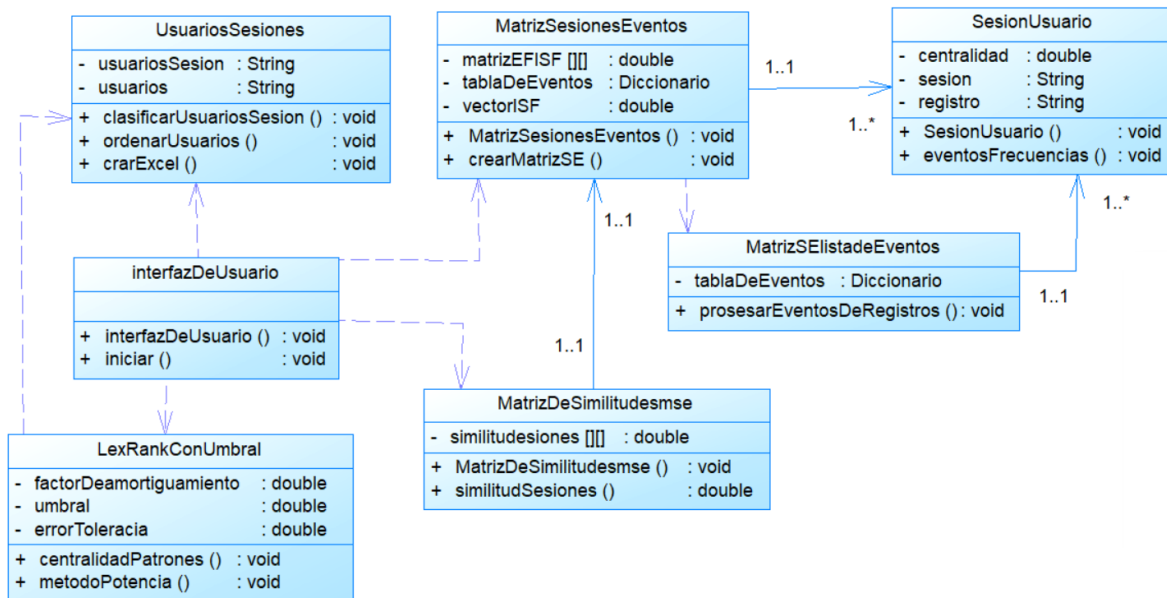


Figura 7. Diagrama de Clases. Fuente Propia.

En Tabla 9 se realza una descripción de las clases.

Tabla 9. Descripción de Clases. Fuente Propia.

Clase	Descripción
interfazDeUsuario	Es la interfaz que se le presenta al usuario para el cálculo de los patrones.
UsuariosSesiones	Clase que desarrolla el pre procesamiento, se para los usuarios de con sus respectivas sesiones.
LexRankConUmbral	Calcula los patrones de navegación EF-ISF por sesión de mayor centralidad de los usuarios y la centralidad de los patrones de navegación EF-ISF por sesión de los usuarios.
MatrizSesioneEventos	Con tiene las sesiones del usuario, representadas en el modelo de espacio vectorial.

MatrizDeSimilitudesmse	Calcula las similitudes de las sesiones del usuario i , representadas en el espacio vectorial.
MatrizSElistaDeEventos	Calcula el peso de los eventos de la sesiones del usuario i .
SesionUsuario	Contiene la sesión j del usuario i con respectivos registros.

REFERENCIAS

- Alario-Hoyos, C., Pérez-Sanagustín, M., Morales, M., Kloos, C. D., Hernández-Rizzardini, R., Román, M., . . . Gütl, C. MOOC-Maker: Tres Años Construyendo Capacidades de Gestión de MOOCs en Latinoamérica.
- Alguliev, R. M., Aliguliyev, R. M., Hajirahimova, M. S., & Mehdiyev, C. A. (2011). MCMR: Maximum coverage and minimum redundant text summarization model. *Expert Systems with Applications*, 38(12), 14514-14522. doi: <http://dx.doi.org/10.1016/j.eswa.2011.05.033>
- Anaconda Anaconda, F. A., Cobos Lozada, C. A., & Mendoza Becerra, M. E. (2015). *Algoritmo para generación automática de resúmenes extractivos genéricos de múltiples documentos basado en consensos*. (Monografía), Universidad del Cauca. (01)
- Anaconda, F., Cobos, C., & Mendoza, M. (2015). *Algoritmo para generación automática de resúmenes extractivos genéricos de múltiples documentos basado en consensos*. (Monografía), Universidad del Cauca. (01)
- Anaconda, F., Solarte, M., & González, G. (2018). Modelo de Espacio Vectorial con ponderación basada en frecuencia relativa de eventos de navegación en una instancia de Open edX para caracterización del estudiantado. *Segunda Conferencia Interanacional MOOC-MAKER Medellín 2018*, 2224, 87-95.
- Anaconda, F., Solarte, M., & Ramírez, G. (2017). Descubrimiento de patrones de navegación en Open edX—una aproximación arquitectónica. *Ingeniería e Innovación*, 5(1).
- Burbano, M., Anaconda, F., Solarte, M., & Ramírez, G. (2016). Informe sobre tecnologías Web Semántica y Social en cursos MOOC. *MOOC-Maker*, 1-17.
- DeBoer, J., & Breslow, L. (2016). *Work in Progress: Student Behaviors Using Feedback in a Blended Physics Undergraduate Classroom*. Paper presented at the Proceedings of the Third (2016) ACM Conference on Learning@ Scale.
- edX. (2013, 14-03-2013). Release of XBlock Source Code. Retrieved 24-09-2016, 2016, from <https://www.edx.org/press/edx-takes-first-step-toward-open-source>
- edX. (2016). Quality education for everyone, everywhere Retrieved 30-10-2016, 2016, from <https://www.edx.org/about-us>
- Erkan, G., & Radev, D. R. (2004). LexRank: graph-based lexical centrality as salience in text summarization. *J. Artif. Int. Res.*, 22(1), 457-479.
- Fasihuddin, H., Skinner, G., & Athauda, R. (2015). Towards adaptive open learning environments: Evaluating the precision of identifying learning styles by tracking learners' behaviours. *Education and Information Technologies*, 1-19.
- Fox, A. (2013). From moocs to spocs. *Communications of the ACM*, 56(12), 38-40.
- Gong, Y., & Liu, X. (2001). *Generic text summarization using relevance measure and latent semantic analysis*. Paper presented at the Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval, New Orleans, Louisiana, USA.
- Grossman, S. I. (2001). *Álgebra Lineal*: Editorial Offset, S.A. de C.V, Durazno No. 1 esq. Ejido, Col. Las Peritas, Tepepan Xochimilco, C.P. 16010 México D.F.
- Guàrdia, L., Maina, M., & Sangrà, A. (2013). MOOC design principles: A pedagogical approach from the learner's perspective. *eLearning Papers*(33).

