

## Adaptación del Modelo de Espacio Vectorial para la Comparación de Redes Semánticas Naturales

### *Adaptation of the Vector Space Model for the Comparison of Natural Semantic Networks*

**González Lara, Aída Lucina & Torres Guerrero, Francisco**

Universidad Autónoma de Nuevo León

Facultad de Ingeniería Mecánica y Eléctrica. San Nicolás de los Garza, N.L.

[aida.gonzalezlr@uanl.edu.mx](mailto:aida.gonzalezlr@uanl.edu.mx)

[francisco.torresgrr@uanl.edu.mx](mailto:francisco.torresgrr@uanl.edu.mx)

**Resumen.** Actualmente la técnica de Redes Semánticas Naturales (RSN) es usada ampliamente para la representación de significados, es de interés comparar redes de diferentes poblaciones acerca de una misma temática. En este trabajo se presenta la adaptación de un modelo de Recuperación de Información denominado Modelo de Espacio Vectorial (MEV) para mejorar el alcance y optimizar la forma de comparar dos RSN. Para realizar esta investigación, se generaron cuatro RSN de la temática de Ecología con la participación de 159 individuos, se efectuó la comparación entre ellas y como resultado se obtuvo el valor de similitud. Se llevó a cabo el análisis de los resultados obtenidos en la comparación el cual indica que sí es posible realizar la comparación entre RSN con este método aun cuando la cantidad de participantes en las redes sea distinta, lo que proporciona una nueva herramienta para la amplia variedad de campos en los que se usan la Redes Semánticas Naturales.

**Palabras clave.** Comparación de Redes Semánticas, Modelo de Espacio Vectorial, Redes Semánticas Naturales, Similitud.

**Abstract. Abstract.** At present the Natural Semantic Networks (NSN) technique is widely used for the representation of meanings, it is of interest to compare networks of different populations regarding the same subject. This work presents the adaptation of an Information Retrieval model called Vector Space Model (VSM) to improve the scope and optimize the way to compare two NSN. In order to carry out this research, five NSN on the subject of Ecology were generated with the participation of 159 individuals, the comparison between them was made, and as a result the similarity value was obtained. The analysis of the results obtained from the comparison was conducted, which indicates that it is possible to perform the comparison between NSN with this method, even if the number of participants in the networks is different, providing a new tool for the wide variety of fields wherein the Natural Semantic Networks are used.

**Keywords.** Comparison of Semantic Networks, Vector Space Model, Natural Semantic Networks, Similarity.

### **Introducción**

El uso de redes semánticas ha permitido dimensionar la representación del conocimiento en diferentes niveles en la aplicación de diversas áreas de investigación, existen distintos métodos para la construcción de redes semánticas, en el caso de las Redes Semánticas Naturales (RSN), esta es una técnica que se basa en la generación de definiciones conceptuales de participantes para, con dichas definiciones, construir una red semántica; la diferencia con otros métodos que producen redes semánticas, está en el origen de las definiciones con las que la red es construida, ya que en las RSN estas definiciones son proporcionadas por el propio participante (Figuroa et al., 1981).

La Recuperación de Información (*Information Retrieval*) es un campo de estudio que se relaciona con la estructura, el análisis, el almacenamiento y la búsqueda de información, tales acciones son requeridas para dar respuesta de manera eficiente a consultas de información que son realizadas por los usuarios (Baeza-Yates & Ribeiro-Neto, 1999; Salton G., 1989).

La finalidad más importante de la Recuperación de Información es mostrar a los usuarios aquella información que es relevante para ellos (Hanani et al., 2001); para lograr lo anterior los sistemas de recuperación de información aplican técnicas de evaluación con el objetivo de comparar diversos documentos con una consulta solicitada por el usuario, tal proceso es desarrollado de acuerdo al Modelo de Recuperación de Información utilizado (Salton & Buckley, 1988; Van Rijsbergen, 1979).

El Modelo del Espacio Vectorial (MEV) es un modelo muy utilizado por los Sistemas de Recuperación de información, una de sus principales características es el uso de técnicas de ponderación de los términos en los documentos ya que posibilita determinar la importancia de un término en un documento específico y permitir la coincidencia parcial entre el documento y la consulta en el proceso de recuperación, también facilita la calificación de la relevancia de los documentos con respecto de las consultas (Salton & McGill, 1986).

Dado que el MEV realiza la comparación entre elementos de texto (palabras) de documentos y consultas para determinar la relevancia de los documentos solicitados, estos modelos pueden ser aplicados para comparar Redes Semánticas Naturales, ya que se trata también de comparación de definidores expresados en palabras.

