



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

UNA PROPUESTA DE PUNTUACIÓN DE
COMPLEJIDAD LEGAL: CASO LEY DEL IMPUESTO
SOBRE LA RENTA

T E S I S

PARA OBTENER EL TÍTULO DE:

ACTUARIO

P R E S E N T A :

OSCAR URIEL PÉREZ SALAZAR

TUTOR

ACT. EDGAR DÍAZ ORDOÑEZ



CIUDAD UNIVERSITARIA, Cd. Mx., 2022



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mi perrito Lemmy.

Agradecimientos

Agradezco a mis padres por brindarme su apoyo durante todo este tiempo, así como el ayudarme a corregir aspectos de mi ser día con día.

Agradezco a mi hermano, una persona que siempre me ha apoyado y que ha sido un gran ejemplo del cual aprendí y sigo aprendiendo bastante.

Agradezco a mis perritos, Lemmy y Rommel, por brindarme su compañía y cariño. Sobre todo a Lemmy, quien es uno de los seres más fieles y atentos que he tenido a mi lado.

Agradezco a mis amistades, las cuales me han aceptado tal como he sido a lo largo de nuestra amistad y me han ayudado a sobrellevar esto que denominamos vida. Algunas amistades que me gustaría destacar son las de Ivan, Araujo, Ivar, Melina, Edith, Mitzi, Geraldine y Vanesa.

Agradezco a mi asesor de tesis y mentor, Edgar Díaz Ordoñez, quien me dio la oportunidad de no sólo realizar este trabajo bajo su tutela, si no también de incursionar en temas que me han resultado interesantes y que han contribuido a mejorar mi formación profesional.

Agradezco a los sinodales por su tiempo y dedicación para indicarme cómo mejorar el presente trabajo.

En fin, a todos ustedes que me han ayudado a ser un mejor hijo, hermano, compañero, amigo, estudiante, una mejor persona que espera ser un poco más amable que de lo que era el día anterior.

Resumen

En el presente trabajo se analizaron algunas características que se pueden obtener de las leyes que integran el marco jurídico mexicano. Algunos ejemplos son las palabras que contienen dichas leyes, las referencias hacía otras leyes u otras partes dentro de la misma ley o el número de fracciones que contiene un artículo, entre otras. Se contemplaron 3 grandes rubros para agrupar dichas características:

- Estructura
- Lenguaje
- Interdependencia

Con base en dichas características, se construyeron puntuaciones de complejidad y con ello se generó una puntuación para cada uno de los rubros mencionados.

Finalmente, se construyó la puntuación de complejidad legal con base en las puntuaciones de complejidad de cada rubro. Lo anterior se aplicó a una ley mexicana en particular, la Ley del Impuesto sobre la Renta. Se analizaron los resultados obtenidos en conjunto con otras cifras relacionadas con dicha ley tales como la recaudación obtenida con el Impuesto sobre la Renta, el universo de contribuyentes dados de alta en el Servicio de Atención Tributaria y los servicios realizados por la Procuraduría de la Defensa del Contribuyente.

Índice general

Agradecimientos	II
Resumen	III
1. Introducción	1
1.1. Contexto	1
1.2. Estado del arte	1
1.2.1. Complejidad legal	1
1.2.2. Estructura de la Ley del Impuesto sobre la Renta (LISR) . . .	6
1.3. La Ley del Impuesto sobre la Renta (LISR)	7
1.4. Objetivo	8
1.5. Estructura	9
2. Métricas y puntuaciones de la complejidad - Teoría	10
2.1. Métricas	10
2.1.1. Estructura	10
2.1.1.1. Densidad de la gráfica	11
2.1.1.2. Grado de un nodo	12
2.1.1.3. Profundidad	13
2.1.2. Lenguaje	14
2.1.2.1. Entropía	15
2.1.3. Interdependencia	16
2.1.3.1. Hyperlink-Induced Topic Search	17
2.2. Puntuación de complejidad con base en las métricas	21
2.2.1. Puntuación de complejidad de la estructura de la ley	22
2.2.1.1. Puntuación de complejidad con base en la densidad de la gráfica	22
2.2.1.2. Puntuación de complejidad con base en el grado de salida de los nodos	23
2.2.1.3. Puntuación de complejidad con base en la profundidad de los nodos	24
2.2.2. Puntuación de complejidad del lenguaje en la ley	26
2.2.2.1. Puntuación de complejidad con base en la entropía de las palabras	26

2.2.3.	Puntuación de complejidad con base en las interdependencias en la ley	27
2.2.3.1.	Puntuación de complejidad con base en las puntuaciones Hub y Authority	27
3.	Implementación de las métricas en la LISR	29
3.1.	Métricas en la LISR	29
3.1.1.	Estructura	29
3.1.1.1.	Nodos en las gráficas	29
3.1.1.2.	Densidad de las gráficas	33
3.1.1.3.	Grado de salida de los nodos	34
3.1.1.4.	Profundidad de los nodos	40
3.1.2.	Lenguaje	42
3.1.2.1.	Palabras en la LISR	42
3.1.2.2.	Entropía	46
3.1.3.	Interdependencias	47
3.1.3.1.	Referencias internas	47
3.1.3.2.	Referencias externas	50
3.2.	Puntuaciones de complejidad con base en las métricas	52
3.2.1.	Puntuación de complejidad con base en la densidad de la gráfica de la LISR	53
3.2.2.	Puntuación de complejidad con base en el grado de salida de los nodos de la LISR	53
3.2.3.	Puntuación de complejidad con base en la profundidad de los nodos hoja de la LISR	55
3.2.4.	Puntuación de complejidad con base en la entropía de las palabras	57
3.2.5.	Puntuación de complejidad con base en las puntuaciones de Hub y Authority obtenidas con el algoritmo HITS	57
3.2.6.	Puntuación de complejidad de la estructura de la LISR	63
3.2.7.	Puntuación de complejidad del lenguaje de la LISR	64
3.2.8.	Puntuación de complejidad de las interdependencias en la LISR	65
3.2.9.	Puntuación de complejidad legal de la LISR	65
3.2.10.	La puntuación de complejidad legal y la recaudación obtenida con el ISR	69
3.2.11.	La complejidad legal y algunos servicios proporcionados por la PRODECON	71
4.	Conclusiones	74
4.1.	Siguientes pasos	75
	Anexos	76
A.	Definiciones	77
A.1.	Gráfica	77
A.2.	Nodos en una gráfica	77

A.3. Aristas en una gráfica	77
A.4. Media geométrica	77
A.5. Media ponderada	77
A.6. Correlación de Pearson	78
A.7. Correlación de Spearman	78
Apéndice	78
A. Proceso de ingesta de datos	79
B. Procesamiento de datos	80
C. Puntuaciones de complejidad con base al año y el título	81
C.1. Puntuación de complejidad de la estructura de la LISR	81
C.2. Puntuación de complejidad del lenguaje de la LISR	82
C.3. Puntuación de complejidad de las interdependencias en la LISR	83
C.4. Puntuación de complejidad legal de la LISR	84

Índice de figuras

1.1. Representación de la estructura de la LISR como un grafo	7
2.1. Ejemplo 1 de densidad en un grafo	11
2.2. Ejemplo 2 de densidad en un grafo en un grafo	11
2.3. Ejemplo 1 de grado de salida en un grafo	12
2.4. Ejemplo 2 de grado de salida en un grafo	12
2.5. Ejemplo 3 de grado de salida en un grafo	12
2.6. Ejemplo 4 de grado de salida en un grafo	13
2.7. Ejemplo de profundidad en un grafo	14
2.8. Ejemplo de puntuación Hub	17
2.9. Ejemplo 1 de puntuación de complejidad con base en el grado de salida	24
2.10. Ejemplo 2 de puntuación de complejidad con base en el grado de salida	24
2.11. Ejemplo 1 de puntuación de complejidad con base en la profundidad .	25
2.12. Ejemplo 2 de puntuación de complejidad con base en la profundidad .	26
3.1. Número de nodos por año en la LISR	30
3.2. Número de títulos por año en la LISR	30
3.3. Número de capítulos por año en la LISR	30
3.4. Número de secciones por año en la LISR	31
3.5. Número de artículos por año en la LISR	31
3.6. Número de fracciones por año en la LISR	31
3.7. Número de incisos por año en la LISR	31
3.8. Número de numerales por año en la LISR	31
3.9. Número de sub incisos por año en la LISR	31
3.10. Número de apartados por año en la LISR	32
3.11. Variación porcentual por tipo de nodo	32
3.12. Cambios en la frecuencia por tipo de nodo	33
3.13. Densidad obtenida por año en la LISR	33
3.14. Distribución del grado de salida considerando todos los nodos	34
3.15. Distribución representativa de los nodos con grado de salida distinto de cero	35
3.16. Porcentaje de nodos con grado de salida distinto de cero cuyo grado de salida no cambió	35
3.17. Cambios en el grado de salida de los nodos durante el periodo analizado	36
3.18. Evolución de la distribución del grado de salida de los nodos	36

3.19. Evolución de las gráficas de cajas del grado de salida por año	37
3.20. Estadísticos de forma de las distribuciones del grado de salida	37
3.21. Grados de salida por tipo de nodo en el 2008	38
3.22. Grados de salida por tipo de nodo en el 2009	38
3.23. Grados de salida por tipo de nodo en el 2010	38
3.24. Grados de salida por tipo de nodo en el 2011	38
3.25. Grados de salida por tipo de nodo en el 2012	38
3.26. Grados de salida por tipo de nodo en el 2013	38
3.27. Grados de salida por tipo de nodo en el 2014	39
3.28. Grados de salida por tipo de nodo en el 2015	39
3.29. Grados de salida por tipo de nodo en el 2016	39
3.30. Grados de salida por tipo de nodo en el 2017	39
3.31. Evolución de la distribución de la profundidad	40
3.32. Gráficas de cajas de la profundidad por año	40
3.33. Cambios en la frecuencia por tipo de nodo	41
3.34. Cambios en la frecuencia por tipo de nodo	41
3.35. Estadísticos de forma de las distribuciones de densidad	42
3.36. Número total de palabras por año	43
3.37. Número total de palabras distintas por año	43
3.38. Variaciones porcentuales por año	43
3.39. Razón entre palabras distintas y palabras totales por año con todas las palabras	44
3.40. Número total de palabras por año	44
3.41. Número total de palabras distintas por año	44
3.42. Razón entre palabras distintas y palabras totales por año usando una lista de palabras vacías	44
3.43. Cambios en el número de palabras totales	45
3.44. Cambios en el número de palabras distintas totales	45
3.45. Términos compuestos más usados durante el periodo analizado	46
3.46. Entropía con todo el universo de palabras	47
3.47. Entropía donde sólo se contemplan las palabras que no formen parte de la lista de palabras vacías.	47
3.48. Referencias internas	47
3.49. Referencias internas en nodos de tipo artículo	48
3.50. Referencias internas en nodos de tipo fracción	48
3.51. Referencias internas en nodos de tipo inciso	48
3.52. Referencias internas en nodos de tipo numeral	48
3.53. Referencias internas que apuntan a nodos de tipo título	49
3.54. Referencias internas que apuntan a nodos de tipo capítulo	49
3.55. Referencias internas que apuntan a nodos de tipo sección	49
3.56. Referencias internas que apuntan a nodos de tipo artículo	49
3.57. Referencias internas que apuntan a nodos de tipo fracción	49
3.58. Referencias internas que apuntan a nodos de tipo inciso	49
3.59. Referencias internas que apuntan a nodos de tipo numeral	50
3.60. Gráfica representativa de referencias internas de 2008 a 2013	50

3.61. Gráfica representativa de referencias internas de 20014 a 2017	50
3.62. Referencias externas	51
3.63. Referencias externas en nodos de tipo artículo	51
3.64. Referencias externas en nodos de tipo fracción	51
3.65. Referencias externas en nodos de tipo inciso	52
3.66. Referencias externas en nodos de tipo numeral	52
3.67. Gráfica representativa de referencias externas	52
3.68. Puntuación de complejidad con base en la densidad	53
3.69. Puntuación de complejidad con base en el grado de salida	54
3.70. Puntuación de complejidad redondeada con base en el grado de salida	55
3.71. Puntuación de complejidad con base en la profundidad	56
3.72. Puntuación de complejidad redondeada con base en la profundidad	56
3.73. Puntuación de complejidad con base en la entropía	57
3.74. Puntuación Hub en las referencias internas	58
3.75. Puntuación Hub de artículos	58
3.76. Puntuación Hub de fracciones	58
3.77. Puntuación Hub de incisos	59
3.78. Puntuación Hub de numerales	59
3.79. Puntuación Authority en las referencias internas	59
3.80. Puntuación Authority de títulos	60
3.81. Puntuación Authority de capítulos	60
3.82. Puntuación Authority de secciones	60
3.83. Puntuación Authority de artículos	60
3.84. Puntuación Authority de fracciones	61
3.85. Puntuación Authority de incisos	61
3.86. Puntuación Authority de numerales	61
3.87. Puntuación de Hub en las referencias externas	62
3.88. Puntuación Hub de fracciones	62
3.89. Puntuación Hub de incisos	62
3.90. Puntuación Hub de fracciones	63
3.91. Puntuación Hub de incisos	63
3.92. Puntuación de complejidad de estructura de la LISR	63
3.93. Correlaciones entre la puntuación de estructura y las puntuación que la conforman	64
3.94. Puntuación de complejidad del lenguaje en la LISR	64
3.95. Puntuación de complejidad de interdependencias en la LISR	65
3.96. Puntuación de complejidad	66
3.97. Puntuación de complejidad legal de 2008 y 2009	66
3.98. Puntuación de complejidad legal de 2010 y 2011	66
3.99. Puntuación de complejidad legal de 2012 y 2013	67
3.100Puntuación de complejidad legal de 2014 y 2015	67
3.101Cambio de puntuaciones preponderantes	68
3.102Puntuación de complejidad legal de 2014 y 2015	68
3.103Puntuación de complejidad legal de 2016 y 2017	68

3.104	Correlaciones entre las puntuaciones de complejidad y la puntuación de complejidad legal	69
3.105	Correlaciones entre las puntuaciones de complejidad y la puntuación de complejidad legal	69
3.106	Recaudación obtenida con el ISR.	70
3.107	Universo de contribuyentes por año	70
3.108	Recaudación por contribuyentes	71
3.109	Correlaciones entre la puntuación de complejidad legal y el monto recaudado así como con el universo de contribuyentes	71
3.110	Servicios realizados por la PRODECON	72
3.111	Razón entre los asuntos atendidos por la PRODECON y el universo de contribuyentes	73
3.112	Correlación entre la puntuación de complejidad y los servicios prestados por la PRODECON	73
C.1.	Puntuaciones de complejidad de la estructura de los títulos de la LISR por año	82
C.2.	Puntuaciones de complejidad del lenguaje en los títulos de la LISR por año	83
C.3.	Puntuaciones de complejidad de las interdependencias en los títulos de la LISR por año	84
C.4.	Puntuaciones de complejidad legal de los títulos de la LISR por año	84

Índice de tablas

1.1. Métricas calculadas por Waltl y Matthes	5
2.1. Métricas contempladas en el presente trabajo	10

Capítulo 1

Introducción

1.1. Contexto

Los ciudadanos mexicanos en su labor diaria eventualmente tendrán que lidiar con algún documento normativo, como el reglamento de tránsito, la ley de algún impuesto como el impuesto sobre la renta o el código civil cuando se requiera redactar el testamento. Si bien existen personas especializadas en dichos documentos las cuales pueden brindar sus servicios para facilitar la comprensión de dichos documentos, no todos los ciudadanos tienen los recursos para poder pagar dichos servicios o su confianza es ciega en tales personas lo que los hace vulnerables a sufrir extorsiones o abusos. Por lo anterior no es de extrañar que algunos entes se esfuercen en difundir y explicar el contenido de los documentos normativos a todos los ciudadanos que puedan alcanzar. No obstante, ¿es sólo el lenguaje de dichos documentos lo que dificulta la comprensión de estos o existen más características que contribuyen a aumentar su comprensión?

En el presente trabajo se intentará construir una puntuación de complejidad legal con base en algunas de las características que comparten las leyes que integran el marco jurídico mexicano. Una vez construida dicha puntuación de complejidad se implementará y se analizará en la Ley del Impuesto Sobre la Renta (LISR), ley de gran importancia en la recaudación de recursos para la nación.

Cabe señalar que también es de interés detectar qué características contribuyen en mayor medida a elevar la complejidad de la LISR, así como la evolución de estas.

1.2. Estado del arte

1.2.1. Complejidad legal

En 1992, Schuck [9] en su artículo titulado *Legal complexity: some causes, consequences, and cures* define a un sistema legal complejo como:

El grado en que sus reglas, procesos, instituciones y cultura poseen cuatro características: densidad, tecnicismos, diferenciación e indeterminación o incertidumbre.

Dichas características las define de la siguiente manera:

- Densidad: la cantidad de aspectos que busca regular una ley.
- Tecnicismos: cuando se requiere determinado expertís para entender y aplicar la ley dada la terminología especializada que se emplea.
- Diferenciación: un sistema legal es institucionalmente diferenciada en la medida del número de estructuras que recurren a diferentes fuentes de derecho, poseen diferentes tipos de inteligencia organizacional y emplea diferentes procesos para crear, elaborar y aplicar las leyes.
- Indeterminación o incertidumbre: el distinto significado que pueden tener las leyes, el cual está en función del contexto en el que se aplican.

Años más tarde, en 1997, Kades [5] en su trabajo titulado *The Laws of Complexity & the Complexity of Laws: The Implications of Computational Complexity Theory for the Law* aborda la complejidad de las leyes con base en la definición existente en la Teoría de la Complejidad Computacional (CCT por sus siglas en inglés), pasando así a usar una rama de conocimiento ajena al Derecho. Afirma que la ley fiscal es un claro ejemplo de complejidad legal y a lo largo de su trabajo menciona aspectos que influyen en la complejidad de la ley. Dichos aspectos son:

- Definiciones de justicia multifacéticas: aquí menciona cómo los intereses de las personas y su definición de justicia pasan a modificar la ley y con ello su complejidad.
- Indeterminación: al igual que Schuck, se refiere a lo ambiguo que puede resultar la ley. Aquí hace mención a una posible solución a este aspecto: la creación de más reglas. No obstante, hacer lo anterior podría aumentar la complejidad de la ley. A esta relación entre la creación de más reglas y la complejidad es denominada como “Complejidad dinámica” por algunos autores.
- Dificultades cognitivas: aborda la relación entre el lenguaje con el que están escritas las reglas y las personas que las producen. Retoma la característica definida como “Tecnicismos” propuesta por Schuck ya que aquí aborda el conocimiento mínimo que debe tener una persona para comprender y aplicar la ley.
- Citas: cuando en una ley o regla se citan otras leyes o reglas las cuales deben abordarse para poder aplicar correctamente la ley.

En cuanto a la complejidad propuesta por la CCT, esta la aborda considerando el tiempo necesario que le tomaría a la computadora cumplir con el objetivo y el riesgo de caer en un gran número de citas o que estas sean circulares. Lo anterior es una propuesta para medir la complejidad con cálculos capaces de replicar y comparar.

Siguiendo la idea de poder medir la complejidad legal mediante alguna medida, en 2007 Bourcier y Mazzega[2] plantearon en su trabajo *Toward measures of complexity*

in legal systems dos posibles clases de medidas de complejidad en sistemas legales: medidas basadas en la estructura y medidas basadas en el contenido. Tras analizar la estructura de algunas leyes, plantearon la idea de que las leyes tienen una estructura jerárquica parecida a un árbol:

Algunos textos legales están organizados en un conjunto de códigos. Por ejemplo, el código Francés del Ambiente esta dividido en siete libros. Cada libro esta dividido en títulos y los títulos en capítulos. Los artículos involucrados en la administración y gobernanza del ambiente se encuentran en diferentes capítulos. Estos a su vez están compuestos de párrafos. No obstante, la organización jerarquía similar a un árbol de este código está conectada a otras estructuras similares a un árbol.

Dado lo anterior, propusieron que algunos conceptos de teoría de gráficas tales como el grado y la centralidad de un nodo podrían ser de interés para describir un sistema legal. Por otro lado, retoman la importancia de las conexiones existentes entre las leyes y dentro de ellas mismas ya que si alguna de las partes conectadas sufre alguna modificación, esa modificación puede influir en todas las partes conectadas a ella.

Años más tarde, en 2010, Daniel Katz y Michael Bommarito[1] analizaron The United States Code¹ en su trabajo titulado *A mathematical approach to the study of the United States Code*. Por un lado, analizaron la estructura de dicho código considerando como un nodo a cada división que lo compone, es decir, a cada título, subtítulo, capítulo, parte, subparte, sección, subsección, párrafo, subpárrafo, clausula o subclausula. Con lo anterior, graficaron la estructura jerárquica del código. Además, también consideraron las referencias existentes en el código.

Por otro lado, analizaron el contenido enfocándose en la distribución del número de palabras que componen a dicho código. No obstante, se dieron cuenta de que sólo las secciones, subsecciones, párrafos, subpárrafos, clausulas y subclausulas contienen texto legal substancial por lo que sólo en dichas partes del código realizaron el conteo de palabras. Además, con dichos conteos calcularon la entropía, una medida desarrollada por Shanon[10] en el campo de la teoría de la información. Esto lo realizaron con el fin de capturar la variabilidad de las palabras que componen a las secciones.

Katz y Bommarito continuaron su análisis del Código de los Estados Unidos de Norte América y en el 2013 publican *Measuring the Complexity of the Law : The United States Code*. En dicho trabajo, analizan la estructura del código con base en las siguientes variables:

- El número de vértices que componen la red.
- La profundidad promedio de los nodos.
- La relación entre la profundidad promedio y el número de elementos que integran los títulos.

¹The United States Code es una consolidación y codificación dividida por temas de las leyes generales y permanentes de Estados Unidos de América. Para más información, consultar la página de internet: <https://uscode.house.gov/>

Por otro lado, analizan el lenguaje del código mediante las siguientes características de las palabras que lo integran:

- El número de palabras.
- La longitud promedio de las palabras.
- La entropía de las palabras.

Respecto a las citas, estas las englobaron como interdependencias y las dividieron en dos tipos:

- Interdependencia entre los títulos: aquellas citas que generan una dependencia entre diferentes títulos del código. Con dichas citas definen la idea de importación de conceptos, cuando A cita a B, y de exportación de conceptos, cuando A es citado por B. Con lo anterior, categorizan a los títulos como exportadores o importadores con base en el número del tipo de citas que tienen:

Si un título es citado por otros títulos más de lo que este cita otros títulos, entonces es un exportador y de igual manera es un importador si tiene un comportamiento inverso al del importador.

El valor de esta variable es el número del tipo de citas con el que fue categorizado el título.

- Interdependencia dentro del título: aquellas citas que generan una dependencia entre los elementos que integran a un mismo título. Con estas citas calculan la proporción de citas dentro del título que hacen referencia a elementos dentro del mismo título.

Finalmente, calcularon una métrica compuesta por las propiedades de la estructura, del lenguaje y de las interdependencias. Cabe señalar que dicha métrica la calcularon a nivel título. Dado que las propiedades no comparten la escala, Katz y Bommarito calcularon la métrica compuesta de dos formas: con las variables normalizadas y sin normalizar.

Un año más tarde, Bernhard Walzl y Florian Matthes[11] hicieron un análisis semejante de las leyes Alemanas en su trabajo titulado *Towards Measures of Complexity: Applying Structural and Linguistic Metrics to German Laws* ya que calcularon métricas estructurales y lingüísticas de dichas leyes y, al igual que Katz Y Bommarito, lo hicieron con el fin de medir la complejidad. Dichas métricas se pueden observar en la figura 1.2.1:

Nombre	Abreviación	Complejidad indicada
Conteo de párrafos	#§	Lingüística y Estructural
Conteo de oraciones	#S	Lingüística y Estructural
Conteo de palabras	\$W	Lingüística y Estructural
Profundidad estructural	D	Estructural
Número de referencias internas	INT	Estructural
Número de referencias externas	EXT	Estructural
Variabilidad del vocabulario	V	Lingüística
Indeterminación	I	Lingüística
Legibilidad (Flesch-Reading-Easy)	FRE	Lingüística

Tabla 1.1: Métricas calculadas por Walth y Matthes[11] para medir la complejidad de las leyes Alemanas.

La mayoría de las métricas contempladas ya habían sido abordadas por trabajos previos de otros autores. No obstante, hay dos métricas que resultaron novedosas: la indeterminación y la legibilidad.

La indeterminación ya había sido mencionada previamente por autores tales como Schuck[9] y Kades[5]. Sin embargo, no se había medido de alguna forma. Para medirla, Walth y Matthes crearon una lista de palabras que, bajo su criterio, son ambiguas y obtuvieron la frecuencia de dichas palabras en cada una de las leyes que analizaron. Con dichas frecuencias calcularon la frecuencia relativa por cada 1000 palabras de este conjunto de palabras consideradas como ambiguas.

Respecto a la legibilidad, la midieron con base a la prueba de legibilidad de Flesch[4]. Esta prueba depende de la longitud promedio de las oraciones, del número promedio de sílabas por palabra y de algunas constantes cuyo valor dependen del idioma en el que esté el texto a analizar. A mayor valor mayor será la facilidad para leer el texto. Originalmente, esta prueba se construyó para medir la legibilidad de los textos en inglés.