- Guo, P. J., & Reinecke, K. (2014). *Demographic differences in how students navigate through MOOCs*. Paper presented at the Proceedings of the first ACM conference on Learning@ scale conference.
- Guo, W. (2014). *From SPOC to MPOC--The Effective Practice of Peking University Online Teacher Training*. Paper presented at the Educational Innovation through Technology (EITT), 2014 International Conference of.
- Hernández, S. (2013). SCOPEO INFORME N°2: MOOC: Estado de la situación actual, posibilidades, retos y futuro.
- Hmedna, B., El Mezouary, A., Baz, O., & Mammass, D. (2017). Identifying and tracking learning styles in MOOCs: A neural networks approach. *International Journal of Innovation and Applied Studies*, 19(2), 267.
- Jaramillo-Morillo, D., Sarasty, M. S., González, G. R., & Pérez-Sanagustín, M. (2017). *Follow-Up of Learning Activities in Open edX: A Case Study at the University of Cauca*. Paper presented at the European Conference on Massive Open Online Courses.
- Kizilcec, R. F. (2013). *Collaborative Learning in Geographically Distributed and In-person Groups*. Paper presented at the Proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED).
- Liyanagunawardena, T., Williams, S., & Adams, A. (2013). The impact and reach of MOOCs: a developing countries' perspective. *eLearning Papers*(33).
- Liyanagunawardena, T. R., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008-2012. *The International Review of Research in Open and Distributed Learning*, 14(3), 202-227.
- Mackness, J., Mak, S., & Williams, R. (2010). The ideals and reality of participating in a MOOC.
- Manning, C. D., Raghavan, P., & Schütze, H. (2009). *An introduction to information retrieval*. Cambridge: Cambridge University Press.
- Martin, F. G. (2012). Will massive open online courses change how we teach? *Communications of the ACM*, 55(8), 26-28.
- Matías González, H., & Pérez Avila, A. (2014). Los Cursos En Línea Masivos Y Abiertos (MOOC) Como Alternativa Para La Educación a Distancia (Massive Open Online Courses (MOOC), an Alternative to Distance Learning).
- McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). The MOOC model for digital practice.
- Palta, R., & Vásquez, J. (2016). *Descubrimiento de patrones de interacción en cursos MOOC en entornos ONLINE: Un enfoque utilizando Minería de Procesos. Caso de Estudio: "Curso de la metodología DICREVOA en Open edX"*. Universidad de Cuenca.
- Pratt, K. S. (2009). *Design Patterns for Research Methods: Iterative Field Research*. Paper presented at the Association for the Advancement of Artificial Intelligence.
- Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Commun. ACM*, 18(11), 613-620. doi: 10.1145/361219.361220
- Semana. (2014). Las clases que no se quedan sin cupos. Retrieved 22-02, 2017, from <http://www.semana.com/educacion/articulo/cursos-en-linea-masivos-abiertos-en-colombia/409366-3>

- Siemens, G. (2004). A learning Theory for the Digital Age. *Elearnspace Everything Learning*, 1-8.
- Singhal, A. (2001). Modern Information Retrieval: A Brief Overview. *IEEE Data Eng. Bull.*, 24(4), 35-43.
- Solarte, M., Ramírez, G. A., & Jaramillo, D. A. (2017). Hábitos de ingreso y resultados en las evaluaciones en cursos en línea masivos con reconocimiento académico. *Ingeniería e Innovación*, 5(1).
- Sonwalkar, N. (2013). *The first adaptive MOOC: A case study on pedagogy framework and scalable cloud Architecture—Part I*. Paper presented at the MOOCs Forum.
- Yadav, R. K. (2014). *Understanding Logs in edX for Monitoring Student Progress*. (Doctoral Thesis), Indian Institute of Technology, Bombay.
- Zapata-Ros, M. (2015). MOOCs, una visión crítica y una alternativa complementaria: La individualización del aprendizaje y de la ayuda pedagógica. *Campus virtuales*, 2(1), 20-38.
- Zazo Rodríguez, Á. F., Figuerola Paniagua, C. G., Berrocal, J. L. A., & Gómez Díaz, R. (2002). Recuperación de Información Utilizando el Modelo Vectorial. Participación en el taller CLEF-2001 (D. d. I. y. Automática, Trans.) (pp. 1-40): Universidad de Salamanca