### ***Justificación***

Actualmente las RSN son usadas ampliamente en diferentes áreas de investigación y aplicación debido a su alcance en la representación del conocimiento, en donde una red acerca de una temática puede reflejar la representación de una muestra; es de interés para los investigadores comparar redes de diferentes poblaciones acerca de una misma temática, actualmente se utiliza un indicador llamado valor Q que representa el consenso grupal (Figuroa et al., 1981).

Para mejorar el alcance y optimizar la forma de comparar dos redes es necesario realizar un estudio que permita diseñar un nuevo modelo de comparaciones desde una perspectiva de recuperación de información, que tenga como base el cálculo de similitud entre las redes.

### ***Definición del Problema***

Los trabajos acerca de RSN se han dirigido principalmente para aplicaciones de psicología social con el objetivo de investigar acerca de los significados, diversos estudios muestran que esta técnica ha sido utilizada para representar significados con distintos temas (Arévalo Aguilera, 2010; Murillo Estepa & Becerra Peña, 2009; De la Cruz Flores & Abreu Hernández, 2012; Trejo-Lucero et al., 2011). Sin embargo existen también investigaciones de RSN en el campo de la evaluación cognitiva en las que se comparan las redes generadas, ya sea entre participantes del mismo grupo, o de participantes de diferentes grupos (Collipal et al., 2004), para lo que se requiere hacer comparaciones entre las redes semánticas generadas

En algunos trabajos se menciona que el método utilizado actualmente para la comparación, mediante el cálculo del valor Q no representa en algunos casos un resultado representativo en la comparación de las redes (García & Jiménez Vidal, 1996; Padilla Montemayor & Rodríguez Nieto, 2011) por lo que se propone adaptar el MEV para comparar redes semánticas naturales.

### ***Objetivo***

Adaptar el Modelo Espacio Vectorial para realizar la comparación entre Redes Semánticas Naturales y evaluar sus resultados.

### ***Hipótesis***

Es posible comparar dos redes semánticas naturales mediante la adaptación del Modelo de Espacio Vectorial.

## **Marco Teórico**

### ***Redes Semánticas Naturales***

La finalidad de esta técnica es estudiar la manera en que se establecen los significados desde la memoria semántica, una de sus características principales es que las propias redes determinan el significado de los conceptos; se describe de manera breve la técnica RSN, la cual se divide en dos partes fundamentales (Reyes-Lagunes, 1993):

- a) Solicitar a los sujetos participantes que definan con claridad a la palabra estímulo que se les presenta, mediante el uso de palabras individuales que consideren que están relacionadas con ésta, a las palabras mencionadas se les llamará definidores; el tiempo que se dará a los sujetos para definir cada concepto es de 60 segundos, debido a que si se les proporciona un tiempo mayor, los participantes tienden a proporcionar asociaciones libres en lugar de definiciones.
- b) Pedir a los participantes que califiquen asignando el número 10 a la palabra más cercana o relacionada con la palabra estímulo, el número 9 a la que sigue en importancia, etc., se asignará 1 al concepto menos relevante, se deben jerarquizar todas las palabras dadas como definidores.

Una vez que se tienen los definidores de la palabra estímulo para la cantidad de sujetos establecida que participa en el estudio, se genera una tabla de frecuencias, en la que se calcula el peso semántico tomando en cuenta la frecuencia con que fue mencionado cada uno de los definidores en cada valor jerárquico.

El paso siguiente consiste en calcular los indicadores propuestos por Figueroa, González y Solís (Figueroa et al., 1981) y Valdez (Valdez M, 2000) que se definen a continuación:

Valor J: Son todas las palabras utilizadas por los sujetos para definir los conceptos que actuaron como estímulo (definidores).

Valor M: Se calcula tomando en cuenta el peso semántico asignado por los sujetos a cada concepto definidor de acuerdo a su importancia o pertinencia.

Grupo SAM (*Semantic Association Memory*): Está formado por los 10 definidores con el Valor  $M$  más elevado de cada uno de los conceptos que actúan como estímulo. Forman el significado de la red sobre el concepto estímulo.

### **Modelo de Espacio Vectorial**

En este Modelo de Recuperación de Información, el texto es representado por un vector de términos, los términos comúnmente son palabras; cualquier texto puede ser representado por un vector en un espacio dimensional (Salton et al., 1975). En el MEV los documentos se representan a partir de vectores, de la siguiente manera (Salton & McGill, 1986):

$$\vec{d}_j = (w_{1j}, w_{2j}, w_{3j}, w_{4j}, \dots, w_{nj}) \quad (1)$$

Donde  $n$  es igual al número total de elementos de representación considerados y por su parte  $w$  indica el peso que el término en concreto tiene para el documento  $j$ , el peso de un término es una medida de su importancia en la representación del documento.