Un año más tarde, Rick Krever y Peter Mellow en su trabajo *Where's the Complexity in Tax Law?*[7] proponen cuatro indicadores para medir la complejidad de una ley de impuestos.

El primer indicador propuesto fue el costo de obtener un “private ruling” o fallo privado por parte de una autoridad legal a un contribuyente. Estos fallos privados se generan cuando un contribuyente, al tener incertidumbre sobre sus responsabilidades fiscales, pide ayuda a un profesional para obtener la interpretación de una autoridad fiscal y con ello cumplir sus obligaciones fiscales con base en la interpretación dada por la autoridad fiscal. Como la información de estos costos es privada, no fue posible usarlo como un indicador objetivo durante su trabajo.

El segundo indicador propuesto fue denominado como “interpretation decisions”. Con esto se refieren a las preguntas comunes que tienden a realizar los contribuyentes para poder cumplir sus obligaciones fiscales. Dichas preguntas las separaron por temas para ubicar los temas fiscales más complejos con base en este indicador.

Por otro lado, el tercer indicador propuesto son los denominados “public ruling” o

fallos públicos emitidos por autoridades fiscales. Estos los dividen en dos:

- “Rulings”: estos son fallos que abarcan varios aspectos.
- “Determinations”: son fallos enfocados en determinados aspectos.

Al igual que con los “private rulings”, los separaron por temas para ubicar los temas fiscales más complejos con base en este indicador.

Por último, el cuarto indicador consiste en las apelaciones de evaluación a tribunales y cortes. Estas apelaciones las usaron para medir la incertidumbre que llegan a tener los contribuyentes dado que dichas apelaciones conllevan un costo el cual los contribuyentes están dispuestos a pagar si creen que existe una interpretación alternativa de la ley que les beneficiará. Al igual que con los indicadores anteriores, los separaron por temas para ubicar los temas fiscales más complejos con base en este indicador.

Los indicadores propuestos por estos autores se enfocan en las consultas realizadas por los contribuyentes y el material generado por las autoridades fiscales para facilitar la comprensión de la ley en vez de las características propias de la ley; el cual es un enfoque que no se había abordado con tanto esfuerzo en los trabajos anteriores.

1.2.2. Estructura de la Ley del Impuesto sobre la Renta (LISR)

Actualmente, la estructura normativa de las leyes que integran el marco jurídico mexicano está comprendida de diversos elementos los cuales Eliseo Muro Ruiz define en su libro *Algunos elementos de técnica legislativa*[8] de la siguiente manera:

- Títulos: se emplean como una de las principales divisiones de los cuerpos legales, pues se refieren a la partición más general de una ley. Generalmente, el título primero consagra las disposiciones generales, y los siguientes aluden instituciones específicas.
- Capítulos: dividen al título e indican las partes de una legislación.
- Secciones: son una subdivisión del capítulo que establece los límites del objeto de la norma.
- Artículos: se enumeran consecutivamente; constituyen una división elemental de los textos normativos; condensan en una o varias frases las especificaciones y los supuestos de la función que cumplen dentro del esquema legal.
- Fracciones: explican y enumeran las autoridades, atributos, obligaciones y facultades que en cada artículo se otorguen. Constituye uno de los medios más usuales para subdividir el o los párrafos que integran cada precepto.
- Inciso: constituyen un desglose de la fracción. Cada inciso debe referirse al texto introductorio del artículo y presentar la misma construcción gramatical. Deben expresar en conjunto con el texto referido, un concepto completo.

Adicionalmente, pueden existir numerales, subincisos y apartados, los cuales son la última distribución de la estructura de las leyes. Al considerar el análisis realizado por los autores previamente mencionados respecto a la estructura de las leyes, la leyes que integran el marco jurídico mexicano presentan una estructura jerárquica dado que cada nivel se divide en subniveles los cuales, a su vez, también pueden dividirse en subniveles. Además, el contenido de cada división depende del nivel del cual se dividió. Cabe señalar que cada subnivel pertenece a un único nivel inicial. En la figura 1.1 se ejemplifica la estructura de una ley que integra el marco jurídico mexicano, la ley del impuesto sobre la renta:

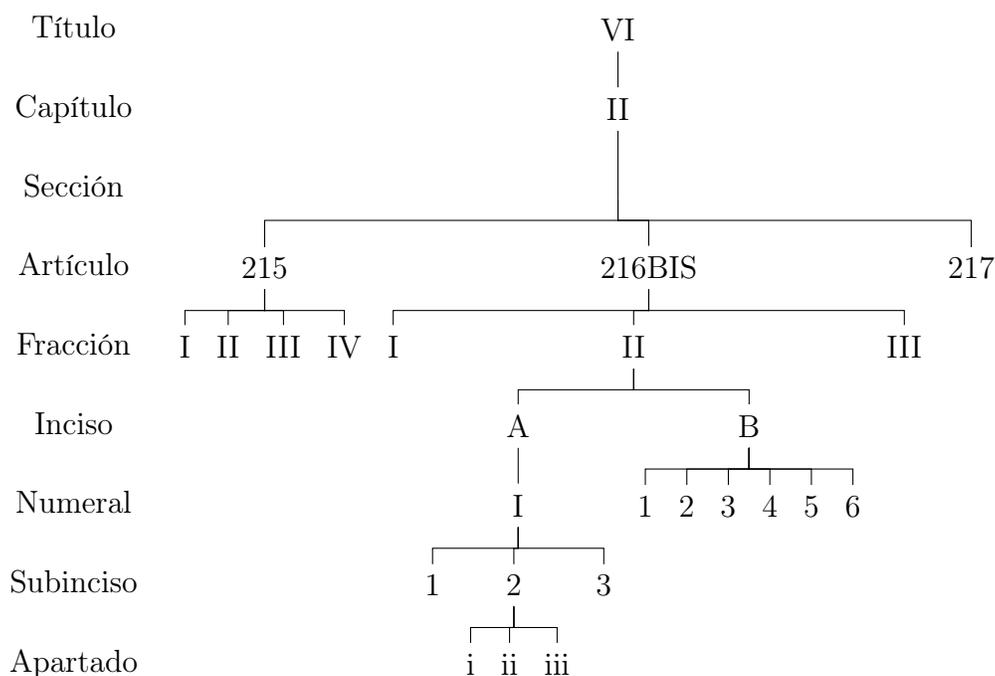


Figura 1.1: Representación gráfica de la estructura de la LISR

Como se muestra en el ejemplo anterior, los arcos entre los nodos pueden omitir subniveles, como es el caso del Capítulo II que omite el subnivel "Sección".

Cabe señalar la existencia de otro tipo de artículos, los artículos transitorios. Estos artículos tienen una vigencia temporal o momentánea y su carácter es secundario dado que complementan a los artículos principales. Por lo anterior, los artículos transitorios no serán considerados en el análisis del presente trabajo.

1.3. La Ley del Impuesto sobre la Renta (LISR)

El Servicio de Administración Tributaria (SAT) define el Impuesto sobre la Renta (ISR) como:

Contribución que grava los ingresos de las personas físicas o morales residentes en el país, así como de las personas residentes en el extranjero por los ingresos atribuibles a sus establecimientos permanentes ubicados en territorio nacional o aquéllos que proceden de fuente de riqueza ubicada en el país.²

Dicho impuesto está regulado por la Ley del Impuesto Sobre la Renta (LISR). Este impuesto tiene antecedentes en México desde inicios del siglo XX, los cuales los menciona Ernesto Zavala[12] en su artículo titulado *Trayectoria del impuesto sobre la renta en México*. El impuesto más remoto que se menciona es el impuesto llamado *Impuesto Centenario*, establecido en la *Ley del Centenario* en 1921, el cual es descrito como:

Impuesto extraordinario y gravaba los ingresos derivados del comercio, de la industria, del ejercicio de una profesión liberal, artística o innominada, del trabajo a sueldo o salario, de la colocación de dinero o valores a crédito, participación o dividendo.

Cuatro años más tarde, en 1925, se expidió la nueva *Ley del Impuesto sobre la Renta*, siendo esta la ley más remota que tiene el nombre de LISR. Durante los años posteriores se crearon nuevas leyes aparte de la LISR, como la *Ley sobre Exportación de Capitales* o la *Ley del Impuesto sobre Utilidades Excedentes* las cuales contemplaban aspectos específicos de los ingresos de las personas y de las empresas. El contenido de dichas leyes pasaría a integrarse en las posteriores versiones de la LISR, como la LISR de 1941 o la LISR de 1953. Así, el nombre de la LISR no cambió y se ha ido modificando con el paso del tiempo hasta la que actualmente está vigente, la cual es la LISR publicada en el Diario Oficial de la Federación (DOF) el 11 de Diciembre de 2013. Dicha LISR ha sido reformada siendo la reforma del 23 de Abril de 2021 la última reforma publicada en el DOF.

1.4. Objetivo

En el presente trabajo se intentará construir una puntuación de complejidad legal con base en algunas de las características que comparten las leyes que integran el marco jurídico mexicano. Una vez construida dicha puntuación de complejidad, se implementará y se analizará en la LISR, hecho esto se atenderán los siguientes puntos:

- ¿La puntuación de complejidad logró detectar algún efecto cuyo origen sea un evento en el marco jurídico Mexicano que repercutiera en la LISR?
- ¿La puntuación de complejidad se puede emplear para estimar la recaudación obtenida con base en el ISR?

²Definición obtenida del glosario elaborado por el SAT ubicado en la página de internet: http://www2.sat.gob.mx/sitio_internet/informe_tributario/itg2014t2/glosario.pdf

- ¿La puntuación de complejidad implementada en la LISR puede emplearse como indicador para prever un aumento o disminución en el número de consultas relativas a la LISR atendidas por parte de organismos que resuelven dudas, quejas o aclaraciones relativas a la LISR?

Cabe señalar que también es de interés detectar qué características contribuyen en mayor medida a elevar la complejidad de la LISR así como la evolución de estas ya que esto permitiría detectar cómo contribuyen a modificar la complejidad las reformas.

1.5. Estructura

La estructura del presente trabajo se conforma de 4 capítulos y un anexo, los cuales se desarrollaron para atender los objetivos previamente planteados.

En el primer capítulo se menciona el trabajo realizado por investigadores que buscaban definir y con ello poder medir la complejidad legal. Por otro lado también contiene información relativa a la ley que se estudiará en el presente trabajo.

Por su parte, en el capítulo dos se mencionan las propiedades de la ley que se contemplaron para las puntuaciones de complejidad que se proponen así como la construcción de dichas puntuaciones.

Con base a lo mencionado en el capítulo dos, el capítulo tres contiene los valores obtenidos de las propiedades contempladas así como las puntuaciones de complejidad calculadas de una ley en particular, la LISR.

El capítulo cuatro contiene el análisis de la puntuación obtenida así como las conclusiones.

Por último, el anexo contiene información complementaria de algunos aspectos teóricos o técnicos empleados en el presente trabajo. Por su parte, el apéndice contiene las puntuaciones de complejidad segregadas con base en el año y el título.

Capítulo 2

Métricas y puntuaciones de la complejidad - Teoría

2.1. Métricas

Con base en la clasificación de métricas realizada por los autores previamente mencionados, se separaron las métricas usadas en el presente trabajo en tres: métricas sobre la estructura de la ley, métricas sobre el lenguaje empleado en ley e interdependencias existentes en el contenido de la ley. La siguiente tabla muestra dichas métricas así como su clasificación:

Métrica	Clasificación
Densidad	Estructura
Profundidad	Estructura
Grado de salida	Estructura
Entropía	Lenguaje
Puntuación de Hub	Interdependencia
Puntuación de Authority	Interdependencia

Tabla 2.1: Métricas contempladas en el presente trabajo

2.1.1. Estructura

Se consideró a cada título, capítulo, sección, artículo, fracción, inciso, numeral, subinciso y apartado de las leyes como un nodo. Además, dado que dichos nodos están contenidos en una versión de la ley la cual depende del año en el que estuvo vigente, se consideró a dicho año como el nodo raíz. Se detectó que existe una jerarquía entre cada uno de los tipos de nodo, esto se deriva de la relación que tienen las distintas divisiones que integran a la ley, aspecto mencionado por Eliseo Muro Ruiz. Por otro lado, hay ausencia de ciclos en la gráfica de dicha ley ya que cada nodo que integran a las leyes está contenido en un único nodo padre. Lo anterior permitió contemplar la

estructura de la ley como la de un árbol. Bajo esta idea se obtuvieron las siguientes propiedades para poder describir a la estructura de la ley:

- Densidad
- Grado
- Profundidad

2.1.1.1. Densidad de la gráfica

En teoría de gráficas se define a la densidad de una gráfica como:

$$Densidad = \frac{2l}{n(n-1)}$$

Donde:

- l : número de aristas en la gráfica,
- n : número de vértices en la gráfica.

En el caso de gráficas con aristas orientadas, se define como:

$$Densidad = \frac{l}{n(n-1)}$$

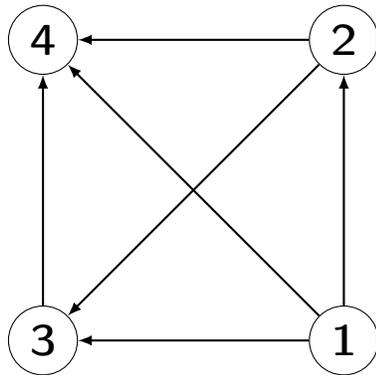


Figura 2.1: Ejemplo de la mayor densidad que puede tener esta gráfica: cuando todos los nodos están conectados entre sí.

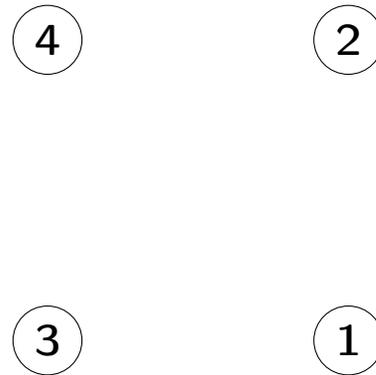


Figura 2.2: Ejemplo de la menor densidad que puede tener esta gráfica: cuando no existen aristas que conecten los nodos.

Entre mayor sea la densidad de una gráfica, mayor será el número de conexiones entre sus nodos. Usualmente, en las leyes todo nodo tiene un único nodo padre con excepción del nodo raíz. Esto hace que el número de aristas $l = n - 1$, por lo que la densidad sólo dependería del número de nodos en cada versión de la ley:

$$Densidad = \frac{l}{n(n-1)} = \frac{n-1}{n(n-1)} = \frac{1}{n}$$

2.1.1.2. Grado de un nodo

El grado de un nodo se define como el número de aristas que conectan al nodo con otros nodos. Por lo anterior, el grado mínimo que puede tener un nodo es 0 -cuando un nodo no está conectado con algún otro nodo- y el grado máximo es $l - 1$ -cuando el nodo está conectado con todos los demás nodos-.

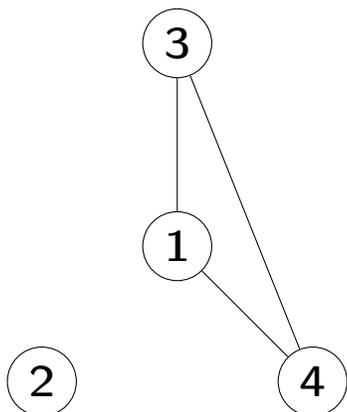


Figura 2.3: El grado del nodo 2 es 0, el grado mínimo que puede tener un nodo.

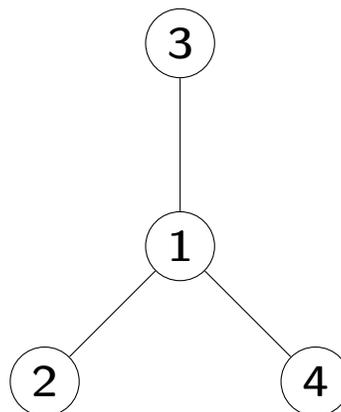


Figura 2.4: El grado del nodo 1 es 3, el grado máximo que puede tener un nodo en esa gráfica.

En gráficas con aristas orientadas, el grado se divide en dos:

- Grado de entrada (In-grade): es el número de aristas donde el nodo analizado es el nodo final.

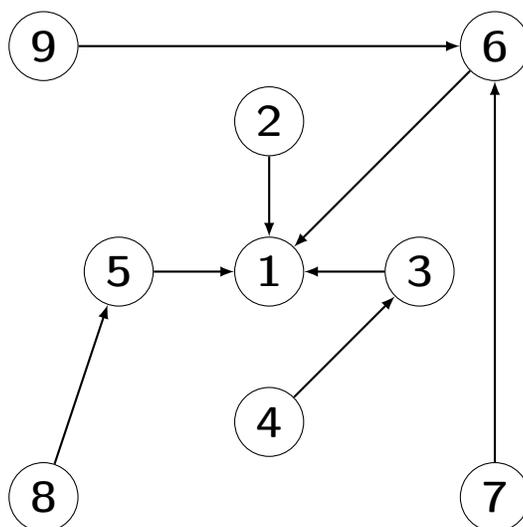


Figura 2.5: El grado de entrada del nodo 1 es 4, mientras que el del nodo 6 es 2.

- Grado de salida (out-degree): es el número de aristas donde el nodo analizado es el nodo inicial.

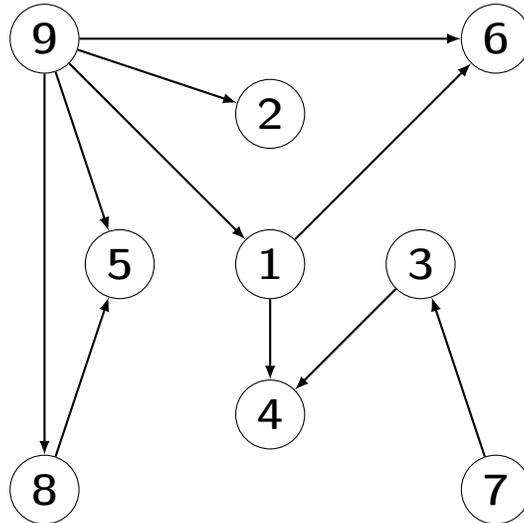


Figura 2.6: El grado de salida del nodo 9 es 5 mientras que el del nodo 4 es 0.

Dado que usualmente en las leyes todos los nodos hijo tienen un único nodo padre, con excepción del nodo raíz, el grado de entrada de todos los nodos sería igual a 1. Por esta razón, se propone no analizar el grado de entrada de los nodos sino el grado de salida de los nodos existentes en la ley.

2.1.1.3. Profundidad

La profundidad o nivel de un nodo es una propiedad definida en teoría de gráficas como la distancia del nodo analizado al nodo raíz, en donde cada nodo intermedio entre el nodo analizado y el nodo raíz aumenta la profundidad. Por ejemplo, el nodo raíz tiene una profundidad igual a 0.

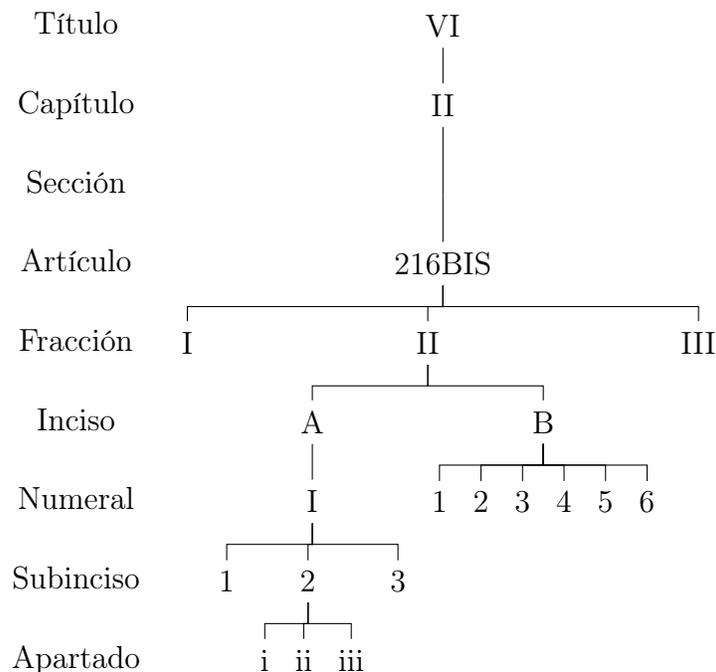


Figura 2.7: En este ejemplo, la profundidad máxima que puede tener un nodo es 7, como es el caso de los apartados i, ii y iii, dado que se omite el nivel "Sección".

2.1.2. Lenguaje

Las palabras contenidas en la ley son consideradas en esta sección para la generación de métricas del lenguaje en la ley. Estas palabras fueron sometidas a un proceso de stemming para encontrar las palabras raíz a la cual pertenecen y se diferencia con base al rol que tengan dichas palabras. Además, se consideraron sinónimos para agrupar en un sólo término a palabras que tengan el mismo significado. Cabe señalar que se hicieron dos análisis:

- En el primer análisis no se empleó una lista de palabras vacías, también conocidas como palabras a ignorar o stop words, por lo que todas las palabras detectadas fueron empleadas para el cálculo de las métricas.
- En el segundo análisis se hizo el cálculo de las métricas con el universo de palabras reducido por la lista de palabras vacías.

En ambos casos no se contemplaron las palabras que son abreviaciones, adverbios, preposiciones y conjunciones. Cabe señalar que los tipos de palabras no mencionados tales como verbos, adjetivos, artículos, etc. si se contemplaron en al menos uno de los dos análisis mencionados anteriormente.

2.1.2.1. Entropía

En 1948, Claude E. Shannon introduce el concepto de entropía en la teoría de la información mediante una función que propone en su trabajo *A Mathematical Theory of Communication*[10]. Con dicha fórmula le interesaba medir la cantidad de información que se 'produce' en un proceso. Estipula que dicha función, H , debe cumplir lo siguiente:

1. H debe ser continua en p_i .
2. Si todos los p_i son iguales, $p_i = \frac{1}{n}$, entonces H debe ser una función de n monótona creciente. Con eventos semejantes hay más alternativas, o incertidumbre, cuando hay más eventos posibles.
3. Si una opción se desglosa en dos opciones sucesivas, la original H debe ser ponderada con la suma de los valores individuales de H .

Demostró que la función que cumple con lo anterior es de la forma:

$$H = -K \sum_{i=1}^n p_i \log p_i$$

Donde:

- K : constante positiva,
- p_i : probabilidad de ocurrencia del evento i ,
- n : número total de eventos.

Shannon comparó la función H con el teorema H de Boltzmann y la definición de entropía en la mecánica estadística, razón por la cual llamó a la función H como entropía.

Cabe señalar que Shannon, al estar interesado en mensajes codificados en binario, usó \log_2 en su función H al calcular la entropía. Recuérdese que en el presente trabajo se midió la entropía a nivel palabra, diferenciando a las palabras que se escriben igual pero cuyo tipo de palabra es diferente. Por lo anterior, se usó la función extendida a D símbolos propuesta por Febres, Jaffe y Gershenson en su trabajo titulado *Complexity measurement of natural and artificial languages*[3]. Dicha extensión de la función de entropía de Shannon es la siguiente:

$$H = - \sum_{r=1}^D \frac{f_r}{L} \log_D \frac{f_r}{L}$$

Donde:

- D : total de símbolos distintos,

- L : total de símbolos,
- f_r : frecuencia del símbolo r .

Al depender la base del logaritmo de D se obtiene que $H \in [0, 1]$, donde $H = 0$ cuando el texto esté compuesto por una única palabra que se repite tantas veces como palabras contenga dicho texto. Por el contrario, ocurre que $H = 1$ cuando todas las palabras del texto son diferentes. Relacionado con lo anterior, Katz señaló lo siguiente:

Cuando un individuo está leyendo un título con alta entropía, es menos probable que ese individuo ya haya sido expuesto a una palabra o concepto extraído aleatoriamente de ese título. Alta entropía indica que el usuario final usualmente encontrará nuevas palabras y nuevos conceptos.[6]

Cabe señalar que si el individuo es alguien experimentado en tratar temas de la ley que está leyendo las palabras poco frecuentes que puede llegar a encontrar no necesariamente le resultarán nuevas pero si encontrará una mayor variedad de palabras en una ley con alta entropía que en una con baja entropía.

2.1.3. Interdependencia

Dado que en algunos artículos se citan a otras leyes u otros artículos dentro de la misma ley, la interpretación de dicho artículo pasa a depender del contenido de las citas. Autores como Katz o Matthes consideran que las citas modifican el esfuerzo necesario que una persona necesita para interpretar aquello que busca en la ley ya que debe estar consciente de todas las referencias que interactúen con ello. Dichos autores engloban dichas citas en dos tipos:

- Referencias internas.

Estas referencias son todas aquellas que se refieren elementos dentro de la misma ley. El siguiente fragmento ejemplifica este tipo de referencias:

Para estos efectos, la diferencia de los saldos de la cuenta de utilidad fiscal neta a que se refiere el inciso a) de la fracción II de este artículo, las pérdidas fiscales pendientes de disminuir, los reembolsos pagados y la diferencia a que se refiere el cuarto párrafo del artículo 88 de esta Ley, todos estos conceptos contenidos en el inciso b) de la citada fracción.

En este tipo de referencias hay dos elementos de interés: el que hace la referencia y el que es referenciado.

- Referencias externas.