Las comparaciones se realizan de acuerdo a lo siguiente: dos vectores pueden ser representados en el hiperplano y pueden ser medidas las diferencias de dirección entre ambos de la misma forma que se haría en un plano bidimensional mediante la comparación del coseno del ángulo que forman (Singhal, 2001). En la Figura 1 se muestra la representación del MEV para la comparación entre la consulta y lo documentos.

Para asignar una puntuación numérica a un documento para una consulta, este modelo mide la similitud entre el vector de consulta representado como  $\vec{q}$  y el vector del documento representado como  $\vec{d}$ , típicamente el ángulo entre dos vectores es usado como una medida de divergencia entre los vectores, y el coseno del ángulo es usado como la similitud numérica (Salton et al., 1975).

El cálculo de la similitud cosenoidal se realiza mediante la siguiente ecuación (Muflikhah & Baharudin, 2009).

$$SimCos(a,b) = \frac{\sum_{i=1}^k a_i * b_i}{\sqrt{\sum_{i=1}^k a_i^2} * \sqrt{\sum_{i=1}^k b_i^2}} \quad (2)$$

Esta ecuación produce un valor entre 0 y 1, en donde un valor de 0 indica la perpendicularidad (ortogonalidad) de los vectores, lo cual indica que se tendrían dos vectores en dirección completamente diferente, y un valor de 1 indica que los vectores tienen una dirección idéntica (Baeza-Yates & Ribeiro-Neto, 1999).

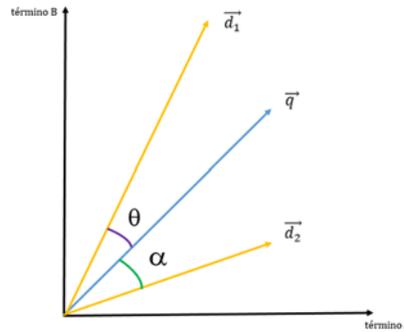


Figura 1. El Modelo de Espacio Vectorial representa documentos y consultas como vectores para compararlos.

### Metodología

A continuación se especifican las actividades que se realizaron para cumplir con el objetivo de la investigación, las cuales serán detalladas posteriormente:

- a. Aplicar la técnica de RSN con el objetivo de generar las redes semánticas a utilizar como material de comparación.
- b. Realizar el pre-procesamiento de las RSN generadas para la preparación para la comparación.
- c. Comparar las RSN mediante el Modelo del Espacio Vectorial.
- d. Análisis de resultados de comparación.

### Generación de Redes Semánticas Naturales

En este trabajo se generan Redes Semánticas Naturales para la propia investigación mediante una aplicación web conectada a una base de datos, participaron 159 individuos, estudiantes de ingeniería de entre 18 y 21 años, familiarizados con el uso de internet.

La temática seleccionada para la representación de significado mediante las RSN fue Ecología. La Tabla 1 muestra la cantidad de participantes por RSN generada.

Los conceptos estímulo utilizados en este estudio para la generación de la RSN fueron: contaminación, reciclaje, ciudad, reforestación, ecología, medio ambiente y naturaleza. Para realizar este proceso se utilizó una aplicación web denominada *RSNweb* (Ingeniería Cognitiva, 2014).

Tabla 1. Representa la cantidad de participantes por cada RSN generada.

RSN generada	Cantidad de participantes
ECOLOGÍA 1	32
ECOLOGÍA 2	47
ECOLOGÍA 3	35
ECOLOGÍA 4	45

El proceso de aplicación de la técnica de RSN se realizó en cuatro grupos de clase en donde cada alumno cuenta con una computadora personal para realizar la prueba. Se solicitó al estudiante que accediera a un enlace de la aplicación que le fue proporcionado para iniciar con el registro de sus datos personales, el siguiente paso es mostrar al participante una pantalla con las instrucciones a seguir, en donde se explica de qué manera debe realizar la prueba y permite hacer un ejercicio de práctica; posteriormente se presenta al participante los conceptos estímulo para que escriba palabras asociadas a estos, el participante dispondrá de 60 segundos para tal acción; finalmente se le solicita que califique cada uno de los definidores de acuerdo a su cercanía o relación con el concepto estímulo.

Los definidores proporcionados son almacenados en una base de datos relacional, la cual es accedida por un programa que permite hacer la limpieza de los definidores, en este proceso se corrige la ortografía de los conceptos y se estandarizan de acuerdo a criterios definidos.

Una vez que se concluye el proceso de limpieza se realiza la identificación de los grupos SAM de la red, tomando los primeros diez definidores con el valor M más alto, se identifica también el valor J denotado para la cantidad de definidores diferentes por estímulo como se muestra en la Figura 2 para la red ECOLOGÍA 1.

### ***Modelo de Espacio Vectorial en Redes Semánticas Naturales***

De acuerdo con la definición del Modelo de Espacio Vectorial, se realiza la búsqueda de términos en documentos; para trasladar lo anterior al campo de las Redes Semánticas Naturales, se considera que una RSN se puede interpretar como un documento, cada definidor como un término del documento y el peso del término sería el valor M del definidor, de tal manera que la ecuación 1 se transforma en:

$$\overrightarrow{rsn}_i = (M_{1i}, M_{2i}, M_{3i}, M_{4i}, M_{ni}) \quad (3)$$

En donde  $n$  es la cantidad de definidores diferentes en la red, es decir el valor  $J$  de la RSN y  $M$  representa el valor M que tiene el definidor en la RSN  $i$ . En el caso de la comparación de redes generadas con diferente cantidad de participantes es necesario normalizar el valor M de los definidores. Para representar la información de una forma adecuada con la finalidad de realizar la comparación mediante el MEV es necesario realizar el pre-procesamiento de la información.

PROYECTO - ECOLOGIA 1

CIUDAD			CONTAMINACION			ECOLOGIA			MEDIO AMBIENTE		
F	Definidor	M	F	Definidor	M	F	Definidor	M	F	Definidor	M
1	EDIFICIO	218	2	BASURA	243	4	VERDE	119	2	NATURAL	133
2	CARRO	198	1	HUMO	109	4	ARBOL	117	4	ARBOL	117
2	CONTAMINACION	142	1	SUCIO	101	1	RECICLAJE	109	3	ANIMAL	96
1	PERSONA	137	1	ENFERMEDAD	76	4	PLANTA	102	2	ECOSISTEMA	96
1	CASA	115	1	SMOG	71	2	NATURAL	92	4	PLANTA	92
1	TRAFICO	82	2	CARRO	66	4	AMBIENTE	86	3	ECOLOGIA	70
1	METROPOLI	66	1	DESECHO	53	3	ANIMAL	79	2	CONTAMINACION	57
1	GENTE	56	4	AMBIENTE	47	1	LIMPIA	77	4	VERDE	51
1	CALLE	49	1	FABRICA	42	2	ECOSISTEMA	53	2	AGUA	49
1	TIENDA	33	1	TOXICO	33	1	CUIDADO	51	2	AIRE	49
J = 97			J = 79			J = 66			J = 85		
G = 18.6			G = 21			G = 8.8			G = 8.4		
NATURALEZA			RECICLAJE			REFORESTACION					
F	Definidor	M	F	Definidor	M	F	Definidor	M			
3	ANIMAL	199	1	REUSAR	202	4	ARBOL	270			
4	PLANTA	150	2	BASURA	176	4	PLANTA	171			
4	ARBOL	136	1	PLÁSTICO	125	2	BOSQUE	126			
4	VERDE	115	1	PAPEL	100	1	OXIGENO	86			
2	AGUA	85	4	AMBIENTE	66	3	ECOLOGIA	70			
1	FLORA	75	1	LATA	55	4	AMBIENTE	59			
2	BOSQUE	70	3	ECOLOGIA	55	1	TIERRA	58			
1	VIDA	63	1	ORGANICO	52	4	VERDE	47			
1	FAUNA	56	1	BOTELLA	45	1	SEMILLA	45			
1	RIO	55	1	BOTE	45	2	AIRE	35			
J = 85			J = 59			J = 73					
G = 14.4			G = 15.7			G = 23.5					

Figura 2. Grupos SAM de RSN del Proyecto ECOLOGÍA1, la columna M representa el valor M calculado para cada definidor.

*Pre-procesamiento de la información*

El primer paso en el pre-procesamiento es vaciar en un vector denominado inicial[70] todos los definidores de los grupos SAM de la RSN y en otro vector se vacían los valores M correspondientes a los definidores, este vector se denomina inicialM[70], en la Figura 3 se presentan estos vectores para la RSN ECOLOGÍA 1.