Contrario a las referencias internas, las referencias externas se refieren a elementos que no pertenecen a la misma ley en la que se citan. Un ejemplo de este tipo de referencias es el siguiente:

Para los efectos de la participación de los trabajadores en las utilidades de las empresas, la renta gravable a que se refiere el inciso e) de la fracción IX del artículo 123 de la Constitución Política de los Estados Unidos Mexicanos y el artículo 120 de la Ley Federal del Trabajo.

En este tipo de referencias el elemento de interés es el que hace la referencia dado que todos los entes referenciados en este tipo de referencias son considerados como entes externos a la ley.

Dado que es de interés en ambos tipos de referencia el elemento que hace la referencia y particularmente con las referencias internas también es de interés el elemento referenciado, se optó por implementar el algoritmo conocido como Hyperlink-Induced Topic Search (HITS).

2.1.3.1. Hyperlink-Induced Topic Search

Es un algoritmo creado por Jon Kleinberg. Este algoritmo es empleado para valorar principalmente páginas web mediante la asignación de dos puntuaciones: una puntuación de Hub y una puntuación de Authority.

Las páginas Hub son las páginas que tienen múltiples vínculos a páginas consideradas como Authorities. Por su parte, las páginas Authority son aquellas que son percibidas o consideradas como una página relevante e importante dada una consulta de información, las cuales son vinculadas por páginas que referencian a otras páginas también consideradas como Authority. Por lo anterior, Kleinberg menciona que Hubs y Authorities tienen una relación de reforzamiento mutuo ya que una buena página Hub es una página que tiene vínculos que apuntan a muchas páginas Authorities: por su parte, una buena página Authority es una página que es referenciada por muchas páginas Hub.

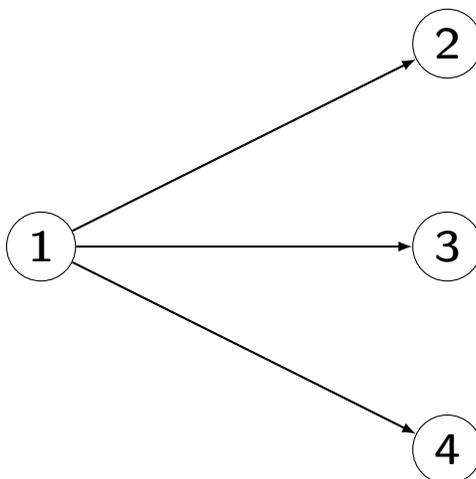
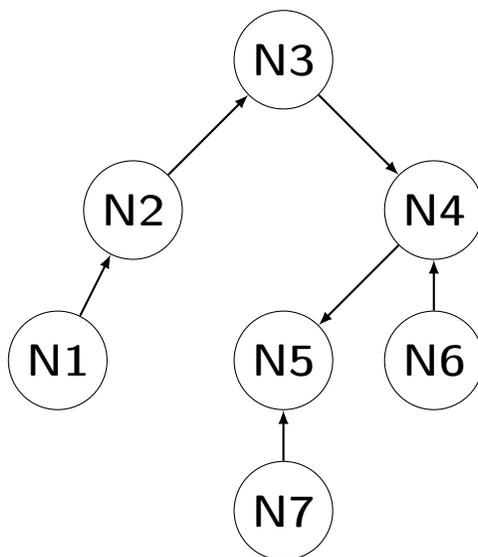


Figura 2.8: El nodo 1 es un nodo Hub que tiene por lo menos un vínculo que apunta a los nodos Authority, los cuales son los demás nodos.

En general el algoritmo se conforma de los siguientes pasos:

1. Escoger un número de iteraciones, k .
2. Asignar a cada nodo una puntuación inicial de Hub y de Authority igual a 1.
3. Se repiten k veces los siguientes pasos:
 - Se actualiza la puntuación de Hub de los nodos sumando la puntuación de Authority de cada uno de los nodos a los que apuntan sus vínculos.
 - Se actualiza la puntuación de Authority de los nodos sumando la puntuación de Hub de cada uno de los nodos cuyos vínculos apuntan al nodo.
 - En caso de querer calcular la puntuación normalizada, se divide la puntuación de Hub entre la suma de todas las puntuaciones Hub. Análogamente, se divide puntuación de Authority entre la suma de todas las puntuaciones Authority.

El siguiente ejemplo ilustra el algoritmo HITS. Considérese la siguiente gráfica y un número de iteraciones $k = 2$.



La matriz de adyacencia de la gráfica sería:

$$A = \begin{matrix} & \begin{matrix} N1 & N2 & N3 & N4 & N5 & N6 & N7 \end{matrix} \\ \begin{matrix} N1 \\ N2 \\ N3 \\ N4 \\ N5 \\ N6 \\ N7 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix}$$

Como el algoritmo lo indica, inicialmente todas las puntuaciones de Hub y de Authority son iguales a 1:

Nodo	Hub	Authority
N1	1	1
N2	1	1
N3	1	1
N4	1	1
N5	1	1
N6	1	1
N7	1	1

Estas puntuaciones de Hub y de Authority se pueden considerar como vectores u y v , respectivamente. Dado lo anterior, empieza el proceso iterativo. Con $k = 1$, se tiene lo siguiente:

- Se actualiza la puntuación de Hub de los nodos. Para el nodo N1 su nueva puntuación de Hub sería igual a 1 dado que sólo apunta a un nodo, el nodo N2. Por su parte, el nodo N5 pasaría a tener una puntuación de Hub igual a 0 ya que no apunta a algún nodo. Estas puntuaciones se pueden calcular de la siguiente forma:

$$A \times v = u$$

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

- Después, se actualiza la puntuación de Authority de los nodos. Para el nodo N1 su nueva puntuación de Authority sería igual a 0 dado que ningún nodo apunta al nodo N1. Por su parte, el nodo N4 pasaría a tener una puntuación de Authority igual a 2 ya que tanto el nodo N3 como el nodo N6 apunta a él. Estas puntuaciones se pueden calcular de la siguiente forma:

$$A^T \times u = v$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

Por lo que las puntuaciones finales en esta iteración son:

Nodo	Hub	Authority
N1	1	0
N2	1	1
N3	1	1
N4	1	2
N5	0	2
N6	1	0
N7	1	0

- Si se decide normalizar la puntuación, las puntuaciones finales serían:

Nodo	Hub	Authority
N1	1/6	0
N2	1/6	1/6
N3	1/6	1/6
N4	1/6	2/6
N5	0	2/6
N6	1/6	0
N7	1/6	0

En este ejemplo coincidió que la suma tanto de la puntuación de Hub así como de la suma de la puntuación de Authority es igual a 6 en la primer iteración. Con lo anterior acaba la primer iteración.

En la siguiente iteración, $k = 2$, se tiene lo siguiente:

- Se actualiza la puntuación de Hub:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1/6 \\ 1/6 \\ 2/6 \\ 2/6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/6 \\ 1/6 \\ 2/6 \\ 2/6 \\ 0 \\ 2/6 \\ 2/6 \end{pmatrix}$$

La suma de las puntuaciones de Hub obtenidas en esta iteración es igual a 10/6.

- Después, se actualiza la puntuación de Authority:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \\ 0 \\ 1/6 \\ 1/6 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/6 \\ 1/6 \\ 2/6 \\ 2/6 \\ 0 \\ 0 \end{pmatrix}$$

Se observa que las puntuaciones de Authority no cambiaron en esta iteración al compararlas con las puntuaciones finales normalizadas de la iteración anterior. Por lo anterior, su suma es igual a 1.

- Las puntuaciones finales, normalizadas, serían:

Nodo	Hub	Authority
N1	1/10	0
N2	1/10	1/6
N3	2/10	1/6
N4	2/10	2/6
N5	0	2/6
N6	2/10	0
N7	2/10	0

De esta forma concluye el algoritmo después de las 2 iteraciones contempladas. En el presente trabajo se obtendrán las puntuaciones Hub y Authority por tipo de referencia, si es interna o externa, y por tipo de nodo, los tipos de nodos que conforman la referencia. Cabe señalar que estas puntuaciones siempre son positivas.

2.2. Puntuación de complejidad con base en las métricas

Con base en las métricas mencionadas anteriormente, se construyeron las puntuaciones que integran esta sección, las cuales se proponen como puntuaciones de complejidad. La puntuación de complejidad legal de la ley del año t estaría dada con base en una media geométrica compuesta por los siguientes elementos:

$$\text{Complejidad legal}_t = \sqrt[3]{PE_t \cdot PL_t \cdot PI_t}$$

Donde:

- PE_t = puntuación de complejidad de la estructura de la ley del año t ,
- PL_t = puntuación de complejidad del lenguaje en la ley del año t ,

- PI_t = puntuación de complejidad de las interdependencias en la ley del año t .

Cada elemento de esta media geométrica está definido con base en las puntuaciones de complejidad de cada tipo, los cuales se describirán más adelante. Cabe señalar que dichos elementos siempre son positivos y distintos de cero. Lo anterior se explicará más adelante. Por otro lado, los valores dados por la media geométrica son positivos y entre más grande sea el valor mayor será la puntuación de complejidad asignada, por lo cual no tendrá límite superior. Respecto al límite inferior, este será un valor asintótico a cero.

2.2.1. Puntuación de complejidad de la estructura de la ley

El objetivo es generar una puntuación de complejidad cuyo valor sea obtenido con base en las métricas de estructura que se contemplaron para las leyes en el presente trabajo. Esta puntuación estaría dada con la siguiente fórmula:

$$PE_t = \sqrt[3]{PD_t \cdot PGS_t \cdot PP_t}$$

Donde:

- PD_t = puntuación de complejidad con base en la densidad de la gráfica de la ley del año t ,
- PGS_t = puntuación de complejidad con base en el grado de salida de los nodos de la ley del año t ,
- PP_t = puntuación de complejidad con base en la profundidad de los nodos de la ley del año t .

2.2.1.1. Puntuación de complejidad con base en la densidad de la gráfica

La puntuación de complejidad con base en la densidad se construyó con la siguiente fórmula:

$$PD_t = 1 - Densidad \cdot F_E$$

Donde:

- PD_t : puntuación de complejidad con base en la densidad de la ley del año t ,
- $Densidad$: densidad de la gráfica,
- F_E : factor de escala de la frecuencia de nodos.

El factor de escala de la frecuencia de nodos es una potencia de 10 tal que la frecuencia de los nodos entre el factor de escala sea igual a un número entre 1 y 10. Por ejemplo, en una gráfica con 15 nodos el factor de escala F_E sería igual a 10 ya que $\frac{15}{10} = 1.5$ el cual es un número entre 1 y 10. Por otro lado, si fuese una gráfica con 1050 nodos, el factor de escala sería igual a 1000 ya que $\frac{1050}{1000} = 1.05$ el cual es un

número entre 1 y 10. Con base en lo anterior y que la frecuencia de los nodos siempre es positiva, el término $Densidad \cdot F_E$ es un valor positivo menor a 1 y con ello la puntuación de complejidad con base en la densidad de la gráfica siempre es positiva y menor a 1. También cumple que entre mayor sea el valor de PD_t , mayor será la puntuación la complejidad dada por la métrica de densidad.

2.2.1.2. Puntuación de complejidad con base en el grado de salida de los nodos

Dado que el grado de salida es una propiedad a nivel nodo, el nivel de detalle de cada tipo de nodo no es el mismo entre ellos. Además, entre mayor sea su grado de salida mayor es el detalle que se le da a determinados aspectos contenidos en el nodo. Por lo anterior, la puntuación del grado de salida de los nodos se consideró como una media ponderada de la siguiente forma:

$$PGS_t = \frac{\sum_{i=1}^N GS_i \cdot w_i}{\sum_{i=1}^N w_i}$$

Donde:

- PGS_t : puntuación de complejidad con base en el grado de salida de los nodos de la ley del año t ,
- GS_i : grado de salida del i -ésimo nodo,
- $w_i = \sum_{j=1}^N PrTGS_{i,j} \cdot PrT_i \cdot PrGS_j$
 - N : total de nodos en la ley del año t con grado de salida distinto de 0,
 - $PrTGS_{i,j}$: proporción del tipo de nodo del i -ésimo nodo que tuvo grado de salida j ,
 - PrT_i : proporción del tipo de nodo del i -ésimo nodo respecto al total de nodos que tienen grado de salida distinto de 0,
 - $PrGS_j$: proporción de nodos que tienen grado de salida j .

Los valores de esta media ponderada siempre son positivos ya que no puede ser negativo el grado de salida contemplado en el presente trabajo, así como las proporciones mencionadas. Por lo anterior la puntuación de complejidad con base en el grado de salida siempre es positiva. Además, entre mayor sea el valor de PGS_t , mayor será la puntuación la complejidad dada por la métrica de grado de salida.

Dado que el grado de salida depende del número de nodos existentes así como de la distribución de estos en las relaciones existentes entre los nodos, los factores

que integran los pesos en la media ponderada se construyeron para que ponderaran la distribución de los distintos grados de salida en cada uno de los distintos tipos de nodo. El siguiente ejemplo ilustrará las distintas puntuaciones que se pueden obtener teniendo el mismo número de nodos:

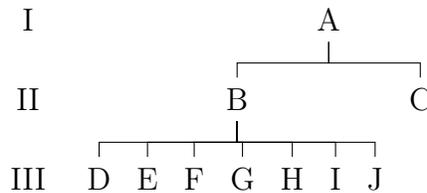


Figura 2.9: El grado de salida del nodo A es igual a 2 mientras que el grado de salida del nodo B es igual a 7. El valor de los factores de ponderación para el nodo A son: $PrGs = .5$, $PrT = .5$ y $PrTGS = 1$ por lo que el valor de $w = .25$. Estos valores son los mismos para el nodo B por lo que la puntuación de grado de salida sería: $PGS = \frac{2 \cdot .25 + 7 \cdot .25}{.25 + .25} = 4.5$

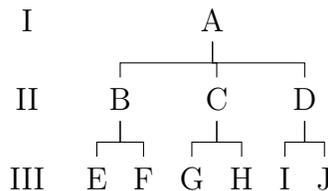


Figura 2.10: El grado de salida del nodo A es 3 mientras que el grado de salida de los demás nodos que no son nodos hoja es 2. El valor de los factores de ponderación para el nodo A son: $PrGs = .25$, $PrT = .25$ y $PrTGS = 1$ por lo que el valor de $w = .0625$. Por otro lado, el valor de los factores de ponderación para el nodo B son: $PrGs = .75$, $PrT = .75$ y $PrTGS = 1$ por lo que el valor de $w = .5625$. Estos valores son los mismos para los nodos C y D por lo que la puntuación de grado de salida sería: $PGS = \frac{3 \cdot .0625 + 3 \cdot 2 \cdot .5625}{.0625 + .5625} = 5.7$

Si bien en el primer ejemplo el grado de salida máximo es mayor que en el segundo ejemplo, el primero obtiene una puntuación menor que el segundo.

2.2.1.3. Puntuación de complejidad con base en la profundidad de los nodos

Al igual que el grado de salida, esta propiedad se calcula a nivel nodo por lo que también se consideró como una media ponderada de la siguiente forma:

$$PP_t = \frac{\sum_{i=1}^N P_i \cdot w_i}{\sum_{i=1}^N w_i}$$

Donde:

- PP_t = puntuación de complejidad con base en la profundidad de los nodos de la ley del año t ,
- P_i = profundidad del i -ésimo nodo hoja,
- $w_i = \sum_{j=1}^N PP_{i,j} \cdot PT_i \cdot PGS_j$
 - N = total de nodos hoja en la ley del año t ,
 - $PTP_{i,j}$ = proporción del tipo de nodo del i -ésimo nodo hoja que tuvo profundidad j ,
 - PT_i = proporción del tipo de nodo del i -ésimo nodo hoja respecto al total de nodos hoja existentes en la ley,
 - PP_j = proporción de nodos hoja que tienen profundidad j .

Análogamente a la media ponderada definida en el grado de salida, los valores de esta puntuación son positivos ya que la profundidad de un nodo es un valor positivo, así como las proporciones mencionadas. Además, entre mayor sea el valor de PP_t , mayor será la puntuación la complejidad dada por la métrica de grado de profundidad. Así mismo, los pesos de esta media ponderada se construyeron para que ponderaran la distribución de las distintas profundidades de los nodos. El siguiente ejemplo ilustrará las distintas puntuaciones que se pueden obtener teniendo el mismo número de nodos:

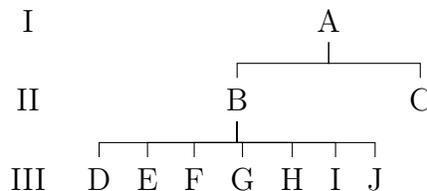


Figura 2.11: La profundidad de los nodos hoja D, E, F, G, H, I, y J es 2 y el valor de sus factores de ponderación son los siguientes: $PTP = 1$, $PT = .875$ y $PP = .875$ por lo que el valor de $w = .765625$ Por su parte, la profundidad del nodo C es 1 y el valor de sus factores de ponderación son los siguientes: $PTP = 1$, $PT = .125$ y $PP = .125$ por lo que el valor de $w = .015625$. Dado lo anterior, la puntuación de profundidad de esta gráfica sería: $PP = \frac{7 \cdot 2 \cdot .765625 + 1 \cdot 1 \cdot .015625}{7 \cdot 1.75 + .25} = 1.997093$

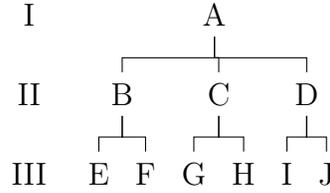


Figura 2.12: La profundidad de los nodos hoja E, F, G, H, I, y J es 2 y el valor de sus factores de ponderación son los siguientes: $PTP = 1$, $PT = 1$ y $PP = 1$ por lo que el valor de $w = 1$. Dado lo anterior, la puntuación de profundidad de esta gráfica sería: $PP = \frac{6 \cdot 1 \cdot 2}{6 \cdot 1} = 2$

En este ejemplo, las dos gráficas presentan una puntuación de profundidad muy similar siendo la primer gráfica la que recibe una puntuación menor por tener un nodo hoja con una profundidad menor, característica que no tiene la segunda gráfica.

2.2.2. Puntuación de complejidad del lenguaje en la ley

El objetivo es generar una puntuación de complejidad cuyo valor se obtenga con base en la métrica del lenguaje considerada para las leyes. Esta puntuación es igual a la única métrica considerada, la entropía:

$$PL_t = PEnt_t$$

Donde:

- $PEnt_t$ = puntuación de complejidad con base en la entropía de las palabras en la ley del año t .

2.2.2.1. Puntuación de complejidad con base en la entropía de las palabras

La puntuación de complejidad de esta métrica es el valor de la entropía obtenida con la función propuesta por Febres, Jaffe y Gershenson:

$$PEnt_t = H_t = - \sum_{i=1}^D \frac{f_i}{L} \log_D \frac{f_i}{L}$$

Donde:

- D : total de palabras distintas de la ley del año t ,
- L : total de palabras en la ley del año t ,
- f_i : frecuencia de la palabra i -ésima.

Recuérdese que los valores obtenidos con dicha fórmula de entropía son valores positivos entre 0 y 1.

2.2.3. Puntuación de complejidad con base en las interdependencias en la ley

El objetivo es generar una puntuación de complejidad cuyo valor se obtenga con base en las referencias internas y referencias externas detectadas en la ley. Esta puntuación se construyó con base en las puntuaciones agrupadas de Hub y de Authority, las cuales se agrupan en una sola utilizando la media geométrica para conformar la puntuación de complejidad de las interdependencias en la ley del año t :

$$PI_t = \sqrt[2]{PHub_t \cdot PAut_t}$$

Donde:

- $PHub_t$: puntuación agrupada de Hub de la ley del año t ,
- $PAut_t$: puntuación agrupada de Authority de la ley del año t .

2.2.3.1. Puntuación de complejidad con base en las puntuaciones Hub y Authority

Estas puntuaciones se construyeron con base en las puntuaciones de Hub y de Authority calculadas con en algoritmo HITS. Dado que las puntuaciones son calculadas por tipo de nodo, se agruparon cada una de las puntuaciones utilizando la media geométrica.

$$PHub_t = \sqrt[n]{\prod_{i=1}^n Hub_i}$$

$$PAut_t = \sqrt[n]{\prod_{i=1}^n Aut_i}$$

Donde:

- n : total de tipos de nodo en la ley del año t ,
- Hub_i : puntuación de Hub del tipo de nodo i -ésimo,
- Aut_i : puntuación de Authority del tipo de nodo i -ésimo,
- $PHub_t$: puntuación agrupada de Authority de la ley del año t ,
- $PAut_t$: puntuación agrupada de Hub de la ley del año t .

Dado que las puntuaciones Hub y Authority que son iguales a cero no se contemplaron en la media geométrica previamente mencionada, se obtuvieron únicamente puntuaciones positivas. Lo anterior se hizo para evitar que los valores iguales a cero hicieran cero las demás puntuaciones.

Estas fueron todas las propiedades y sus respectivas puntuaciones que se construyeron en el presente trabajo para elaborar una puntuación de complejidad legal. Cabe señalar que las propiedades consideradas se pueden obtener de cualquier ley que integre el marco jurídico en México y con ello la puntuación propuesta por lo que se pueden emplear para comparar distintas leyes o distintas versiones de una misma ley. Este último caso será el uso que se le dará a la puntuación de complejidad legal en el presente trabajo, lo cual se desarrolló en el siguiente capítulo.

Capítulo 3

Implementación de las métricas en la LISR

En este capítulo se presentarán los valores obtenidos de las propiedades mencionadas en el capítulo anterior aplicadas a la LISR, así como las puntuaciones de complejidad que se obtuvieron con base en dichas propiedades.

El periodo analizado de la LISR comprende desde el 2008 hasta el 2017, periodo en el cual se dieron los siguientes eventos notorios que afectaron a la LISR:

- 2008: Creación del Impuesto Empresarial a Tasa Única (IETU) y el Impuesto a los Depósitos en Efectivo (IDE).
- 2013: Derogación del IETU y el IDE en la LISR.

3.1. Métricas en la LISR

Siguiendo el orden que se usó para presentar las propiedades consideradas en el capítulo anterior, primero se presentarán los resultados obtenidos en las propiedades estructurales, después se dará paso a lo obtenido con la propiedad del lenguaje y finalmente se mostrarán los resultados con las propiedades de interdependencia.

3.1.1. Estructura

3.1.1.1. Nodos en las gráficas

En la siguiente tabla se muestra el número de nodos que integran a la LISR, desde el 2008 hasta el 2017:

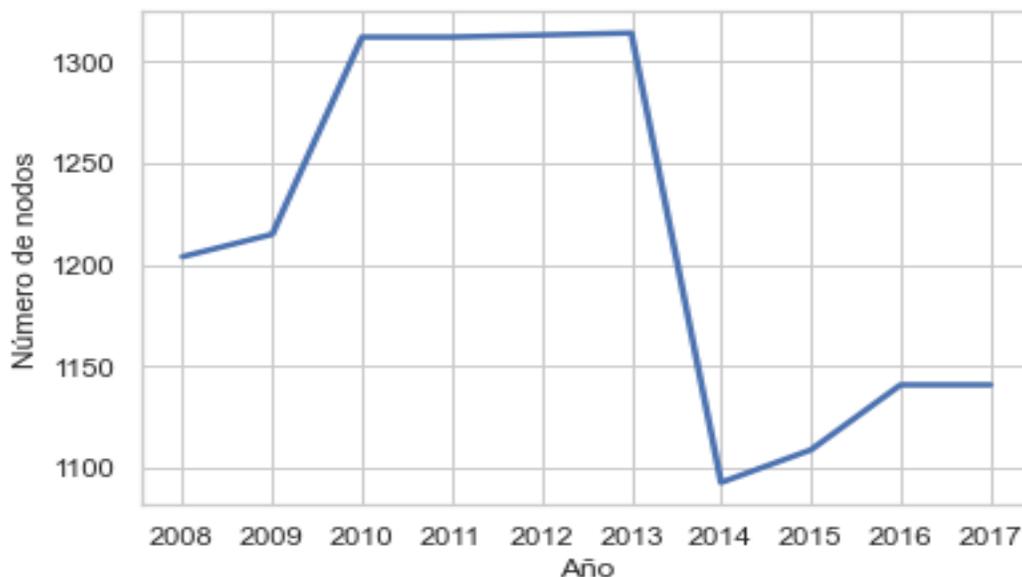


Figura 3.1: Número de nodos por año en la LISR

Se observó que el cambio más notorio en el número de nodos fue de 2013 a 2014 dado que pasó de 1325 nodos a 1126 nodos. Este cambio se debió a que en la reforma del 2014 se eliminaron dos impuestos de la LISR, el IETU y el IDE.