inicial[70]	inicialM[70]		
AGUA	49	ENFERMEDAD	76
AGUA	85	FABRICA	42
AIRE	49	FAUNA	56
AIRE	35	FLORA	75
AMBIENTE	47	GENTE	56
AMBIENTE	86	HUMO	109
AMBIENTE	66	LATA	55
AMBIENTE	59	LIMPIA	77
ANIMAL	79	MONTERREY	51
ANIMAL	96	NATURAL	92
ANIMAL	199	NATURAL	133
ARBOL	117	ORGANICO	52
ARBOL	117	OXIGENO	86
ARBOL	136	PAPEL	100
ARBOL	270	PERSONA	137
BASURA	243	PLANTA	102
BASURA	176	PLANTA	92
BOSQUE	70	PLANTA	150
BOSQUE	126	PLANTA	171
BOTE	45	PLÁSTICO	125
BOTELLA	45	RECICLAJE	109
CALLE	49	REUSAR	202
CARRO	198	RIO	55
CARRO	66	SEMILLA	45
CASA	115	SMOG	71
CONTAMINACION	142	SUCIO	101
CONTAMINACION	57	TIENDAS	33
CUIDADO	51	TIERRA	58
DESECHO	53	TOXICO	33
ECOLOGIA	70	TRAFICO	82
ECOLOGIA	55	VERDE	119
ECOLOGIA	70	VERDE	51
ECOSISTEMA	53	VERDE	115
ECOSISTEMA	96	VERDE	47
EDIFICIO	218	VIDA	63

Figura 3. Formación de vectores inicial e inicialM con los definidores y los valores M de la RSN para pre-procesamiento de la información.

Para obtener el formato necesario para la comparación mediante el Modelo de Espacio Vectorial es necesario agrupar los definidores iguales y calcular una sumatoria de sus valores M para posteriormente realizar la normalización de los valores, para este proceso se desarrolló un programa en *Scilab*, el cual tiene como entrada los vectores *inicial[70]* e *inicialM[70]* y generará como salida los siguientes vectores:

- *final[j]* Vector con los definidores sin repetición
- *finalM[j]* Vector con la suma de los valores M por definidor
- *finalMNor[j]* Vector con la suma de los valores M normalizados

La Figura 4 expone los vectores generados por el programa para la RSN ECOLOGIA 1.

<b>final[j]</b>	<b>finalM[j]</b>	<b>finalMNor[j]</b>
AGUA	134	0.166392
AIRE	84	0.08402
AMBIENTE	258	0.370675
ANIMAL	374	0.561779
ARBOL	640	1
BASURA	419	0.635914
BOSQUE	196	0.268534
BOTE	45	0.019769
BOTELLA	45	0.019769
CALLE	49	0.026359
CARRO	264	0.38056
CASA	115	0.135091
CONTAMINACION	199	0.273476
CUIDADO	51	0.029654
DESECHO	53	0.032949
ECOLOGIA	195	0.266886
ECOSISTEMA	149	0.191104
EDIFICIO	218	0.304778
ENFERMEDAD	76	0.07084
FABRICA	42	0.014827
FAUNA	56	0.037891
FLORA	75	0.069193
GENTE	56	0.037891
HUMO	109	0.125206
LATA	55	0.036244
LIMPIA	77	0.072488
MONTERREY	51	0.029654
NATURAL	225	0.31631
ORGANICO	52	0.031301
OXIGENO	86	0.087315
PAPEL	100	0.110379
PERSONA	137	0.171334
PLANTA	515	0.794069
PLÁSTICO	125	0.151565
RECICLAJE	109	0.125206
REUSAR	202	0.278418
RIO	55	0.036244
SEMILLA	45	0.019769
SMOG	71	0.062603
SUCIO	101	0.112026
TIENDAS	33	0
TIERRA	58	0.041186
TOXICO	33	0
TRAFICO	82	0.080725
VERDE	332	0.492586
VIDA	63	0.049423

Figura 4. Muestra la salida del programa de pre-procesamiento para obtener los vectores para la comparación.

### ***Comparación de RSN mediante el Modelo de Espacio Vectorial***

Al trasladar el proceso que sigue el Modelo de Espacio Vectorial hacia la comparación entre Redes Semánticas Naturales, se considera que la comparación se realizaría entre dos redes semánticas en donde los definidores de los grupos SAM se compararán y el peso de los términos estaría dado por los valores M normalizados de los

definidores, las técnicas de comparación son las indicadas para MEV y el resultado de la similitud sería interpretado como el parecido que existe entre las redes comparadas; la analogía anterior es representada en la Figura 5.

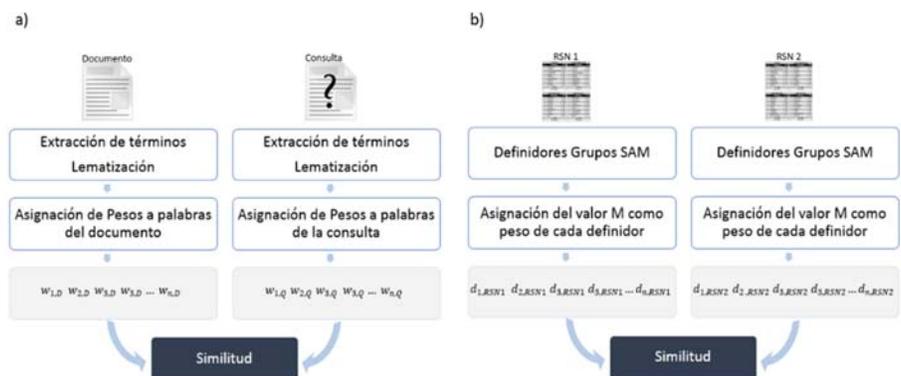


Figura 5. Representa la analogía entre el proceso del Modelo de Espacio Vectorial clásico de Recuperación de Información y el proceso del Modelo de Espacio Vectorial aplicado a la comparación de RSN. En a) se muestra el proceso de comparación entre un documento y una consulta y en b) se muestra el proceso de comparación entre dos RSN.

Para realizar la comparación entre dos RSN se estructura una matriz en donde las columnas representan los definidores y las filas las RSN a comparar; cada elemento representa el Valor M normalizado del definidor  $j$  en la RSN  $i$ . Se realiza la comparación entre las RSN mediante la medida de similitud cosenoidal,

Se desarrolló un programa en *Scilab* para realizar el proceso de comparación, el cual tiene como salida una matriz denominada  $matrizcomp[2,j]$ , que contiene los definidores y los valores M normalizados correspondientes por definidor, a aquellos definidores que sólo existen en una de las redes se les asigna un valor M de cero en el espacio del definidor correspondiente y son mostrados con espacios vacíos. La Tabla 2 muestra el ejemplo de la matriz de salida para los primeros diez definidores en la comparación entre dos RSN.

Tabla 2. Ejemplo de matriz de salida para la comparación de RSN mediante MEV.

	AGUA	AIRE	ALUMINIO	AMBIENTE	ANIMAL	ARBOL	BASURA	BOSQUE	BOTE	BOTELLA
PROYECTO 1	0.166	0.084		0.371	0.562	1.000	0.636	0.269	0.020	0.020
PROYECTO 2	0.359	0.181	0.139	0.213	0.840	0.965	0.518	0.174		

Una vez que se tiene definida la matriz de salida  $matrizcomp[2,j]$ , se realizan las operaciones de cálculo de similitud para obtener el parecido entre estas dos redes.

## Resultados

Aplicando la metodología definida se ejecutó el programa para la generación de la matriz de comparación entre las cuatro RSN de principiantes generadas; en la Tabla 3 se muestran las matrices de comparación obtenidas entre la RSN de ECOLOGÍA 1 con la RSN de ECOLOGÍA 2, así como la comparación de la RSN de ECOLOGÍA 1 con la RSN de ECOLOGÍA 4.

La Tabla 4 expone los resultados del cálculo de similitud de todas las comparaciones entre las RSN.

**Análisis de resultados de comparación entre RSN**

En este caso de la comparación, los valores de similitud obtenidos son muy parecidos como se puede observar en la Tabla 4, los valores de similitud cosenoidal varían entre 0.913 y 0.954, el valor mínimo de similitud se obtuvo en la comparación entre la RSN ECOLOGÍA 3 y la RSN ECOLOGÍA 4 y el valor máximo de similitud se dio entre la RSN ECOLOGÍA 2 y la RSN ECOLOGÍA 4.

En la Tabla 5 se expone información con respecto a los valores de similitud y la proporción de definidores coincidentes y discrepantes para las comparaciones entre principiantes que obtuvieron la similitud máxima y mínima, como se puede observar, existe un valor muy pequeño de diferencia entre estos resultados (0.041), asimismo se puede identificar que la cantidad y porcentaje de definidores coincidentes es menor en la comparación que obtuvo la mayor similitud (ECOLOGÍA 2-ECOLOGÍA 4), como se había mencionado, en el MEV, el peso que se da a los términos (en este caso el valor normalizado M de los definidores) tiene un impacto en el resultado de la similitud entre RSN, por lo que se puede inferir que las diferencias en los valores de similitud calculados están derivadas de las diferencias existentes en los valores normalizados de M.

Tabla 3. Muestra la matriz de comparación entre la RSN ECOLOGÍA 1 con la RSN ECOLOGÍA 2 y de la RSN ECOLOGÍA 1 con la RSN ECOLOGÍA 4.