Por otro lado, en las siguientes gráficas se puede observar la frecuencia de los nodos por tipo y año:

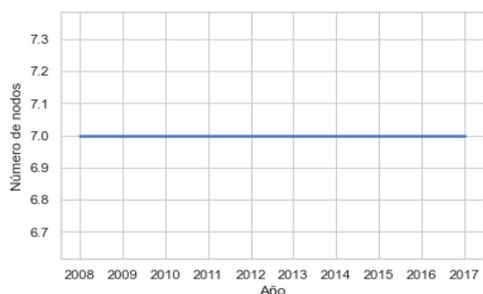


Figura 3.2: Número de títulos por año en la LISR

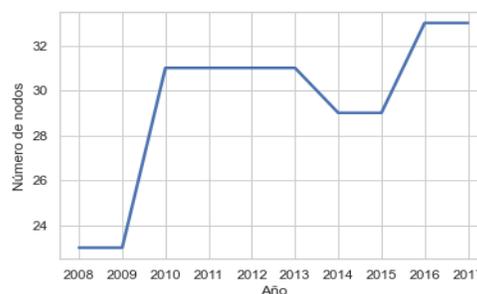


Figura 3.3: Número de capítulos por año en la LISR

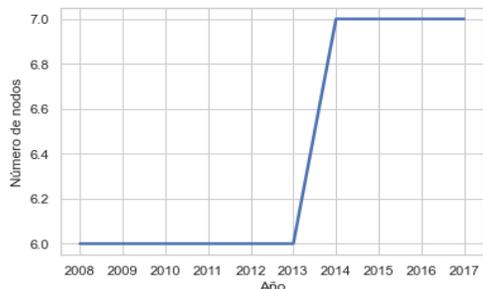


Figura 3.4: Número de secciones por año en la LISR

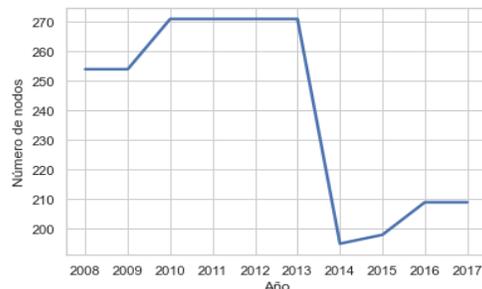


Figura 3.5: Número de artículos por año en la LISR

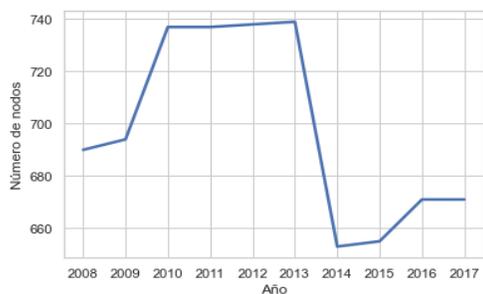


Figura 3.6: Número de fracciones por año en la LISR

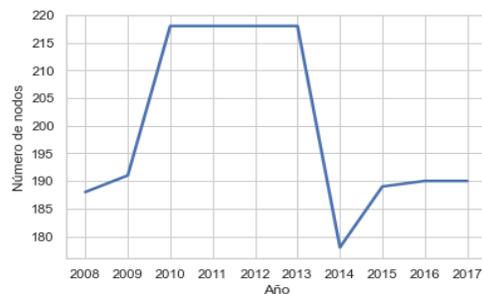


Figura 3.7: Número de incisos por año en la LISR

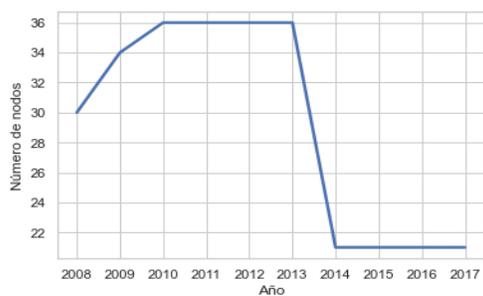


Figura 3.8: Número de numerales por año en la LISR

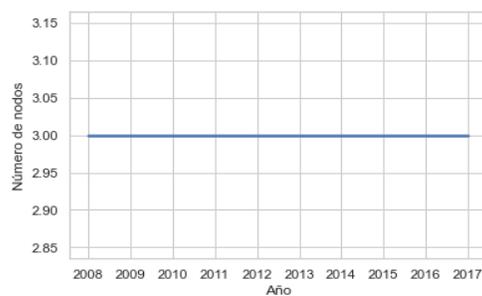


Figura 3.9: Número de sub incisos por año en la LISR

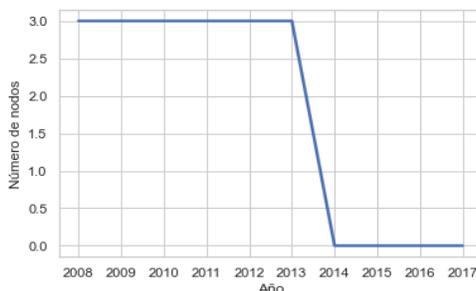


Figura 3.10: Número de apartados por año en la LISR

Es notorio que los nodos de tipo fracción son los que más contribuyen al total de nodos en todas las versiones de la LISR dado que las frecuencias de este tipo de nodo varía entre 650 y 750, mientras que los que menos contribuyen son los nodos de tipo título. Lo anterior es un comportamiento esperado dado que los nodos de tipo fracción están contenidos en artículos en una relación de muchos a uno, los cuales a su vez están contenidos en títulos o capítulos y secciones los cuales están contenidos en una relación de muchos a uno. Un comportamiento que se da en la mayoría de los nodos es un decremento en la frecuencia de 2013 a 2014 por lo que la reestructuración que se aplicó a la LISR de 2014 afectó a la mayoría de los tipos de nodo. Los tipos de nodo que tuvieron un mayor cambio en variaciones porcentuales en el periodo de tiempo analizado fueron los nodos de tipo Apartado y de tipo Numeral:

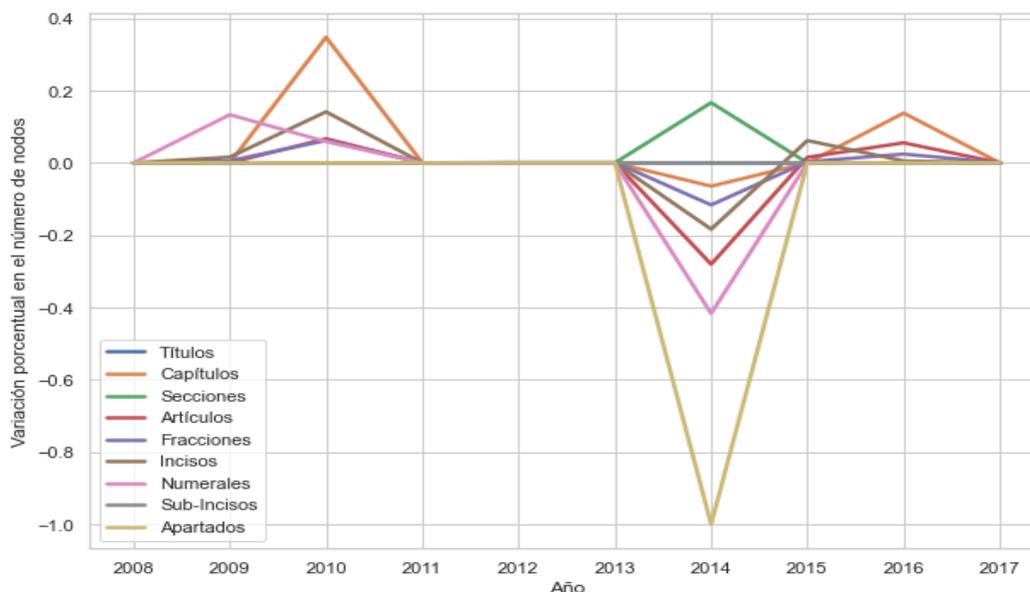


Figura 3.11: Variación porcentual por tipo de nodo

No obstante, si se observan los cambios en la frecuencia de los tipos de nodo en el periodo analizado los tipos de nodo con mayores cambios en la frecuencia no son los mismos que los que presentan un mayor cambio en variaciones porcentuales:

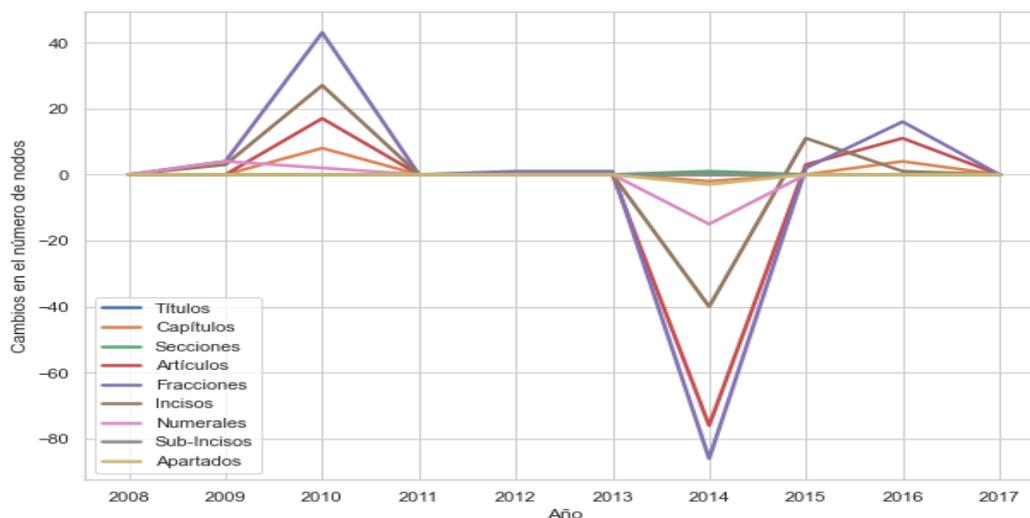


Figura 3.12: Cambios en la frecuencia por tipo de nodo

Los tipos de nodo con un mayor cambio en la frecuencia son tipos de nodo cuyas frecuencias tienen mayores escalas, tales como los artículos y las fracciones.

3.1.1.2. Densidad de las gráficas

La densidad por año de la LISR se puede observar en la siguiente gráfica:

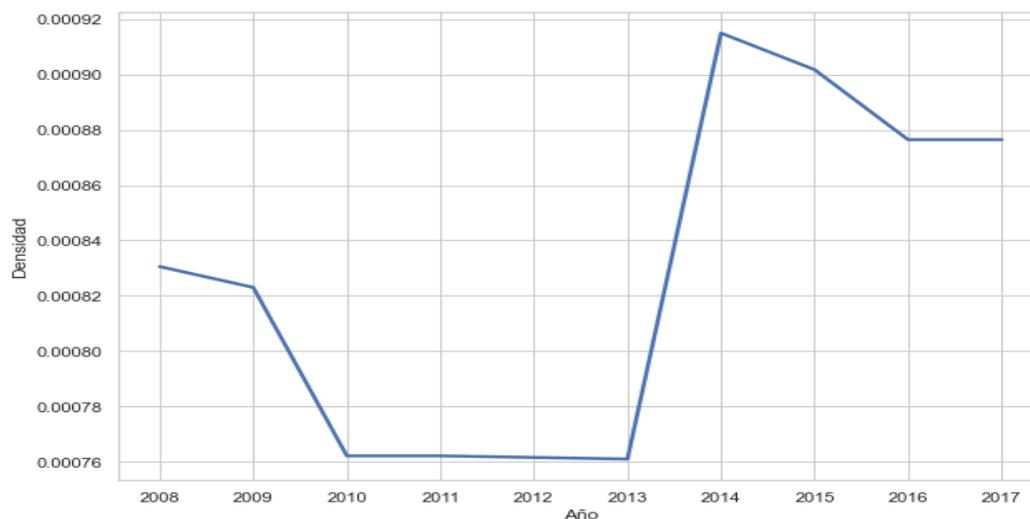


Figura 3.13: Densidad obtenida por año en la LISR

Es notorio un cambio de densidad en el 2014. Esto era de esperarse si se considera que la densidad de la gráfica de la LISR se calculó como $\frac{1}{n}$, con n igual al número de nodos, y la LISR del 2014 cuenta con el mínimo número de nodos durante el periodo analizado. Se observa que, con excepción del año 2014, esta propiedad tiene una tendencia decreciente en el periodo de tiempo analizado la cual era de esperarse dada la relación inversamente proporcional que tiene con la tendencia creciente en el número de nodos.

3.1.1.3. Grado de salida de los nodos

Una vez calculado el grado de salida de todos los nodos existentes en la LISR, se observó que la mayoría de los nodos, alrededor del 80 %, tienen grado de salida igual a 0. Lo anterior se puede observar en el siguiente ejemplo:

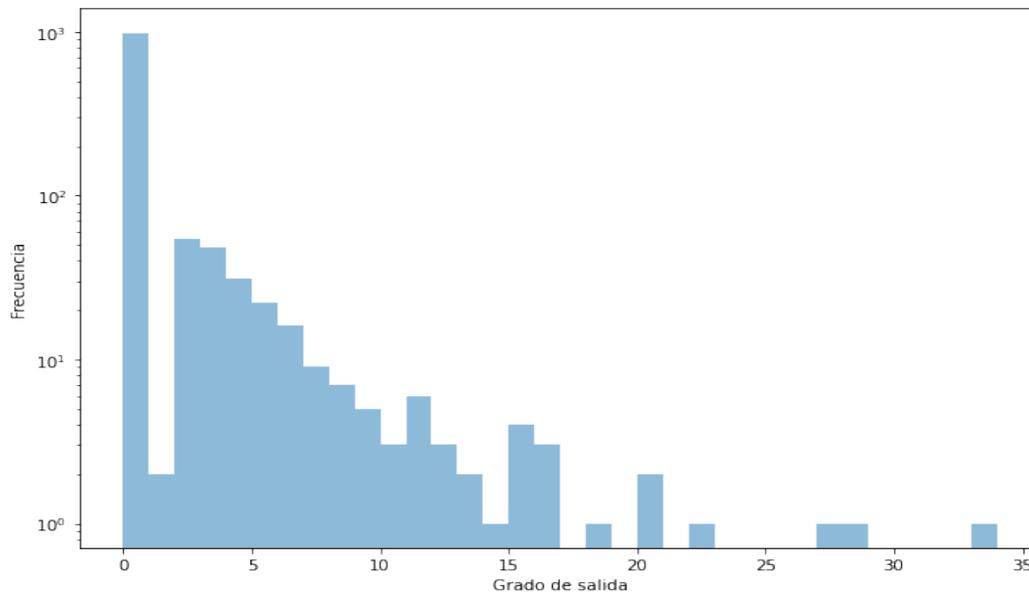


Figura 3.14: Distribución del grado de salida considerando todos los nodos

Esto indica que la mayoría son nodos de tipo hoja. Por lo anterior, se optó por contemplar únicamente los nodos con grado de salida distinto de cero. Al generar las distribuciones de estos nodos se observó que las distribuciones casi no cambiaban. Todos los años analizados presentaron una distribución muy similar a la siguiente:

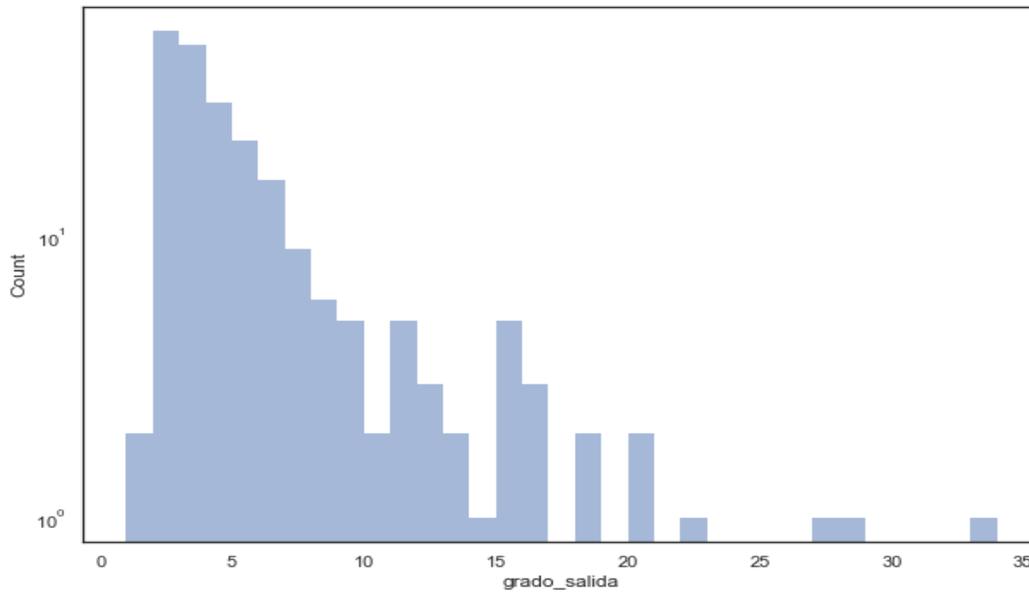


Figura 3.15: Distribución representativa de los nodos con grado de salida distinto de cero

Por ello, se identificó que la mayoría de dichos nodos no tuvo un cambio en su grado de salida durante el periodo analizado. Esto se puede apreciar en la siguiente gráfica:

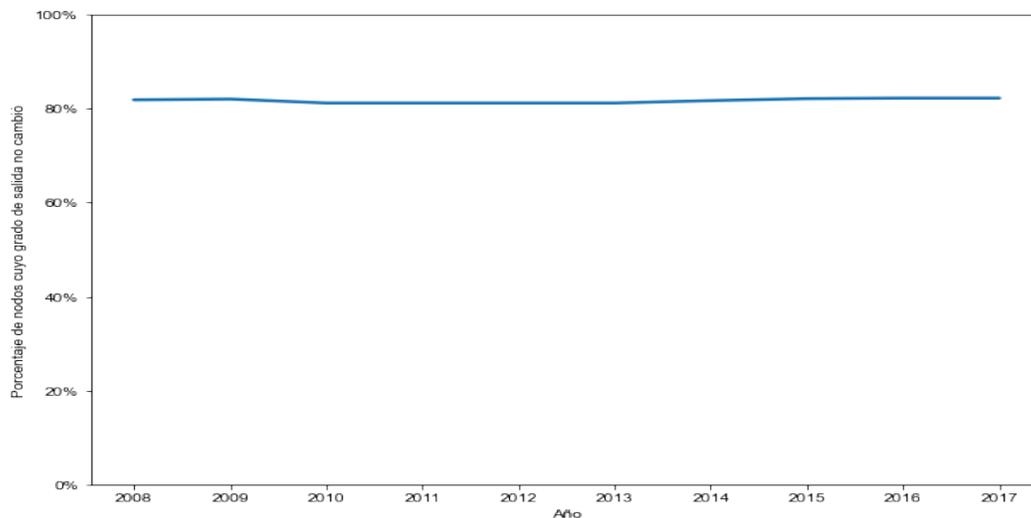


Figura 3.16: Porcentaje de nodos con grado de salida distinto de cero cuyo grado de salida no cambió

También se detectó que la mayoría de los nodos cuyo grado de salida se modificó sólo tuvieron un cambio en dicha propiedad:

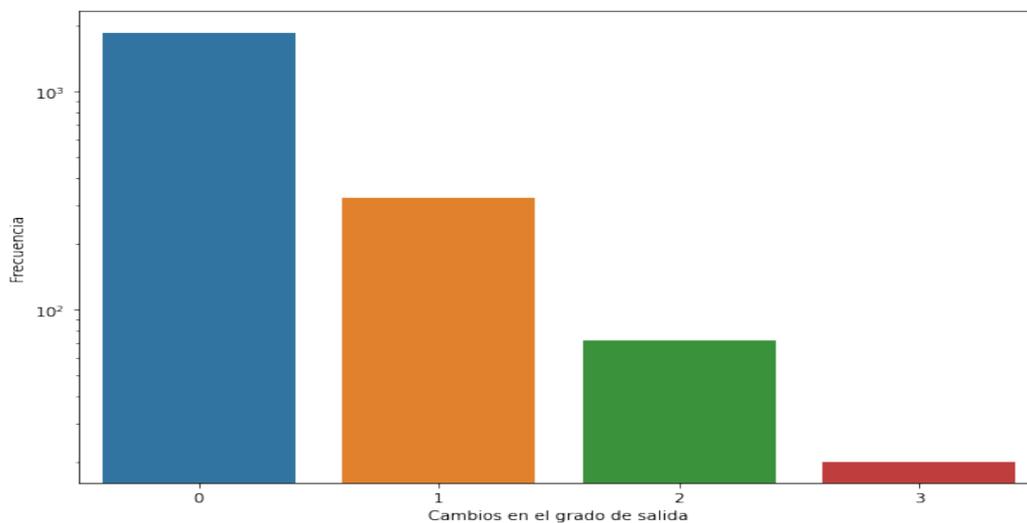


Figura 3.17: Cambios en el grado de salida de los nodos durante el periodo analizado

Por lo anterior, se optó por analizar la evolución de la distribución del grado de salida de los nodos que fue distinto de cero:

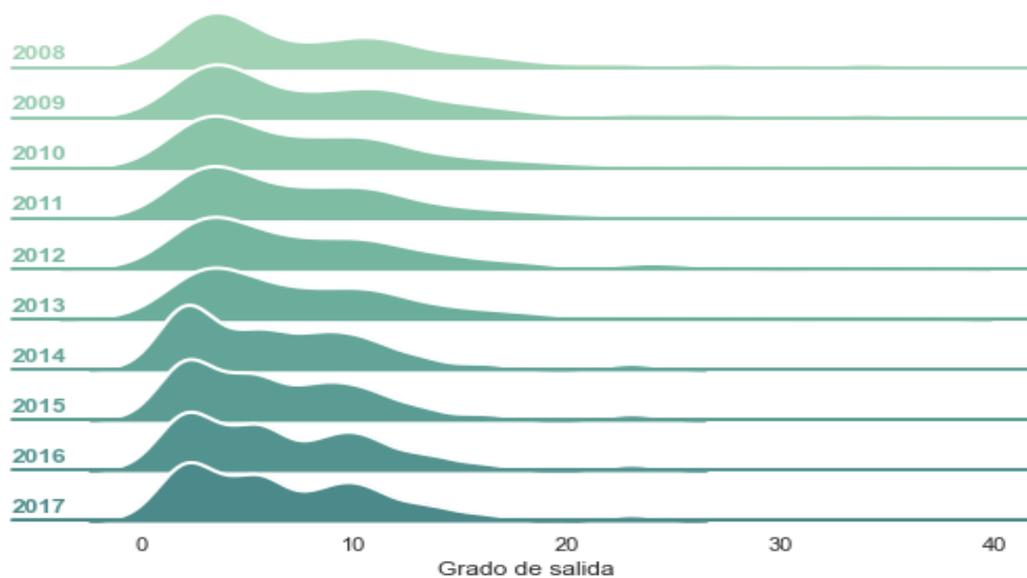


Figura 3.18: Evolución de la distribución del grado de salida de los nodos

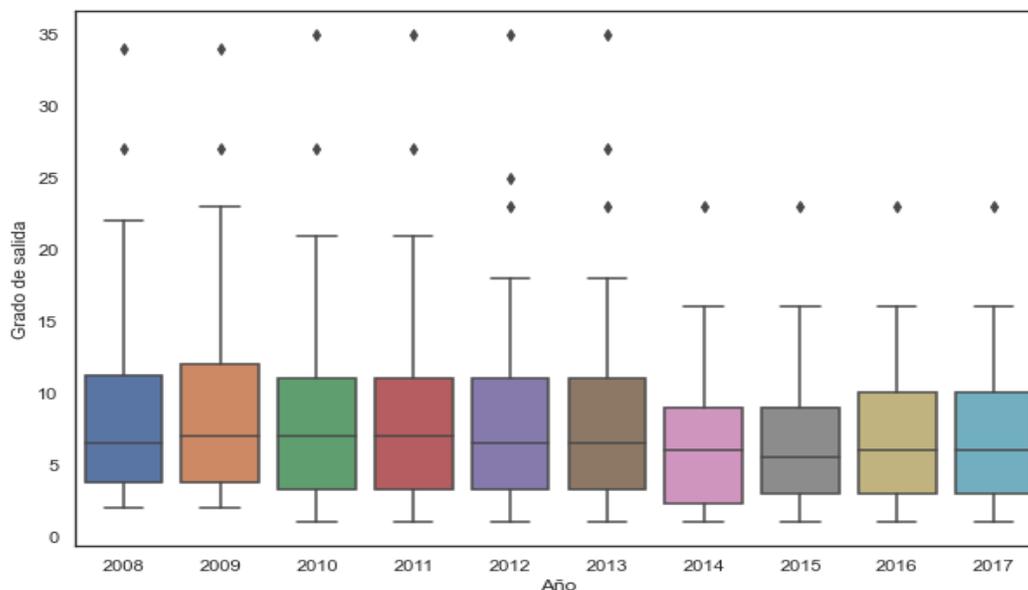


Figura 3.19: Evolución de las gráficas de cajas del grado de salida por año

Año	Q1	Q2	Q3	Media	Coefficiente de sesgo	Curtosis
2008	3.5	6.5	11.5	8.825	1.7139728	3.4403147
2009	3.5	7	12	8.925	1.693800564	3.32091943
2010	3	7	11	8.47826087	1.870818186	4.53016941
2011	3	7	11	8.47826087	1.870818186	4.53016941
2012	3	6.5	11	8.5	1.82470217	4.31034222
2013	3	6.5	11	8.52173913	1.861665358	4.32877252
2014	2	6	9	6.55263158	1.290345697	2.38217799
2015	3	5.5	9	6.60526316	1.316958552	2.51414691
2016	3	6	10	6.84615385	1.151339887	1.78391941
2017	3	6	10	6.84615385	1.151339887	1.78391941

Figura 3.20: Estadísticos de forma de las distribuciones del grado de salida

Visualmente se observa que las distribuciones muestran un sesgo a la derecha, característica que también se puede detectar con base en el valor del coeficiente de sesgo ya que en todos los años es mayor a 0. Por otro lado, el valor de la curtosis es mayor a 0 en todos los años, lo cual indica una forma leptocúrtica. El valor de los dos estadísticos de forma mencionados anteriormente indica una concentración en los valores, la cual se puede apreciar que es en los valores pequeños de la distribución. Cabe señalar que a partir del año 2014 empieza a generarse otra concentración al rededor del valor 10, lo cual da la impresión de que la distribución empezó a evolucionar a una distribución bimodal. También se detectó la existencia de valores atípicos y que el valor de los primeros 2 cuartiles casi no cambió. Por lo anterior, se optó por

considerar únicamente a los nodos con grado de salida distinto de 0 en posteriores análisis del grado de salida así como en la construcción de la puntuación del grado de salida.