	ECO 1	ECO 2		ECO 1	ECO 4
AGUA	0.166	0.359	AGUA	0.166	0.327
AIRE	0.084	0.181	AIRE	0.084	0.235
ALUMINIO		0.139	ALUMINIO		0.057
AMBIENTE	0.371	0.213	AMBIENTE	0.371	0.195
ANIMAL	0.562	0.840	ANIMAL	0.562	0.598
ARBOL	1.000	0.965	ARBOL	1.000	1.000
BASURA	0.636	0.518	BASURA	0.636	0.379
BOSQUE	0.269	0.174	BOSQUE	0.269	0.052
BOTE	0.020		BOTE	0.020	
BOTELLA	0.020		BOTELLA	0.020	0.024
CALLE	0.026		CALLE	0.026	0.014
CARRO	0.381	0.375	CARRO	0.381	0.292
CARTÓN		0.092	CASA	0.135	0.043
CASA	0.135	0.084	CIUDAD		0.054
CONTAMINAC...	0.273	0.306	CONTAMINAC...	0.273	0.373
CUIDADO	0.030	0.055	CUIDADO	0.030	0.059
DESECHO	0.033		DESECHO	0.033	
ECOLOGIA	0.267		ECOLOGIA	0.267	0.030
ECOSISTEMA	0.191	0.149	ECOSISTEMA	0.191	0.224
EDIFICIO	0.305	0.164	EDIFICIO	0.305	0.241
EMPRESA		0.015	ENFERMEDAD	0.071	0.022
ENFERMEDAD	0.071		FABRICA	0.015	0.027
FABRICA	0.015	0.039	FAUNA	0.038	
FAUNA	0.038		FLORA	0.038	0.014
FLORA	0.069		GENTE	0.038	0.004
GENTE	0.038	0.039	GRANDE		0.022
GRANDE		0.020	HUMO	0.125	0.041
HUMO	0.125	0.076	LATA	0.036	
INORGANICO		0.031	LIMPIA	0.072	0.009
LATA	0.036	0.041	MEDIO		0.014
LIMPIA	0.072	0.140	METROPOLI	0.030	
METROPOLI	0.030		NATURAL	0.316	0.432
NATURAL	0.316	0.383	ORGANICO	0.031	
ORGANICO	0.031		OXIGENO	0.087	0.067
OXIGENO	0.087	0.027	PAPEL	0.110	0.114
PAPEL	0.110	0.168	PERSONA	0.171	0.098
PERSONA	0.171	0.154	PLANTA	0.794	0.621
PLANTA	0.794	1.000	PLÁSTICO	0.152	0.121
PLÁSTICO	0.152	0.268	POBLACION		0.057
RECICLAJE	0.125	0.138	RECICLAJE	0.125	0.090
REUSAR	0.278	0.191	REUSAR	0.278	0.130
RIO	0.036		RIO	0.036	
RUIDO		0.074	RUIDO		0.017
SEMILLA	0.020		SEMILLA	0.020	
SMOG	0.063	0.037	SMOG	0.063	0.042
SUCIO	0.112	0.117	SUCIO	0.112	0.065
TALA		0.001	TALA		0.001
TIENDAS	0.001		TIENDAS	0.001	
TIERRA	0.041	0.264	TIERRA	0.041	0.138
TOXICO	0.001		TOXICO	0.001	
TRAFICO	0.081	0.015	TRAFICO	0.081	
VERDE	0.493	0.463	VERDE	0.493	0.205
VIDA	0.049	0.130	VIDA	0.049	0.208
VIDRIO		0.029			

Tabla 4. Resultado de similitud obtenido en la comparación entre las RSN.

COSENOIDAL				
	eco1	eco2	eco3	eco4
eco1	1.000	0.942	0.930	0.927
eco2	0.942	1.000	0.913	0.954
eco3	0.930	0.913	1.000	0.939
eco4	0.927	0.954	0.939	1.000

Tabla 5. Valores relacionados con la similitud entre principiantes.

Comparación	Similitud	Definidores coincidentes		Definidores no coincidentes	
		Cantidad	Porcentaje	Cantidad	Porcentaje
ECOLOGÍA 2- ECOLOGÍA 3	0.913	35	71.43%	14	28.57%
ECOLOGÍA 2- ECOLOGÍA 4	0.954	34	70.83%	14	29.17%

Para comprobar lo anterior se realizó la sumatoria de las diferencias entre los valores normalizados M entre las redes comparadas y se obtuvo que para la comparación entre ECOLOGÍA 2 y ECOLOGÍA 4 el valor calculado fue 2.693 y para ECOLOGÍA 2 y ECOLOGÍA 3 el valor fue 3.340, es decir, para la comparación con la mayor similitud la diferencia entre los valores M es menor.

### Conclusiones

Fue posible adaptar el MEV trasladando el proceso que sigue este modelo hacia la comparación entre Redes Semánticas Naturales, las técnicas de comparación son las indicadas para MEV y el resultado de la similitud sería interpretado como el parecido que existe entre las redes comparadas, en la Figura 5 se representó la analogía para el proceso anterior.

Los resultados de la comparación entre las redes de principiantes son muy altos, sus valores van de 0.913 a 0.954, lo anterior es posible dado que las cuatro RSN fueron generadas con alumnos del mismo perfil, son alumnos de ingeniería de la misma institución y carrera, entre las redes hubo una coincidencia entre definidores también muy parecida que va del 70.83% al 71.43%.

Por lo anterior se considera que sí es posible comparar Redes Semánticas Naturales mediante el Modelo de Espacio Vectorial el cual es un modelo clásico de Recuperación de Información, esta comparación se puede realizar entre RSN de diferente cantidad de participantes ya que utiliza valores normalizados.

Se puede concluir que el Modelo de Espacio Vectorial adaptado a la comparación de RSN, es un método que se puede aplicar para realizar la comparación entre Redes Semánticas Naturales ya que se pudo comprobar que los resultados de similitud obtenidos reflejan el parecido que existe entre las RSN. La ventaja de este método es que puede ser aplicado entre redes de diferente tamaño ya que los valores son normalizados.

## Referencias

- Arévalo Aguilera, H. E. (2010). El uso de las redes semánticas naturales en las representaciones sociales de la responsabilidad. *Revista Internacional de Psicología*, 11(2).
- Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: ACM Press.
- Collipal, E., Cabalín, D., Vargas, J., & Silva, H. (2004). Conceptualización semántica del término anatomía humana por los estudiantes de medicina. *International Journal of Morphology*, 22(3), 185-188.
- De la Cruz Flores, G., & Abreu Hernández, L. F. (2012). Atributos de tutores de posgrado por campo disciplinario. La perspectiva de estudiantes de la Universidad Nacional Autónoma de México. *Perfiles educativos*, 34(138), 10-27.
- Figuroa, J. G., González, E. G., & Solís, V. (1981). Una aproximación al problema del significado: Las redes semánticas. *Revista Latinoamericana de Psicología*, 13(3), pp:447-458.
- García, C. B., & Jiménez Vidal, S. (1996). Redes semánticas de los conceptos de presión y flotación en estudiantes de bachillerato. *Revista Mexicana de Investigación Educativa*, 1(2), 343-361.
- Hanani, U., Shapira, B., & Shoval, P. (2001). Information filtering: Overview of Issues, Research and Systems. *User Modeling and User-Adapted Interaction*, 11(3), 203-259.
- Ingeniería Cognitiva. (2 de Agosto de 2014). *RSNweb*. (FIME UANL) Recuperado el 6 de Marzo de 2015, de <http://ingenieriacognitiva.com/developer/redes/registro/>
- Muflikhah, L., & Baharudin, B. (2009). Document clustering using concept space and cosine similarity measurement. *6th International Conference on Computer Technology and Development*. Hongkong: IEEE .
- Murillo Estepa, P., & Becerra Peña, S. (2009). Las percepciones del clima escolar por directivos, docentes y alumnado mediante el empleo de "redes semánticas naturales": su importancia en la gestión de los centros educativos. *Revista de Educación*, 350, 375-379.
- Padilla Montemayor, V. M., & Rodríguez Nieto, M. C. (2011). Evaluación de la representación del conocimiento de un curso . *Congreso Latinoamericano Ciencias de la Educación*. Cochabamba.
- Reyes-Lagunes, I. (1993). Las Redes Semánticas Naturales: Su Conceptualización y su utilización en la construcción de instrumentos. *Revista de Psicología Social y Personalidad*, 9(1), 83-87.
- Salton, G. (1989). *Automatic text processing: the transformation, analysis, and retrieval of information by computer*. Boston: Addison-Wesley Longman Publishing Co.
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.
- Salton, G., & McGuill, M. J. (1986). *Introduction to modern information retrieval*. New York: Mc Graw Hill.
- Salton, G., Wong, A., & Yang, C. S. (1975). A Vector Space Model for Information Retrieval. *Communications of the ACM*, 18(11), 613-620.
- Singhal, A. (2001). Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.*, 24(4), 35-43.
- Trejo-Lucero, H., Camacho-Beiza, R., & Herrera-Villalobos, J. (2011). Significado semántico de «lactancia materna» y «lactancia artificial» en mujeres y hombres. *Revista Mexicana de Pediatría*, 78(1), 10-15.
- Valdez M, J. L. (2000). *Las redes semánticas naturales, usos y aplicaciones en psicología social*. México: Universidad Autónoma del Estado de México.
- Van Rijsbergen, C. (1979). *Information Retrieval*. Newton: Butterworth-Heinemann