Por otro lado, cabe señalar que esto contempla el grado de salida de todos los tipos de nodo igual. Si se considera el grado de salida por tipo de nodo y año se obtiene lo siguiente:

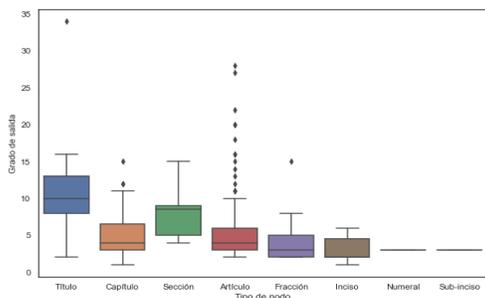


Figura 3.21: Grados de salida por tipo de nodo en el 2008

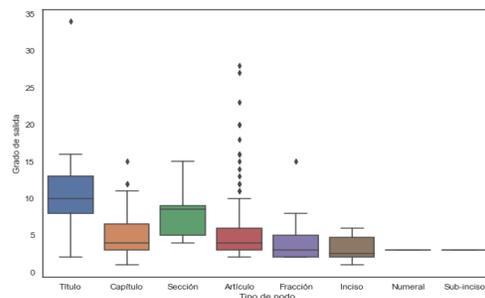


Figura 3.22: Grados de salida por tipo de nodo en el 2009

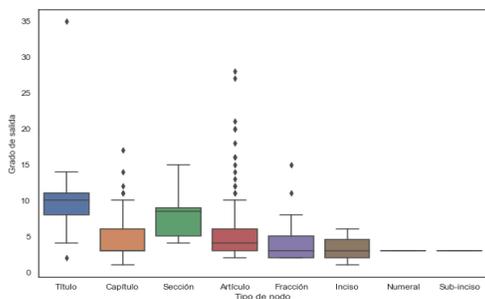


Figura 3.23: Grados de salida por tipo de nodo en el 2010

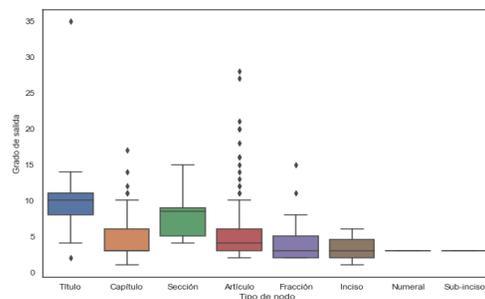


Figura 3.24: Grados de salida por tipo de nodo en el 2011

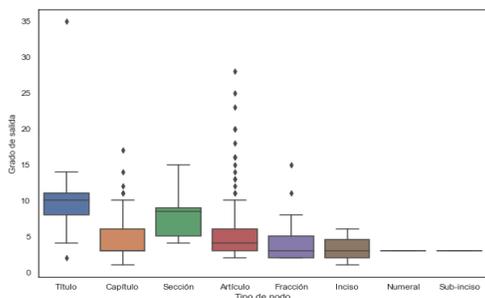


Figura 3.25: Grados de salida por tipo de nodo en el 2012

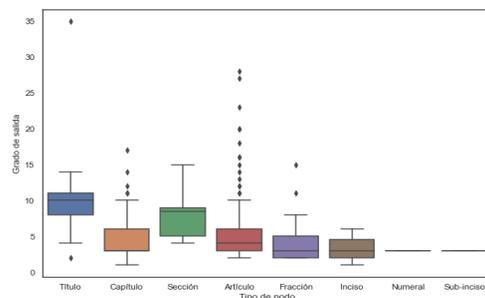


Figura 3.26: Grados de salida por tipo de nodo en el 2013

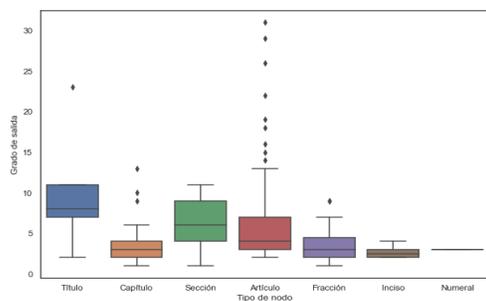


Figura 3.27: Grados de salida por tipo de nodo en el 2014

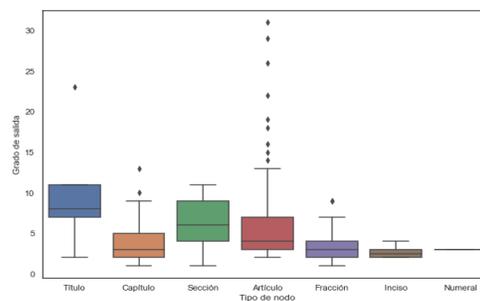


Figura 3.28: Grados de salida por tipo de nodo en el 2015

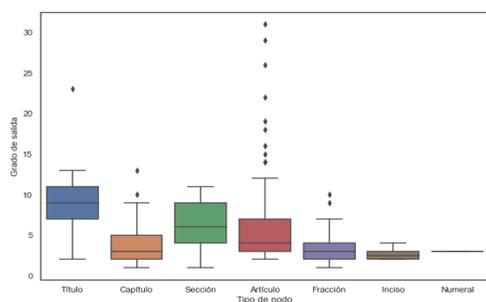


Figura 3.29: Grados de salida por tipo de nodo en el 2016

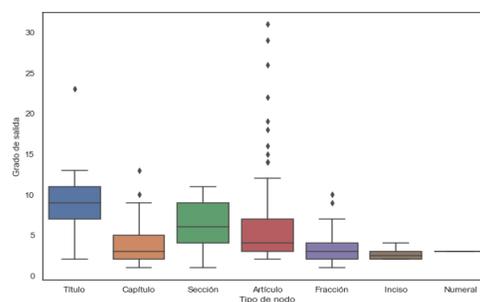


Figura 3.30: Grados de salida por tipo de nodo en el 2017

Es notorio que los títulos y secciones presentan grados de salida grandes respecto al resto del tipo de nodos. Esto se conserva con el pasar de los años. Como se mencionó en el capítulo I, los títulos se emplean como una de las principales divisiones de los cuerpos legales, pues se refieren a la partición más general de una ley, a su vez los capítulos dividen al título y las secciones son subdivisiones del capítulo. Lo anterior explica por qué el grado de salida de estos tipos de nodo presentan los valores más grandes respecto a los valores de los demás tipos de nodo. Por otro lado, se detecta la existencia de observaciones anómalas, sobre todo en los nodos de tipo artículo. Este tipo de nodo es el que concentra el mayor número de observaciones anómalas, característica que mantiene en el periodo de tiempo analizado. Esto indica la existencia de artículos cuya información debe ser detallada en más divisiones que el resto de los artículos.

3.1.1.4. Profundidad de los nodos

Al contrario de la propiedad anterior, en esta propiedad se consideró sólo a los nodos hoja. La distribución de esta propiedad presentó las siguientes características:

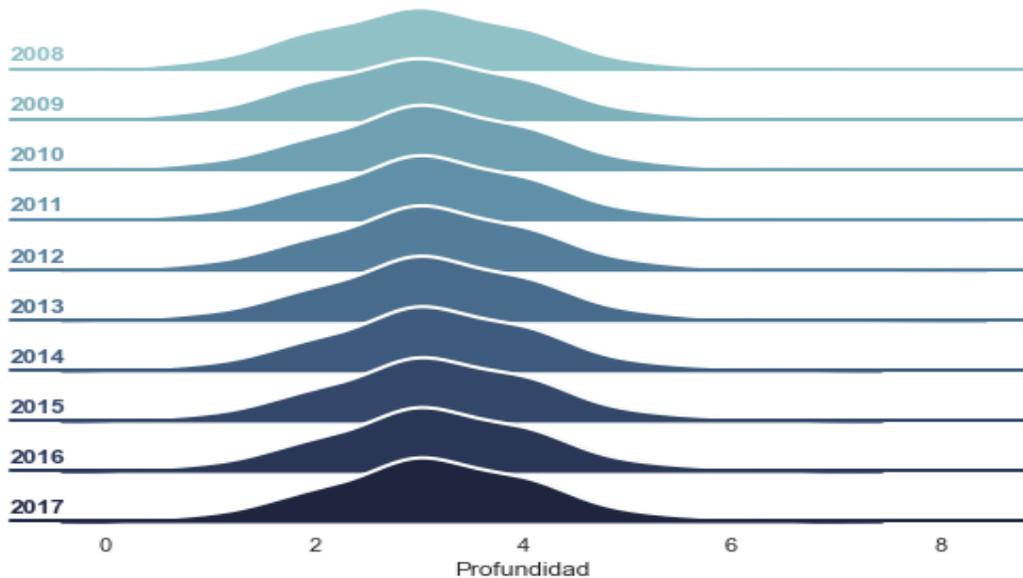


Figura 3.31: Evolución de la distribución de la profundidad

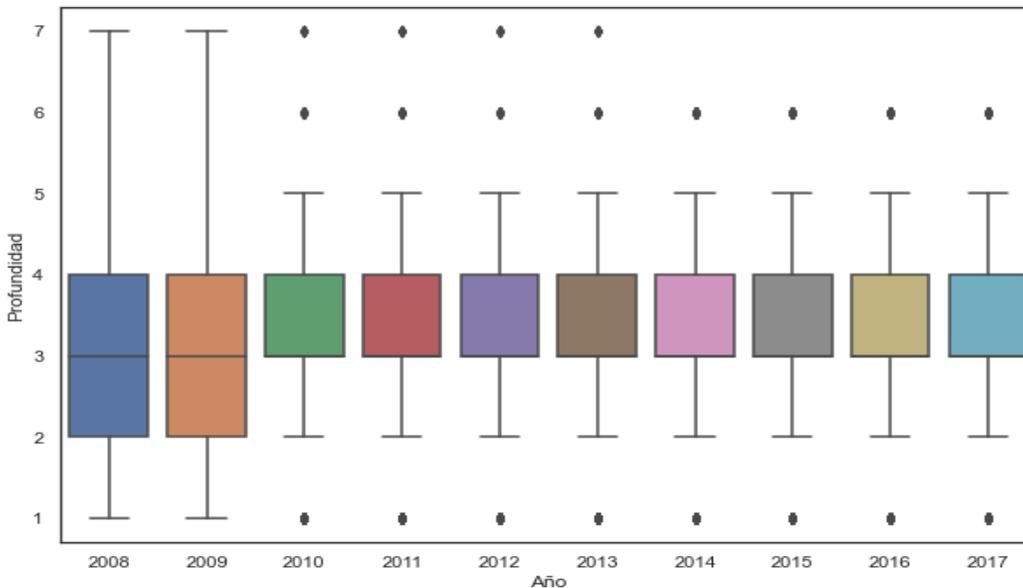


Figura 3.32: Gráficas de cajas de la profundidad por año

Se observó que las distribuciones casi no cambiaron durante el periodo analizado, así como las gráficas de cajas. Por ello, se analizaron los cambios que tuvieron los tipos de nodo que conforman a los nodos hoja. Dichos cambios se apreciaron en la siguiente gráfica:

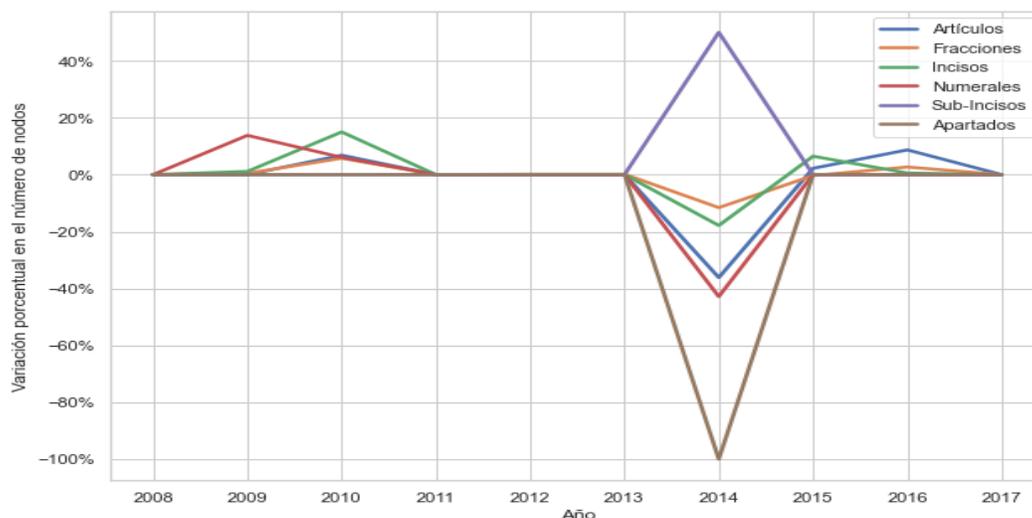


Figura 3.33: Cambios en la frecuencia por tipo de nodo

Se observó que los tipos de nodo que tuvieron un mayor cambio fueron los apartados, los numerales y los sub-incisos. No obstante, los cambios en la frecuencia que tuvieron estos tipos de nodo es baja comparada con la que tuvieron los demás tipos de nodo. Este detalle se pudo observar en la siguiente gráfica:

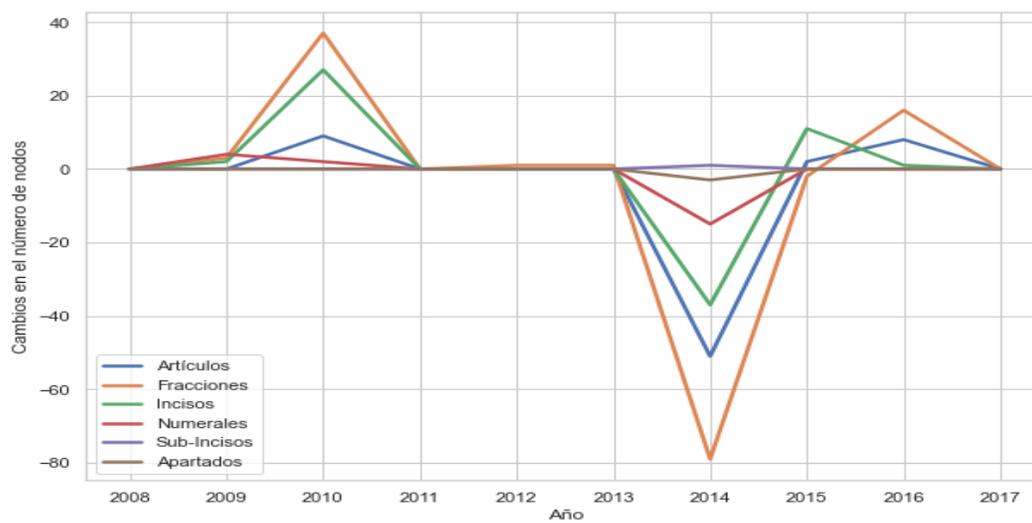


Figura 3.34: Cambios en la frecuencia por tipo de nodo

Dado que los nodos de tipo artículo o fracción fueron el mayor tipo de nodo entre los nodos hoja, no resultó extraño ver que la profundidad propia de estos tipos de nodo sea la que concentre la mayor cantidad de nodos hoja. Lo anterior se pudo observar en las gráficas de caja así como en los valores de la siguiente tabla:

Año	Q1	Q2	Q3	Media	Coefficiente de sesgo	Curtosis
2008	2	3	4	3.0315	0.2202	0.459
2009	2	3	4	3.0463	0.2598	0.461
2010	3	3	4	3.1387	0.2145	0.5927
2011	3	3	4	3.1387	0.2145	0.5927
2012	3	3	4	3.1386	0.215	0.596
2013	3	3	4	3.1403	0.2105	0.5883
2014	3	3	4	3.1774	0.193	0.351
2015	3	3	4	3.1864	0.1832	0.3011
2016	3	3	4	3.1716	0.1863	0.3146
2017	3	3	4	3.1716	0.1863	0.3146

Figura 3.35: Estadísticos de forma de las distribuciones de densidad

Con base en los valores del coeficiente de sesgo, los cuales son mayores a 0, se detectó que las distribuciones presentaron un pequeño sesgo a la derecha, característica que se puede notar en los valores de la media y la mediana los cuales no presentan grandes diferencias. Por otro lado, los valores de la curtosis, los cuales son mayores a 0, indican que las distribuciones presentan una forma leptocúrtica, característica que se puede apreciar visualmente en las distribuciones. Lo anterior se mantiene en todos los años del periodo de tiempo analizado.

Por otro lado, pese a que la mediana no cambia, el rango intercuartil se hace más pequeño a partir del año 2010 dado que el primer cuartil cambió de 2 a 3. Esto indica que aumentó el número de hojas que pasaron a presentar profundidades mayores a 2. Además de la reducción en el rango intercuartil aparecieron observaciones atípicas, las cuales pasan a disminuir desde 2014 por lo que la reestructuración de ese año también le afectó a la profundidad de nodos hoja considerados como atípicos.

3.1.2. Lenguaje

3.1.2.1. Palabras en la LISR

Durante el periodo analizado se identificó el siguiente número de palabras distintas y palabras totales:

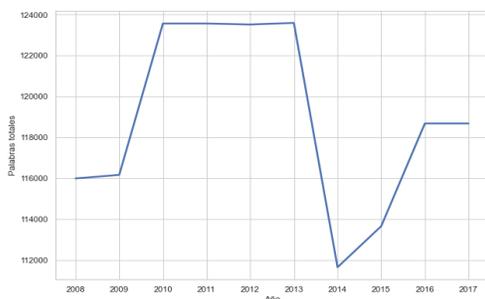


Figura 3.36: Número total de palabras por año

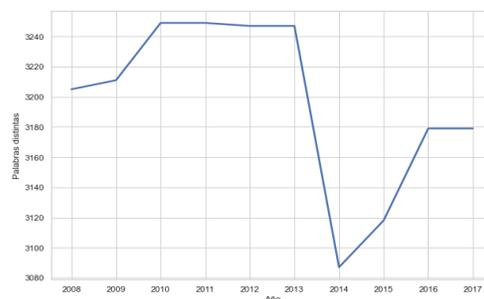


Figura 3.37: Número total de palabras distintas por año

Semejante a lo observado en las métricas de estructura, en el año 2014 hubo un cambio sobresaliente tanto en las palabras totales como en las palabras distintas que integran la LISR. Por lo anterior, el cambio en la LISR que hubo del 2013 al 2014 no sólo reestructuró la LISR disminuyendo el número de nodos que la integraban, también disminuyó las palabras que la integraban. No obstante, el cambio no fue en la misma proporción para las palabras totales y las palabras distintas ya que el número de palabras totales decreció al rededor de 10 % mientras que el número de palabras distintas decreció al rededor de 5 %. Esto se apreció en la siguiente gráfica:

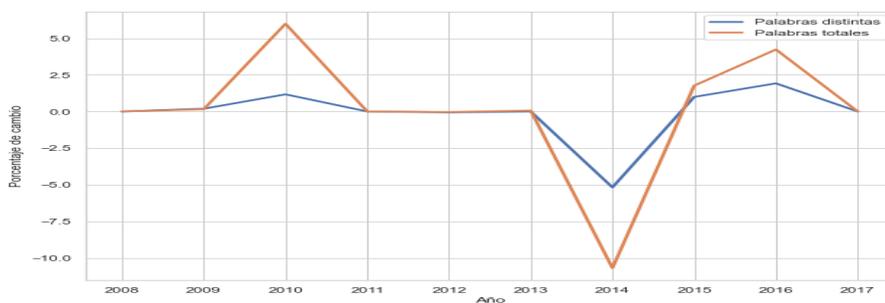


Figura 3.38: Variaciones porcentuales por año

El comportamiento anterior se dio en otros dos años, 2010 y 2016, pero la diferencia en el cambio es menos grande que en el 2014. Estos cambios también se pueden apreciar en la siguiente gráfica:

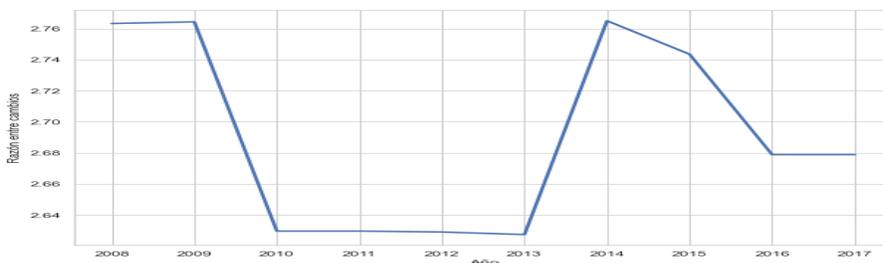


Figura 3.39: Razón entre palabras distintas y palabras totales por año con todas las palabras

Cabe señalar que lo anterior se obtuvo sin emplear una lista de palabras vacías pero si se empleó una lista de sinónimos así como multi-términos. Si se emplea una lista de palabras vacías, se obtiene lo siguiente:



S

Figura 3.40: Número total de palabras por año

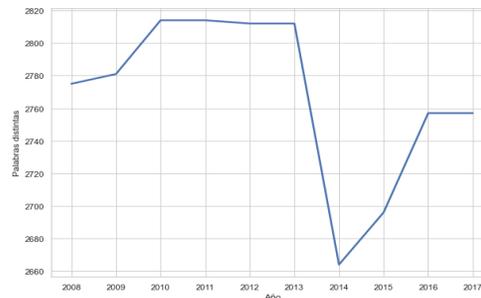


Figura 3.41: Número total de palabras distintas por año

Los años con cambios sobresalientes son los mismos que sin usar una lista de palabras vacías pero la razón entre palabras distintas y palabras totales cambió considerablemente ya que de presentar valores entre 2 y 3 pasó a presentar valores entre 8 y 9 tras usar una lista de palabras vacías:

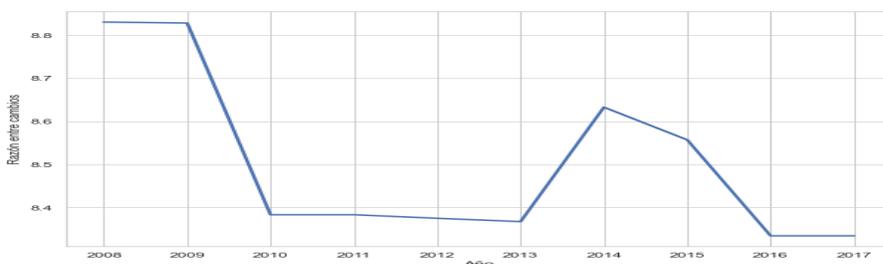


Figura 3.42: Razón entre palabras distintas y palabras totales por año usando una lista de palabras vacías

Lo anterior se esperaba dado que las palabras más frecuentes, sin usar una lista de palabras vacías, son las siguientes:

ser este o ley referir
 e a l con no de en y
 su con no ejercicio un
 del que para articulo al ingreso
 se

Se observa que varias de las palabras más frecuentes son palabras que son artículos, adjetivos posesivos o conjunciones. Dado que dichas palabras son consideradas en la lista de palabras vacías, al usar dicha lista se deja de considerar muchas palabras que tienen las frecuencias más altas por lo que el número de palabras totales pasó a disminuir en mayor proporción que el número de palabras distintas. Lo anterior ocurre en todos los años analizados y se puede apreciar en las siguientes gráficas:

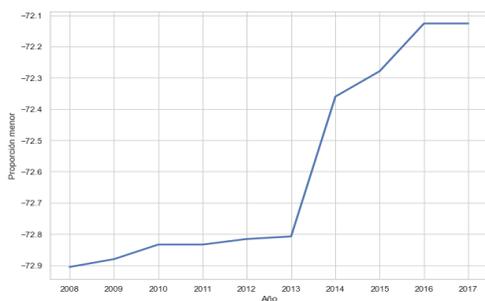


Figura 3.43: Cambios en el número de palabras totales

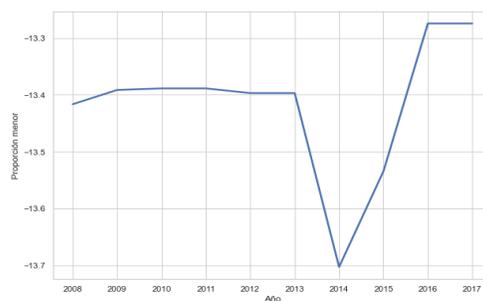


Figura 3.44: Cambios en el número de palabras distintas totales

Es notoria la diferencia en el número de palabras totales y el número de palabras distintas al emplear una lista de palabras vacías por lo que también se espera una diferencia notoria al calcular la entropía de las palabras en la LISR con y sin una lista de palabras vacías.

Por otro lado, las palabras más frecuentes tras usar la lista de palabras vacías pasaron a ser las siguientes:



Estas palabras conservan las primeras posiciones durante todo el periodo de tiempo analizado. Algunas de estas palabras, tales como “persona” o “ejercicio”, son palabras que se esperarían observar junto con otras formando términos compuestos de más de una palabra. Dichos términos si están contemplados en el presente análisis, como por ejemplo “persona moral”, pero son pocos los términos de este tipo con frecuencias altas, razón por la cual no aparecen en la nube de palabras anterior. Si se contemplan únicamente los términos compuestos por más de una palabra, los que tienen las mayores frecuencias por año son los siguientes:

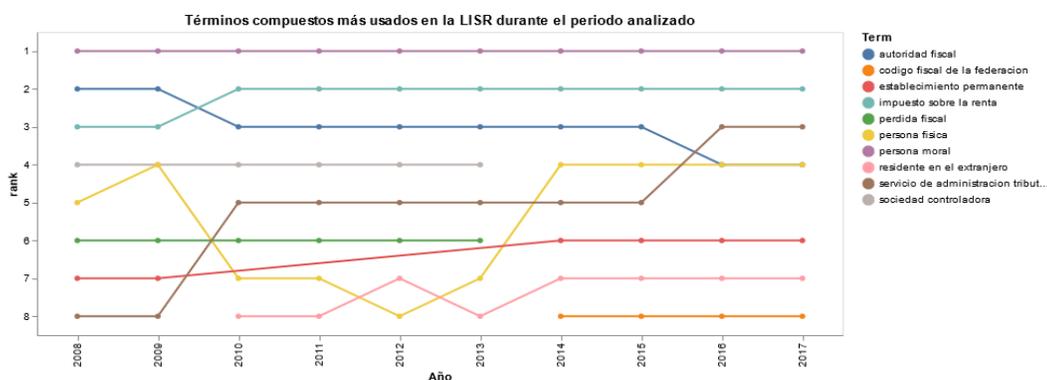


Figura 3.45: Términos compuestos más usados durante el periodo analizado

Cabe señalar que el número de palabras distintas obtenido también se verá afectado por este tipo de términos ya que la detección de este tipo de términos depende del software empleado así como de la persona que esté realizando el análisis y si contempla o no su propia lista de términos de este tipo. Lo anterior implica que la entropía de las palabras no será la misma si se contempla o no los términos de más de una palabra.

3.1.2.2. Entropía

Se calculó la entropía considerando los dos universos de palabras planteados anteriormente, uno en donde se consideran todas las palabras detectadas y uno donde sólo se contemplan las palabras que no formen parte de la lista de palabras vacías. Se obtuvo lo siguiente:

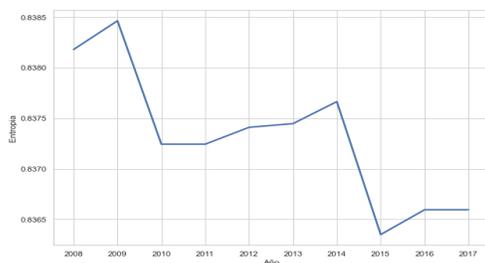


Figura 3.46: Entropía con todo el universo de palabras

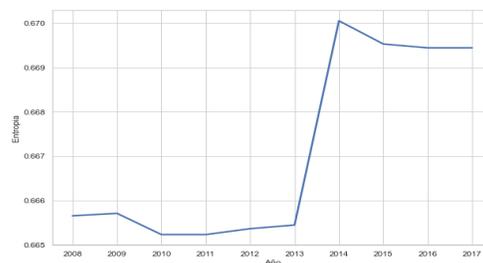


Figura 3.47: Entropía donde sólo se contemplan las palabras que no forman parte de la lista de palabras vacías.

Se observa que las entropías presentan una evolución distinta ya que una tiene una tendencia creciente mientras que la otra presentan una tendencia decreciente. Mientras que la entropía calculada con todas las palabras presentó un aumento notorio en su valor en el año 2014, la entropía calculada con un universo reducido de palabras presentó decrementos notorios en su valor en los años 2010 y 2015. Dado que el universo de palabras que contempla la lista de palabras vacías depende de lo estricta que sea la lista para conservar únicamente determinadas palabras, se optó por considerar la entropía calculada con base a todo el universo de palabras.

3.1.3. Interdependencias

3.1.3.1. Referencias internas

El número total de referencias internas detectadas en la LISR en el periodo analizado es el siguiente:

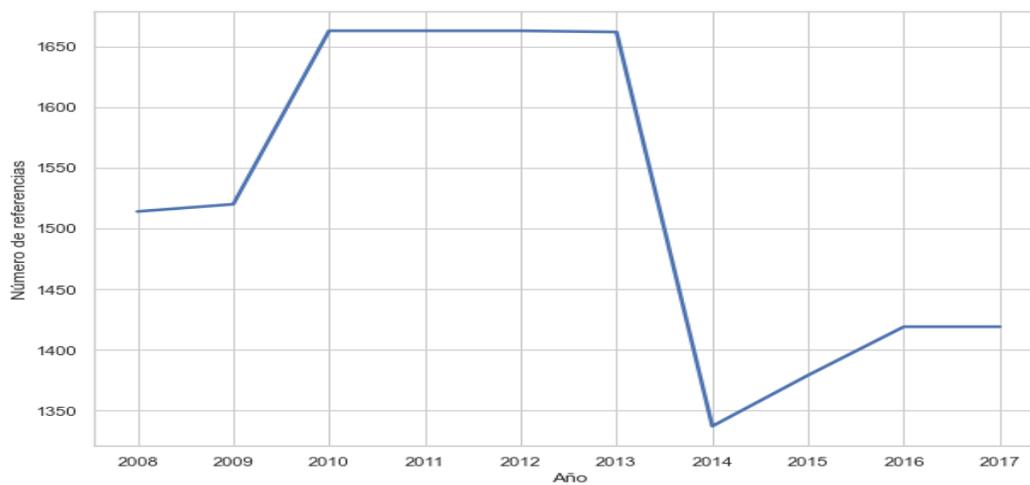


Figura 3.48: Referencias internas

Se observa que, al igual que el número de nodos y de palabras, el número de referencias internas disminuyó en el 2014 y con excepción de ese año presenta una tendencia creciente. Si se agrupan las referencias internas por el tipo de nodo que hace la referencia, se observa que en todos los tipos de nodo se refleja el cambio notorio ocurrido en 2014:

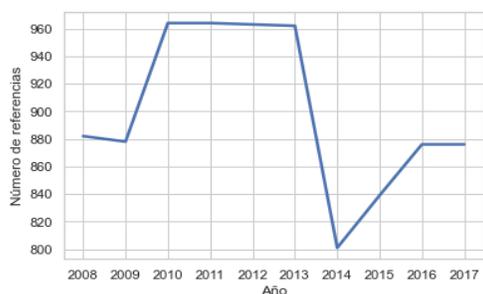


Figura 3.49: Referencias internas en nodos de tipo artículo

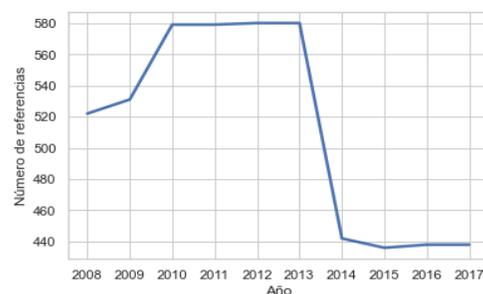


Figura 3.50: Referencias internas en nodos de tipo fracción

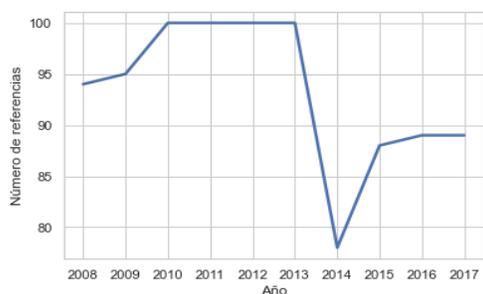


Figura 3.51: Referencias internas en nodos de tipo inciso

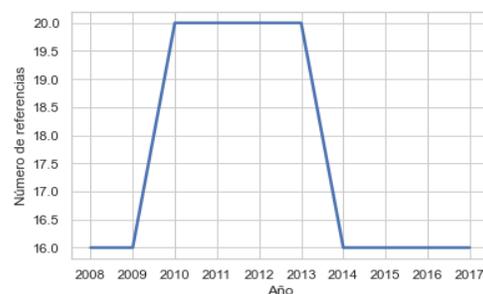


Figura 3.52: Referencias internas en nodos de tipo numeral

No obstante, si se agrupan las referencias internas por el tipo de nodo al que apunta la referencia, se podrá observar que los tipos de nodo título y capítulo presentaron una tendencia creciente incluso desde el 2014, año en el que los demás tipos de nodo presentaron una disminución en el número de referencias que apuntaban a ellos:

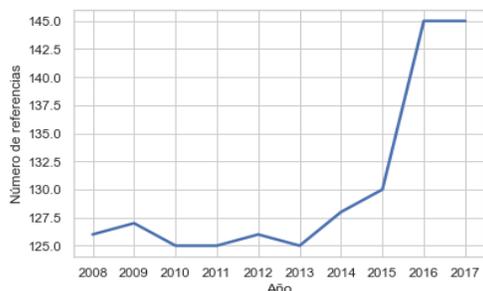


Figura 3.53: Referencias internas que apuntan a nodos de tipo título

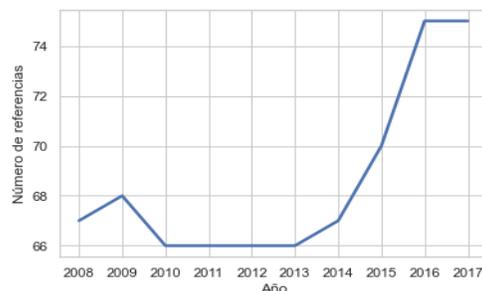


Figura 3.54: Referencias internas que apuntan a nodos de tipo capítulo

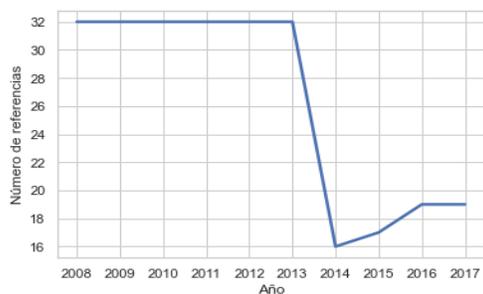


Figura 3.55: Referencias internas que apuntan a nodos de tipo sección

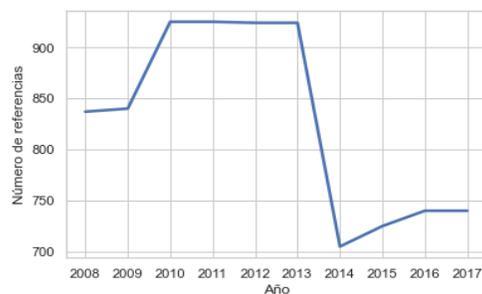


Figura 3.56: Referencias internas que apuntan a nodos de tipo artículo

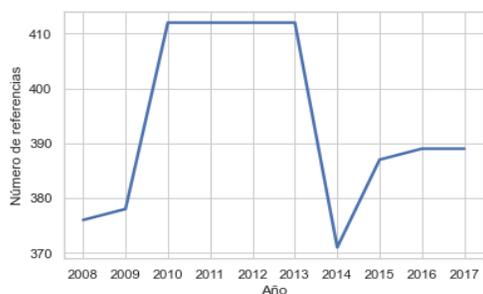


Figura 3.57: Referencias internas que apuntan a nodos de tipo fracción

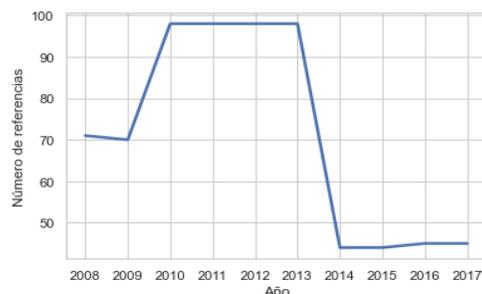


Figura 3.58: Referencias internas que apuntan a nodos de tipo inciso

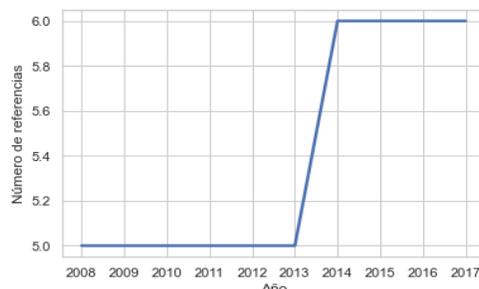


Figura 3.59: Referencias internas que apuntan a nodos de tipo numeral

Lo anterior indica que el cambio del 2014 hizo que las referencias internas apuntaran a divisiones menos profundas dentro de la ley, es decir, a divisiones con información menos específica en un tema en particular.

Por otro lado, durante el periodo analizado la gráfica que está compuesta únicamente por las referencias internas es la siguiente:

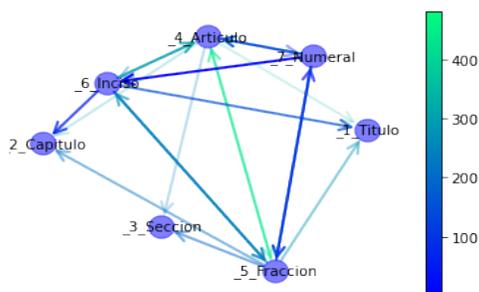


Figura 3.60: Gráfica representativa de referencias internas de 2008 a 2013

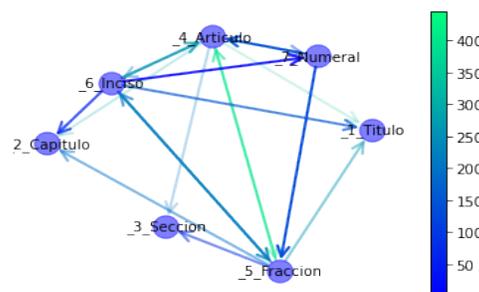


Figura 3.61: Gráfica representativa de referencias internas de 2014 a 2017

Si bien el número de referencias cambiaron con el pasar de los años, del 2013 al 2014 se da un cambio en la estructura de la red de referencias internas, desaparece el vínculo entre las fracciones y los numerales. Este es el único cambio que se da en la estructura.

3.1.3.2. Referencias externas

El número total de referencias externas detectadas en la LISR en el periodo analizado es el siguiente:

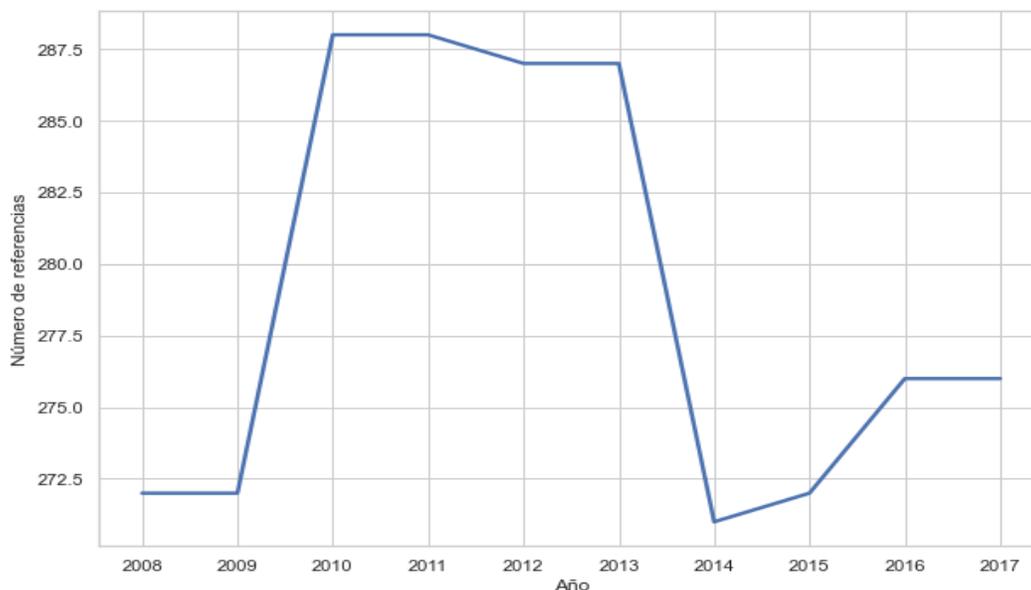


Figura 3.62: Referencias externas

Contrario a las referencias internas, las referencias externas presentan dos años con un cambio notorio, en el 2010 aumentaron y el 2014 disminuyeron, pero a diferencia de los cambios observados en el número de nodos o de palabras el cambio en las referencias externas hizo que el número de estas volviera a frecuencias muy cercanas a las obtenidas en el 2009 al disminuir en el 2014. Si se agrupan las referencias externas por el tipo de nodo que tiene la referencia, se obtiene lo siguiente:

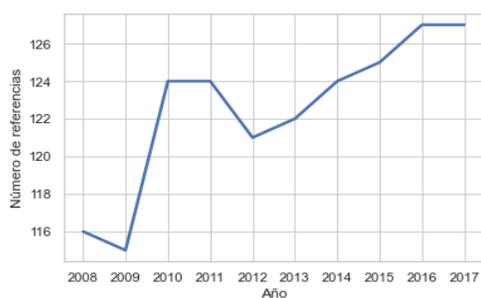


Figura 3.63: Referencias externas en nodos de tipo artículo

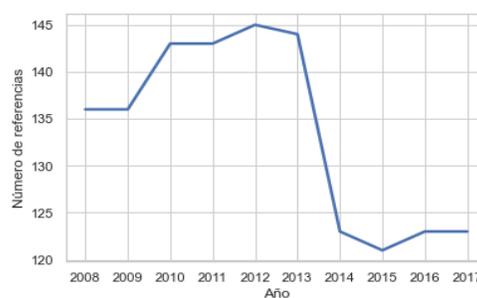


Figura 3.64: Referencias externas en nodos de tipo fracción

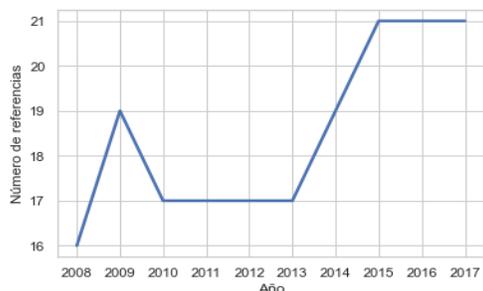


Figura 3.65: Referencias externas en nodos de tipo inciso

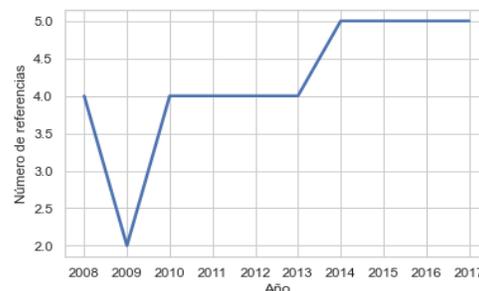


Figura 3.66: Referencias externas en nodos de tipo numeral

Se puede notar que, con excepción de las fracciones, los tipos de nodos presentaron un aumento en el número de referencias externas. Pese a lo anterior, las fracciones continuaron concentrando la mayor parte de las referencias externas. Esto indica que la mayoría de las referencias externas se emplean cuando se da información detallada sobre un tema. La siguiente imagen ejemplifica la red que estas referencias generan:

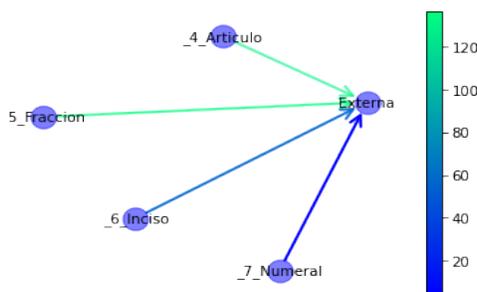


Figura 3.67: Gráfica representativa de referencias externas

Contrario a la estructura de la red de referencias internas, la estructura de la red de referencias externas no presentó cambios durante el periodo analizado.

3.2. Puntuaciones de complejidad con base en las métricas

Una vez obtenidos los valores de las propiedades consideradas, se calcularon las puntuaciones de complejidad propuestas en el capítulo anterior.

3.2.1. Puntuación de complejidad con base en la densidad de la gráfica de la LISR

Las puntuaciones de complejidad obtenidas con base en la densidad de la gráfica son las siguientes:

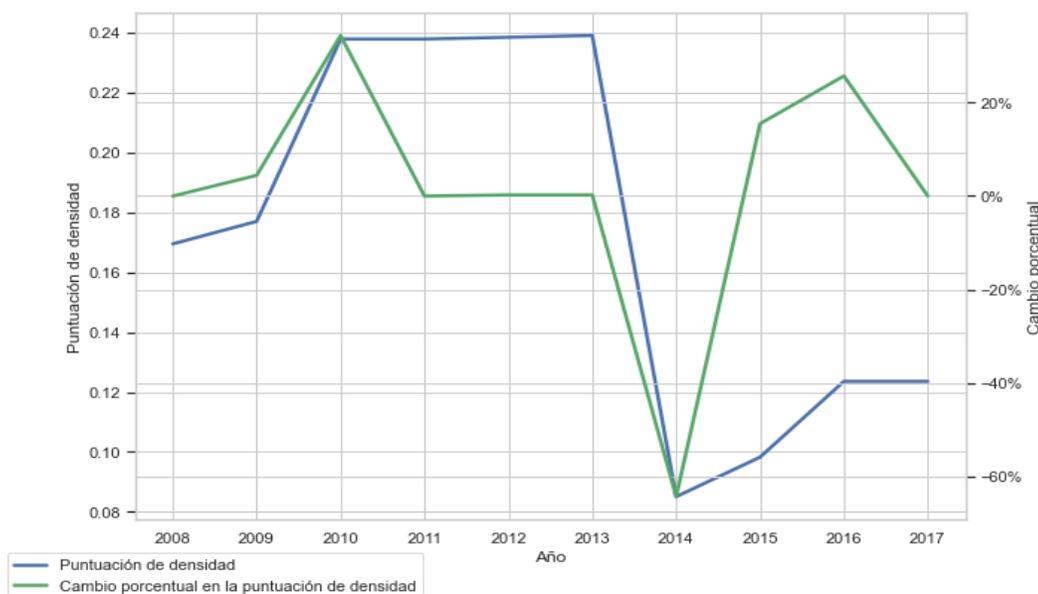


Figura 3.68: Puntuación de complejidad con base en la densidad

Se observa que el año con mayor puntuación de densidad es el 2013 mientras que el año con menor puntuación es el 2014. Con excepción del año 2014, se presenta una tendencia creciente en esta puntuación la cual es acorde al comportamiento creciente en el número de nodos que se observó anteriormente. Lo anterior indica que esta ley tiende a crecer en el número de nodos que la integran a menos que pase por una reforma que la reestructure considerablemente, como el evento del 2014.

3.2.2. Puntuación de complejidad con base en el grado de salida de los nodos de la LISR

Las puntuaciones de complejidad obtenidas con base en el grado de salida fueron las siguientes:

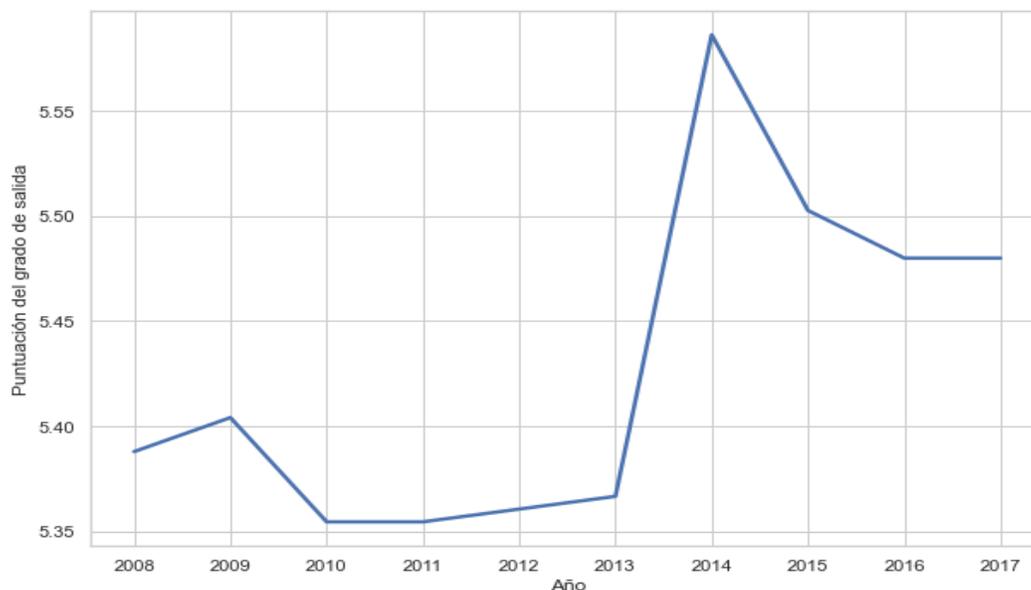


Figura 3.69: Puntuación de complejidad con base en el grado de salida

Se observa que el año con mayor puntuación de grado de salida es el año 2014, mientras que el año con la menor puntuación es el 2010 y el 2011. Contrario a lo observado en la puntuación de densidad, la puntuación del grado de salida pasó a presentar su valor más alto cuando la LISR tuvo el menor número de nodos. Lo anterior indica que el cambio notorio en la estructura de la LISR que se dio del 2013 al 2014 hizo que cambiara la distribución del grado de salida en los distintos tipos de nodos de tal forma que, con base en la puntuación propuesta, se hizo más compleja pese a que el número de nodos decreció. También marca un cambio en la tendencia que se presentaba en los años anteriores al 2014 ya que, con excepción del 2010, se presentó una tendencia creciente en la puntuación propuesta pero después del 2014 se presentó una tendencia decreciente. Por otro lado, se construyó una variante de esta puntuación de complejidad, la puntuación del grado de salida redondeada al entero más cercano. Tiene por objetivo indicar un grado de salida representativo por año que pertenezca al dominio del grado de salida de los nodos. Con dicha puntuación se obtuvieron los siguientes resultados:

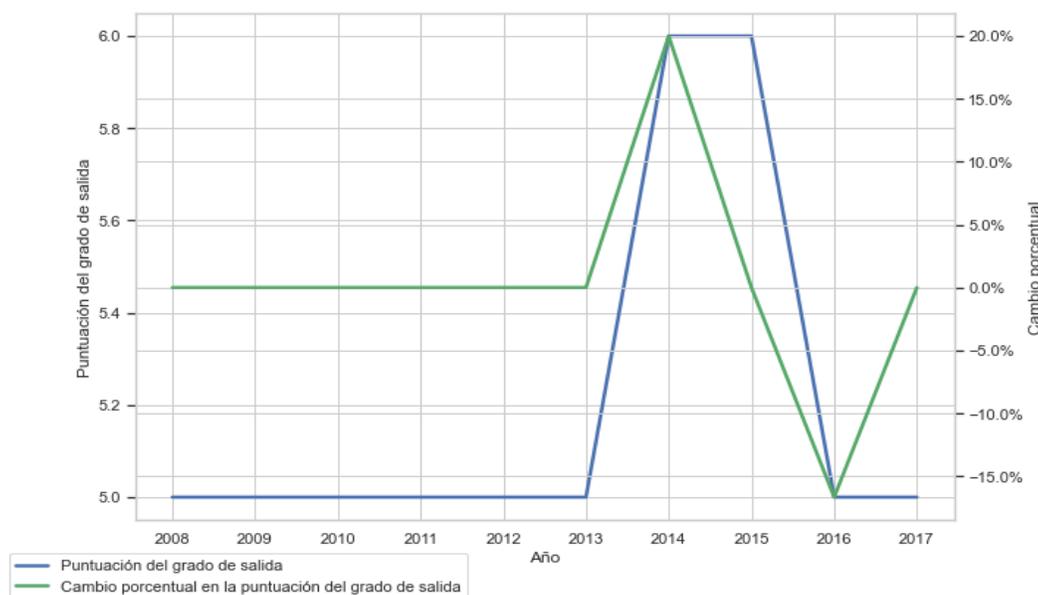


Figura 3.70: Puntuación de complejidad redondeada con base en el grado de salida

Se observó que los años con mayor puntuación de grado de salida son los años 2014 y 2015 mientras que los demás años mantienen la misma puntuación de grado de salida por lo que sólo en los años mencionados la distribución del grado de salida cambió lo suficiente como para que el grado de salida representativo aumentara. Lo anterior indica que durante estos años existieron más divisiones en los tipos de nodos con mayor proporción dentro de la ley. Lo anterior indica una presencia mayor de nodos que mencionan información más específica de algún punto en particular, es decir, se hace más énfasis en la información de algunos nodos en particular.

3.2.3. Puntuación de complejidad con base en la profundidad de los nodos hoja de la LISR

Con base en la fórmula de la puntuación de complejidad propuesta se obtuvo lo siguiente:

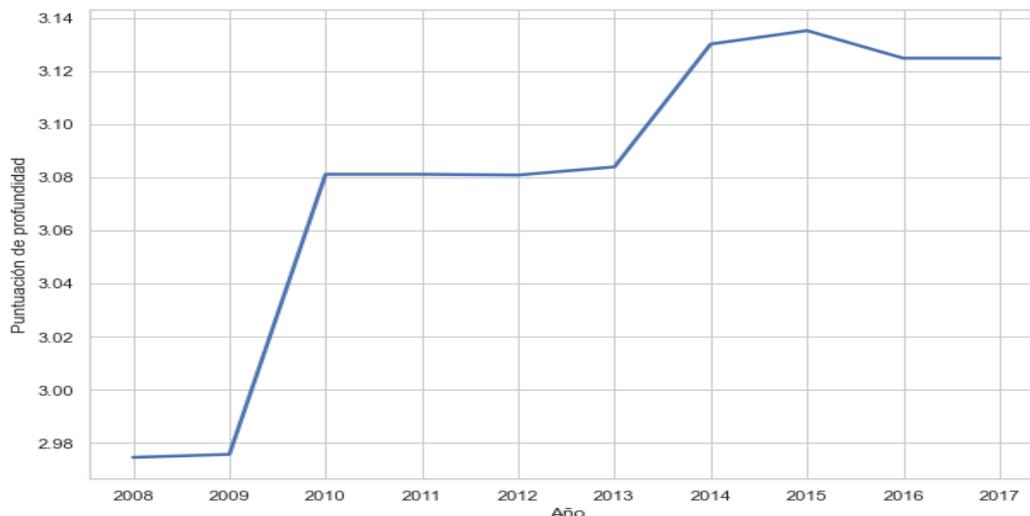


Figura 3.71: Puntuación de complejidad con base en la profundidad

Con excepción de los últimos dos años la puntuación presenta una tendencia creciente y contrario a las dos puntuaciones anteriores, no se presenta un cambio en la evolución de la puntuación entre los años 2013 y 2014. La tendencia creciente indica que aumentó la proporción de los nodos cuyas profundidades son mayor a 3 por lo que se fue haciendo más énfasis en algunos puntos en específico de la ley. No obstante, los valores obtenidos no se alejan lo suficiente del 3 como para que, al ser redondeados al entero más cercano, haya una que alcance un valor igual a 4. Esto se puede apreciar en la gráfica donde se redondea al entero más cercano la puntuación.

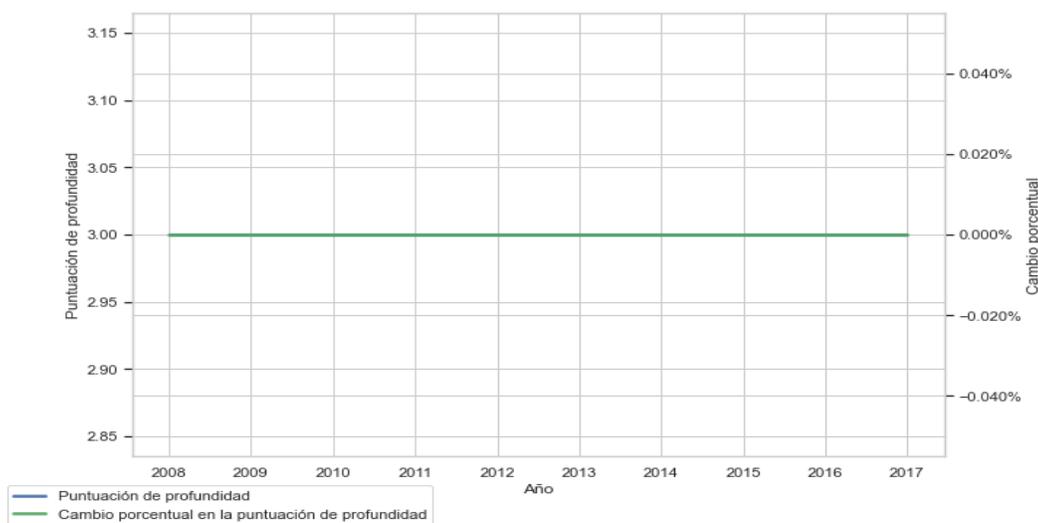


Figura 3.72: Puntuación de complejidad redondeada con base en la profundidad

Similar a la puntuación anterior, el objetivo fue indicar una profundidad representativa por año que pertenezca al dominio de la profundidad de los nodos. Si se recuerdan las distintas divisiones que integran la ley y se contempla la puntuación obtenida se podrá notar que la puntuación indica que los nodos hoja tienden a ser artículos, dado que no hay secciones que no tengan artículos y existen artículos que no pertenecen a una sección, o fracciones cuyos artículos no pertenecen a una sección ni a un capítulo.

3.2.4. Puntuación de complejidad con base en la entropía de las palabras

Anteriormente ya se había observado que la entropía calculada con todo el universo de palabras presenta un cambio notorio en el 2014, año en el que varias de las puntuaciones de estructura también presentaron un cambio.

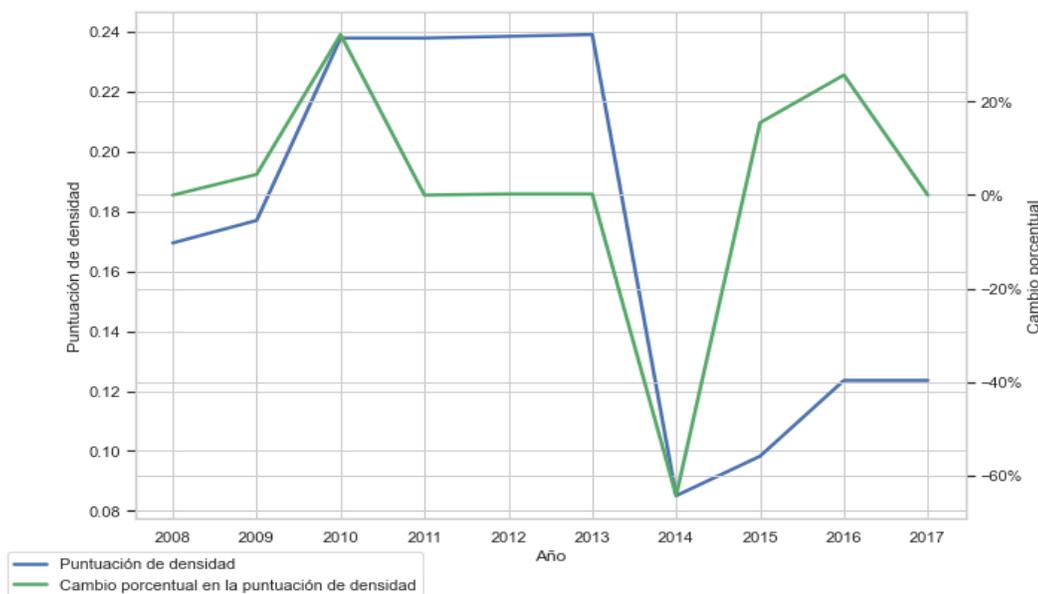


Figura 3.73: Puntuación de complejidad con base en la entropía

Además, presenta la tendencia creciente con excepción del 2014. Recuérdese que el valor obtenido de la entropía no se transformó para considerarlo como la puntuación de complejidad.

3.2.5. Puntuación de complejidad con base en las puntuaciones de Hub y Authority obtenidas con el algoritmo HITS

Las puntuaciones Hub obtenidas para las referencias internas son las siguientes:

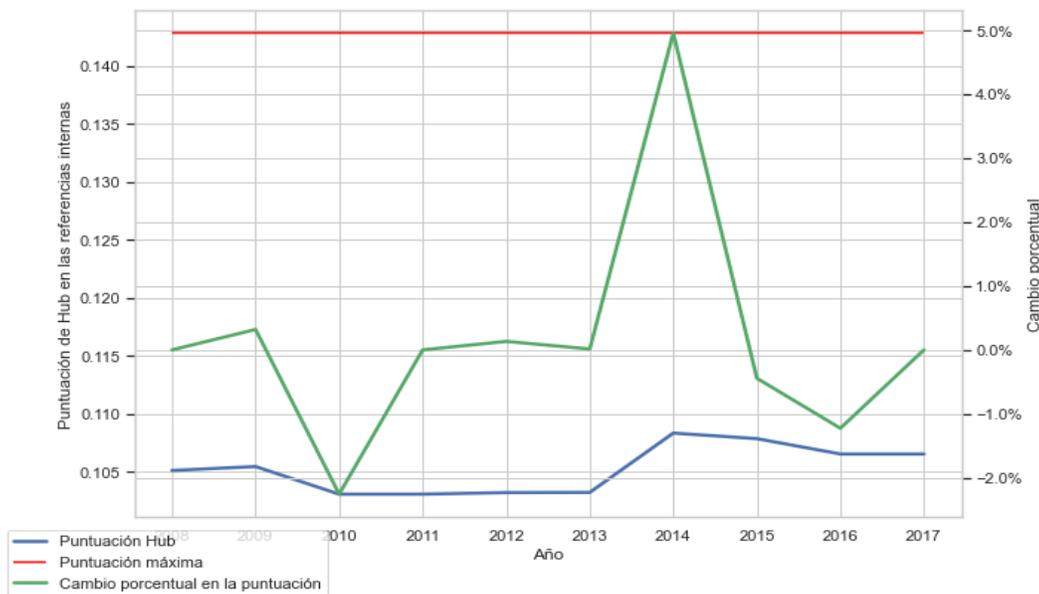


Figura 3.74: Puntuación Hub en las referencias internas

El periodo de tiempo en el que esta puntuación presenta sus valores más bajos es durante los años en los que se presentaron los mayores números de referencias internas en la LISR. Por otro lado, se observa que las puntuaciones tienen un valor similar a dos tercios de aquellas que se obtendrían si todos los tipos de nodo tuviesen la misma importancia como Hub. Esto indica que la importancia como Hub de los tipos de nodo que contenían las referencias internas no es la misma entre ellos, por lo que durante los años en los que existieron más referencias la diferencia en la importancia como Hub de los tipos de nodo fue mayor. Por otro lado, en 2014 se observa el mayor valor de la puntuación. Lo anterior indica que el cambio en la LISR del 2014 hizo que la diferencia en la importancia como Hub entre los distintos tipos de nodo fuese menor en este año, siendo la menor durante el periodo analizado.

Las puntuaciones Hub, por tipo de nodo, obtenidas son las siguientes:

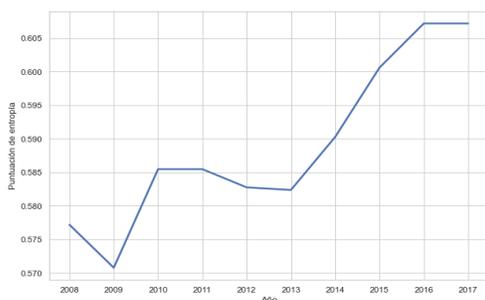


Figura 3.75: Puntuación Hub de artículos

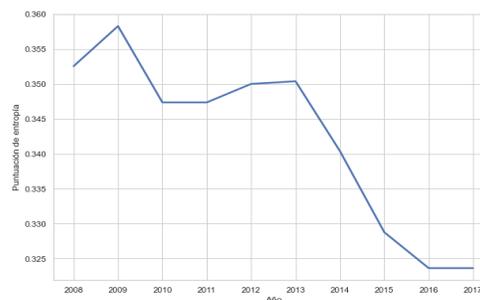


Figura 3.76: Puntuación Hub de fracciones

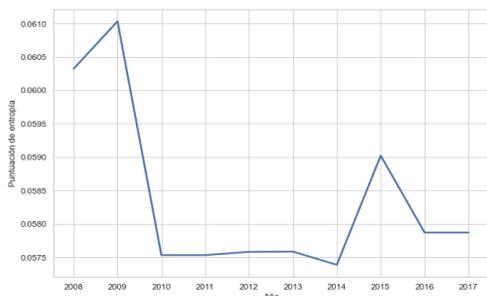


Figura 3.77: Puntuación Hub de incisos

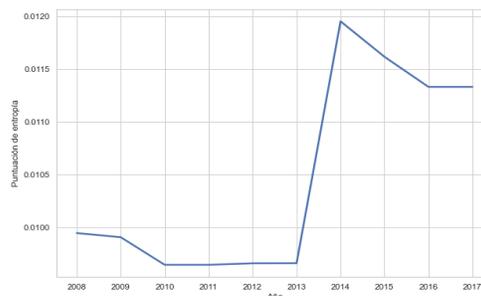


Figura 3.78: Puntuación Hub de numerales

Es muy notoria la diferencia entre las puntuaciones obtenidas en los artículos contra las obtenidas en incisos o numerales, ya que los primeros tienen una puntuación que varía entre .057 y .61 mientras que los últimos tienen puntuaciones menores a .7 y .2, respectivamente. Esto indica que el artículo es el tipo de nodo que tiende a tener referencias a nodos cuya puntuación authority es alta. También se corrobora lo mencionado a las diferencias que hubo entre las puntuaciones Hub de los distintos tipos de nodo.

Por otro lado, las puntuaciones Authority calculadas son las siguientes:

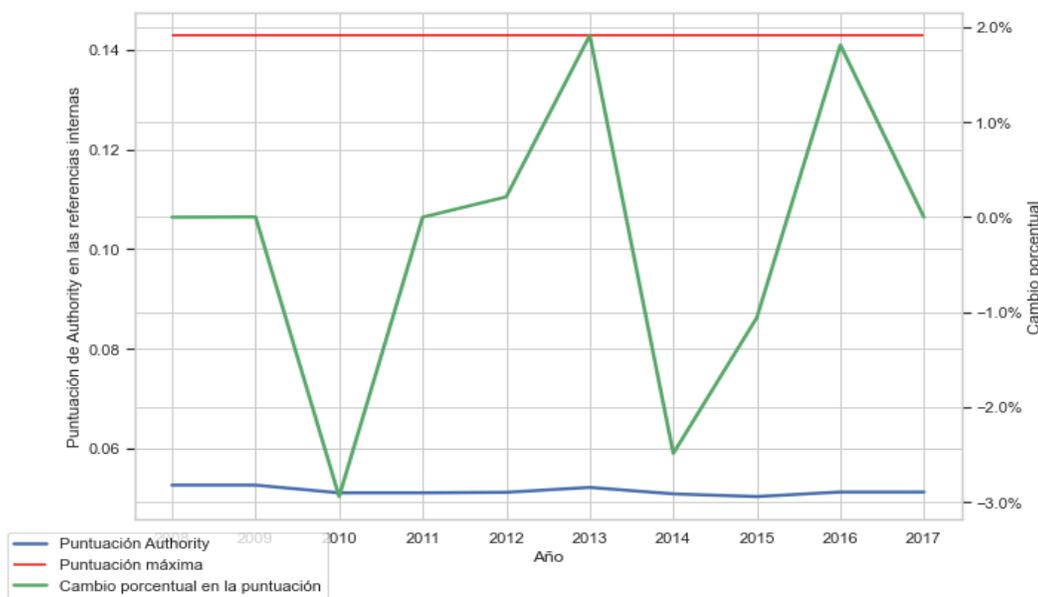


Figura 3.79: Puntuación Authority en las referencias internas

Al igual que en la puntuación Hub, en todos los años la puntuación Authority es menor que aquella que se obtendría si todos los tipos de nodo tuviesen la misma importancia como Authority, es menos de la mitad. Además, el cambio del 2014 en la LISR que había afectado notoriamente a las demás puntuaciones parece no haber afectado notoriamente esta puntuación dado que su valor mínimo se da en el 2015, año con el segundo menor número de referencias internas. Esto indica que la puntuación Authority no se vio afectada notoriamente por el número de referencias internas existentes como en la puntuación Hub.

Lo observado anteriormente indica que los tipos de nodo que contenían las referencias internas presentaron una mayor diferencia en su importancia como Authority que como Hub.

Las puntuaciones Authority, por tipo de nodo, obtenidas son las siguientes:

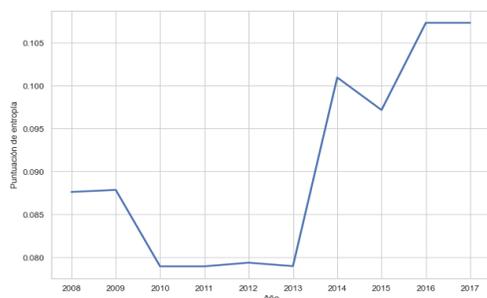


Figura 3.80: Puntuación Authority de títulos

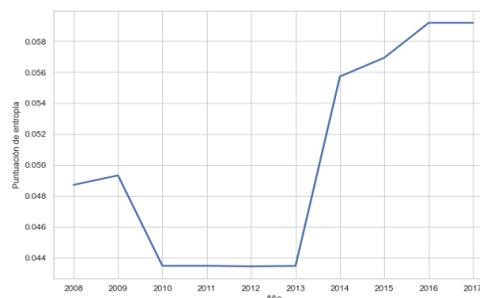


Figura 3.81: Puntuación Authority de capítulos

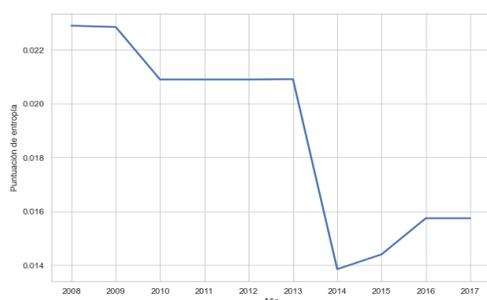


Figura 3.82: Puntuación Authority de secciones

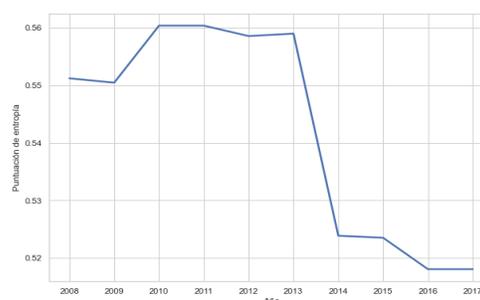


Figura 3.83: Puntuación Authority de artículos

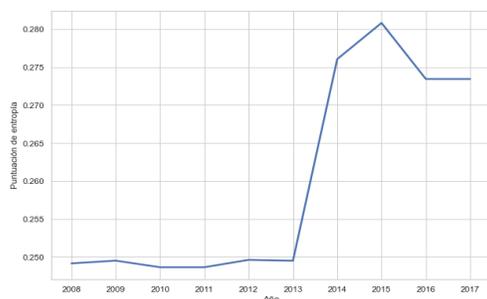


Figura 3.84: Puntuación Authority de fracciones

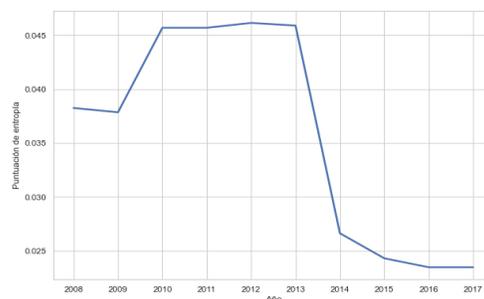


Figura 3.85: Puntuación Authority de incisos

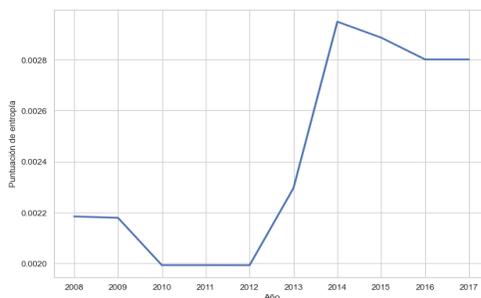


Figura 3.86: Puntuación Authority de numerales

Al igual que lo observado en las puntuaciones Hub, los artículos son el tipo de nodo con mayor puntuación y la diferencia entre la puntuación de estos y la puntuación de otros tipos de nodo, tales como numerales o incisos, es notoria. Contrario a lo observado en la puntuación Authority agregada, las puntuaciones por tipo de nodo se ven afectadas por el cambio del 2014, aspecto que también afectó a otras puntuaciones. Cabe señalar que algunas puntuaciones aumentaron mientras que otras disminuyeron en dicho año.

Por su parte, las puntuaciones de Hub de las referencias externas fueron las siguientes:

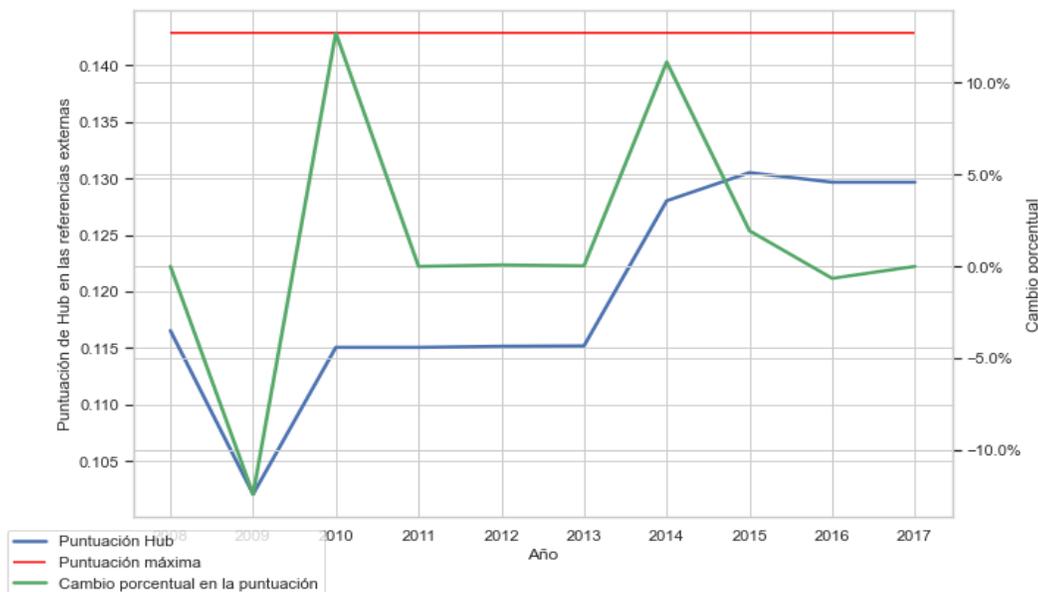


Figura 3.87: Puntuación de Hub en las referencias externas

Se observa que esta puntuación presentó un incremento el 2014, detalle que contrasta a lo observado en el número de referencias externas ya que en este año presentó su valor más bajo. También se detecta que las puntuaciones son menores que aquella que se obtendría si todos los tipos de nodo tuviesen la misma importancia como Hub pero es la que presenta valores más cercanos a la puntuación máxima y al igual que en las puntuaciones anteriores, se observa que los tipo de nodo artículo y fracción presentan las mayores puntuaciones aunque la diferencia entre las puntuaciones de los artículos y las fracciones no son tan diferentes como en las anteriores:

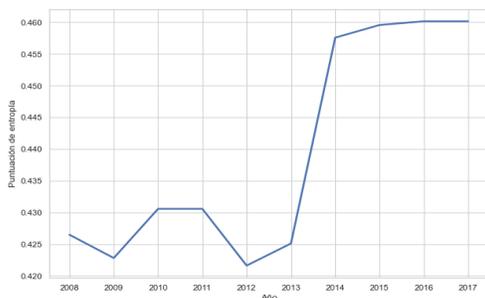


Figura 3.88: Puntuación Hub de fracciones

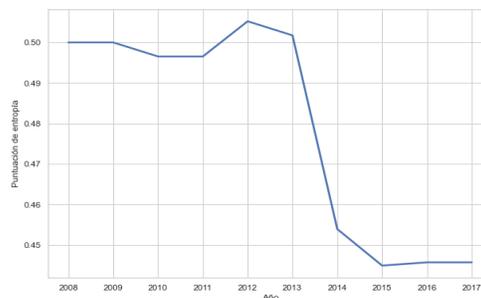


Figura 3.89: Puntuación Hub de incisiones

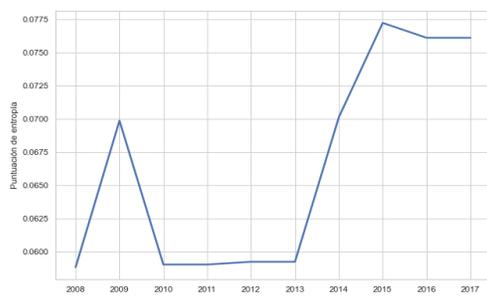


Figura 3.90: Puntuación Hub de fracciones

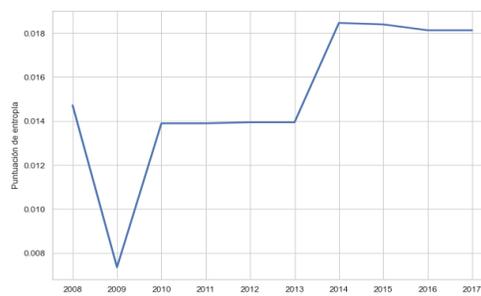


Figura 3.91: Puntuación Hub de incisos

Lo anterior indica que la importancia como Hub de referencias externas es similar entre artículos y fracciones. Esta característica de presentar las mayores puntuaciones es una característica que los artículos concentraban en las dos puntuaciones de referencias internas.

3.2.6. Puntuación de complejidad de la estructura de la LISR

Con base en las puntuaciones de métricas relativas a la estructura de la LISR, se obtuvieron las siguientes puntuaciones de complejidad de estructura:

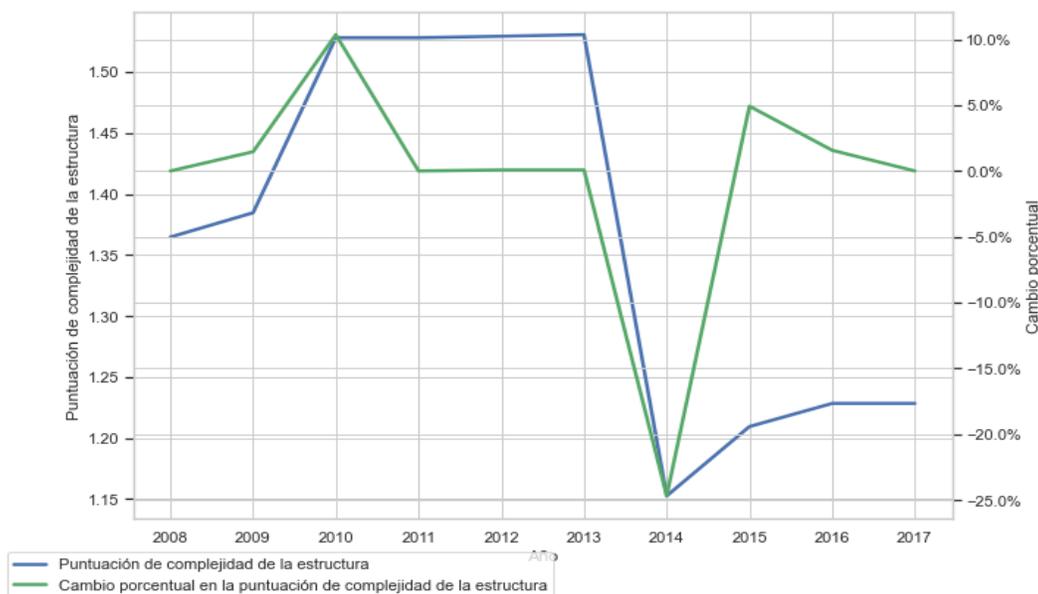


Figura 3.92: Puntuación de complejidad de estructura de la LISR

No resulta extraño ver que esta puntuación presenta un cambio en el 2014 dado que dos de las tres métricas de estructura presentaron un cambio en dicho año. No obstante, una de dichas métricas, la del grado de salida, presentó un aumento en su valor mientras que la otra, la de densidad, presentó una disminución como lo observado en la gráfica anterior. Lo anterior genera la impresión de que tuvo un mayor peso en la puntuación la disminución en la puntuación de densidad que el aumento en la puntuación de grado de salida. Esto se puede validar con la siguiente tabla que contiene las correlaciones de Pearson y de Spearman:

	Correlación de Pearson	P-Valor de Correlación de Pearson	Correlación de Spearman	P-Valor de Correlación de Spearman
Puntuación de complejidad estructural y puntuación de densidad	0.996815	4.49E-10	1.000000	0.000000
Puntuación de complejidad estructural y puntuación de grado de salida	-0.639766	0.046361	-0.700569	0.024041
Puntuación de complejidad estructural y puntuación de profundidad	NA	NA	NA	NA

Figura 3.93: Correlaciones entre la puntuación de estructura y las puntuación que la conforman

Con lo anterior se puede afirmar que pese a que la reestructuración del 2014 hizo que la LISR tuviese una mayor presencia de nodos que mencionan información más específica de algún punto en particular, la disminución en el número de nodos fue mayor y la hizo menos compleja.

3.2.7. Puntuación de complejidad del lenguaje de la LISR

Con base en la puntuación de la métrica relativa al lenguaje en la LISR, se obtuvo la siguiente puntuación de complejidad de lenguaje:

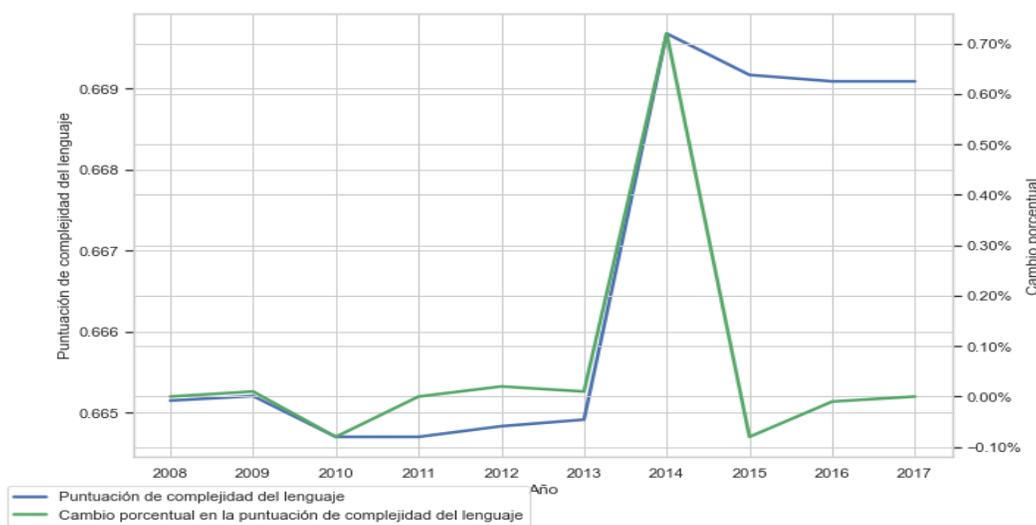


Figura 3.94: Puntuación de complejidad del lenguaje en la LISR

Esta puntuación es igual a la entropía dado que depende únicamente de ella y no se le aplicó transformación alguna.

3.2.8. Puntuación de complejidad de las interdependencias en la LISR

Con base en las puntuaciones de las métricas relativas a las interdependencias en la LISR, se obtuvieron las siguientes puntuaciones de complejidad de interdependencia:

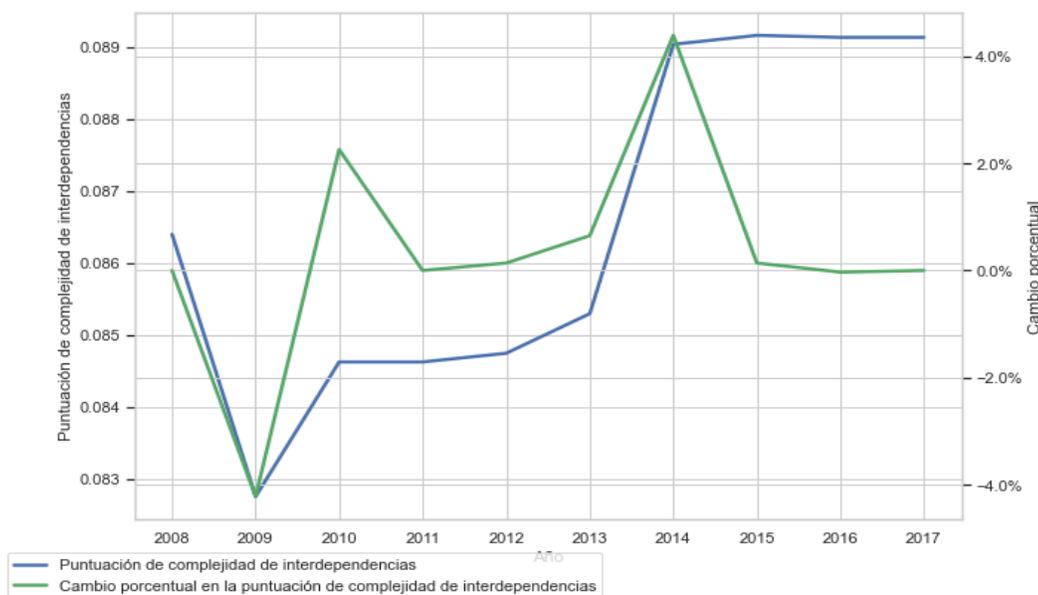


Figura 3.95: Puntuación de complejidad de interdependencias en la LISR

Esta puntuación mantiene un aspecto observado en las puntuaciones que la conforman el cual es el aumento relativamente notorio del 2013 al 2014, año en el que se presentó el único cambio en la estructura de referencias internas en la LISR. Esto indica que esta puntuación fue más sensible a cambios en la estructura de la red de referencias internas y externas que al número total de referencias internas y externas.

3.2.9. Puntuación de complejidad legal de la LISR

Las puntuaciones de complejidad legal obtenidas durante el periodo analizado son las siguientes:

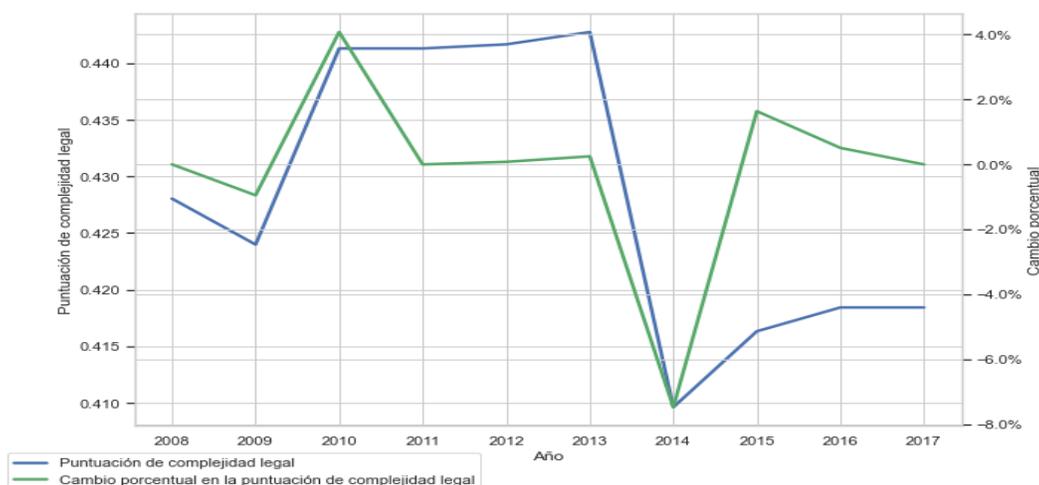


Figura 3.96: Puntuación de complejidad

Se observa que la puntuación de complejidad legal tiene un comportamiento similar al observado en la puntuación de complejidad de la estructura dado que ambas presentan sus valores más altos desde el 2010 hasta el 2013 y su valor más bajo se presenta en el 2014, año en el ocurrió la reestructuración de la LISR. Lo anterior indica que la puntuación de complejidad legal que se construyó se vio más influencia por las características de la estructura de la ley que por aquellas enfocadas en su contenido. También se observa que en la mayoría de los años la puntuación tiene una tendencia creciente, característica que se da tanto en la puntuación de complejidad de estructura así como en la puntuación de complejidad de interdependencias. Lo anterior indica que los cambios que se aplican a la ley, tales como las reformas, tienden a aumentar su complejidad. En las siguientes gráficas se puede visualizar la evolución de las tres puntuaciones que integran la puntuación de complejidad:

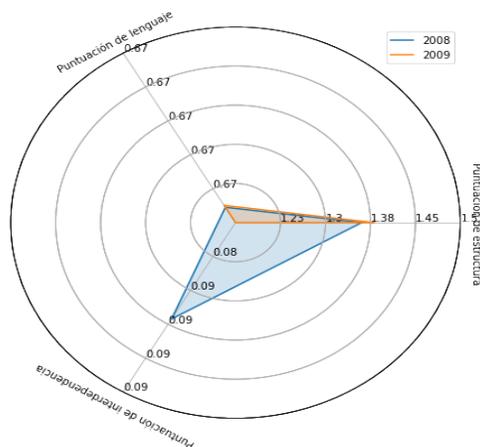


Figura 3.97: Puntuación de complejidad legal de 2008 y 2009

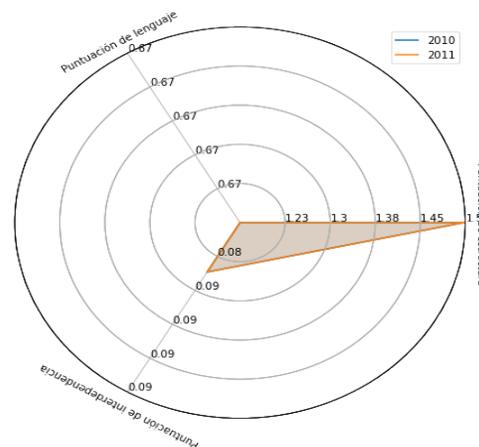


Figura 3.98: Puntuación de complejidad legal de 2010 y 2011

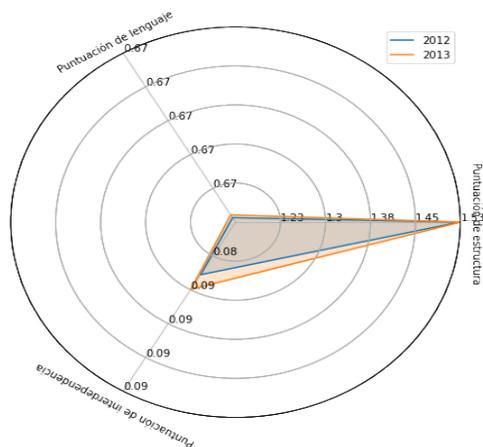


Figura 3.99: Puntuación de complejidad legal de 2012 y 2013

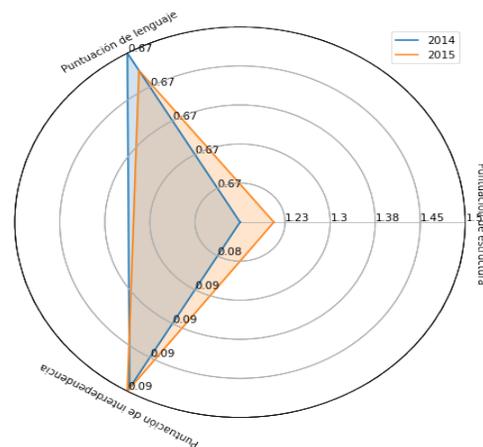


Figura 3.100: Puntuación de complejidad legal de 2014 y 2015

Se observa que de 2008 a 2009 bajó la puntuación de complejidad del lenguaje y la de interdependencias lo cual hizo que la puntuación de estructura fuese la puntuación preponderante durante este último año. Durante los dos años siguientes aumentó tanto la puntuación de complejidad de estructura como la de interdependencias no obstante se mantiene la de estructura como la puntuación preponderante. Ocurre algo similar en los siguientes años, 2012 y 2013, ya que crecieron las puntuaciones de complejidad de interdependencias así como la de lenguaje pero se mantiene la de estructura como la puntuación preponderante.

No obstante, en el 2014 cambian drásticamente las puntuaciones de complejidad preponderantes ya que disminuyó la de estructura mientras que las otras dos aumentaron haciendo que las puntuaciones preponderantes sean la de lenguaje y la de interdependencias:

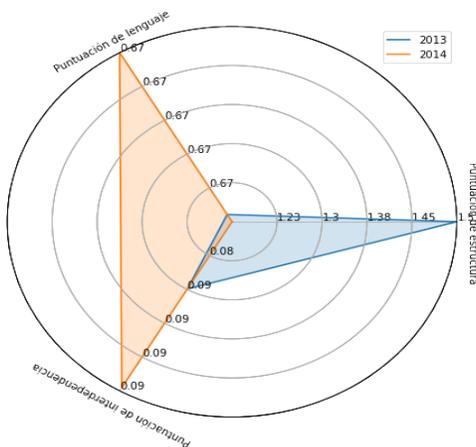


Figura 3.101: Cambio de puntuaciones preponderantes

Lo anterior indica que la derogación del IETU y el IDE hizo que la LISR tuviese una estructura menos compleja pero el contenido de los artículos se hizo más complejo, tanto por las referencias internas y externas que contienen así como por el texto que contienen. Estas puntuaciones preponderantes se mantuvieron durante los últimos dos años analizados pese a que la puntuación de complejidad del lenguaje disminuyó y aumentó la de estructura:

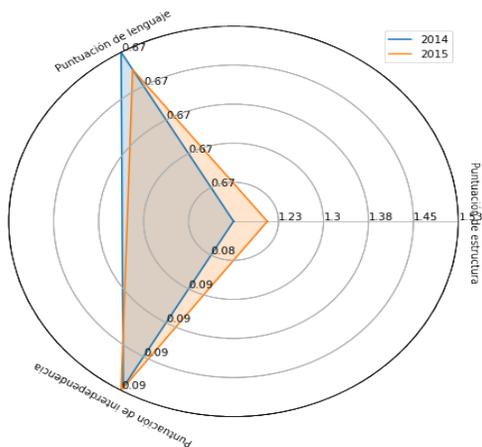


Figura 3.102: Puntuación de complejidad legal de 2014 y 2015

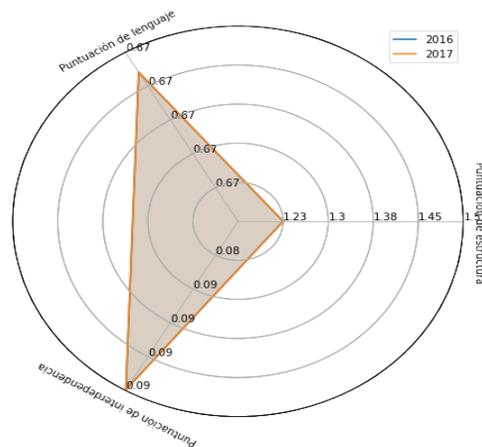


Figura 3.103: Puntuación de complejidad legal de 2016 y 2017

Cabe señalar que la característica cuyas correlaciones fueron las más grandes, en valor absoluto, con la puntuación de complejidad legal fue la puntuación de complejidad de estructura, esto se puede observar en la siguiente tabla:

	Correlación de Pearson	P-Valor de Correlación de Pearson	Correlación de Spearman	P-Valor de Correlación de Spearman
Puntuación de complejidad legal y puntuación de complejidad estructural	0.985945	1.68E-07	0.987730	9.77E-08
Puntuación de complejidad legal y puntuación de complejidad del lenguaje	-0.893002	0.000503	-0.889571	0.000568
Puntuación de complejidad legal y puntuación de complejidad de las interdependencias	-0.743316	0.013745	-0.631902	0.049998

Figura 3.104: Correlaciones entre las puntuaciones de complejidad y la puntuación de complejidad legal

Además, tanto el coeficiente de correlación de Pearson como el coeficiente de correlación de Spearman son significativos. Por otro lado, la puntuación de complejidad de las interdependencias fue la que tuvo la menor correlación con la puntuación de complejidad legal, tanto con la correlación de Pearson como con la correlación de Spearman.

Finalmente, la puntuación de complejidad de estructura y la puntuación de complejidad del lenguaje fueron las puntuaciones que más se correlacionaron entre las 3 puntuaciones que conforman la puntuación de complejidad legal. Esto se puede observar en la siguiente tabla:

	Correlación de Pearson	P-Valor de Correlación de Pearson	Correlación de Spearman	P-Valor de Correlación de Spearman
Puntuación de complejidad estructural y puntuación de complejidad del lenguaje	-0.942940	4.33E-05	-0.877301	0.000853
Puntuación de complejidad estructural y puntuación de complejidad de las interdependencias	-0.843895	0.002142	-0.693252	0.026219
Puntuación de complejidad del lenguaje y puntuación de complejidad de las interdependencias	0.922680	0.000142	0.742331	0.013939

Figura 3.105: Correlaciones entre las puntuaciones de complejidad y la puntuación de complejidad legal

Los datos anteriores indican que, con base en las puntuaciones de complejidad construidas, entre más compleja sea la estructura de la LISR menos complejo será el texto que contiene así como la red de interdependencias que genera. Por otro lado, entre más complejo sea el lenguaje de la LISR, más compleja será la red de interdependencias de la LISR.

3.2.10. La puntuación de complejidad legal y la recaudación obtenida con el ISR

A continuación se puede observar la recaudación obtenida con el ISR¹ durante el periodo analizado:

¹Datos obtenidos de la página: http://omawww.sat.gob.mx/cifras_sat/Paginas/inicio.html

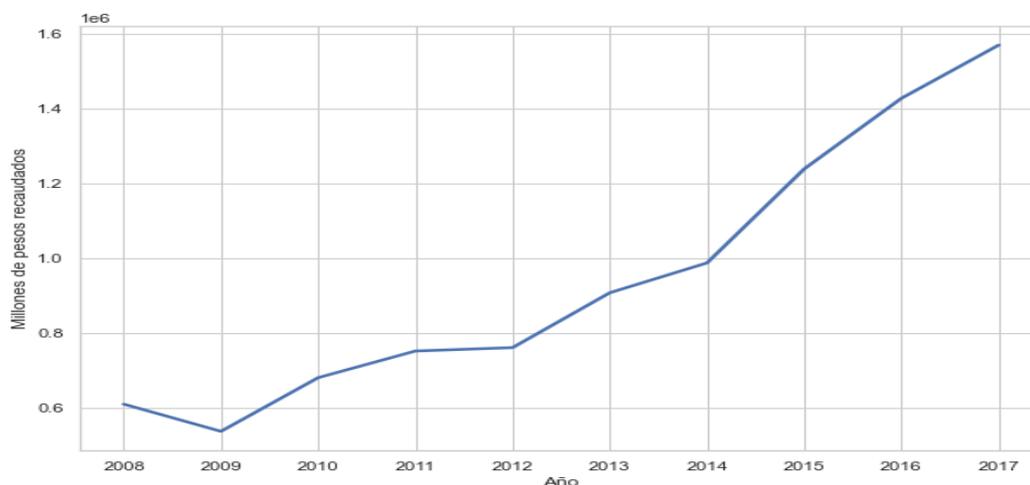


Figura 3.106: Recaudación obtenida con el ISR.

Se observa que la recaudación obtenida con el ISR tiene una tendencia creciente con excepción del 2009, característica que comparte con el universo de contribuyentes registrados:

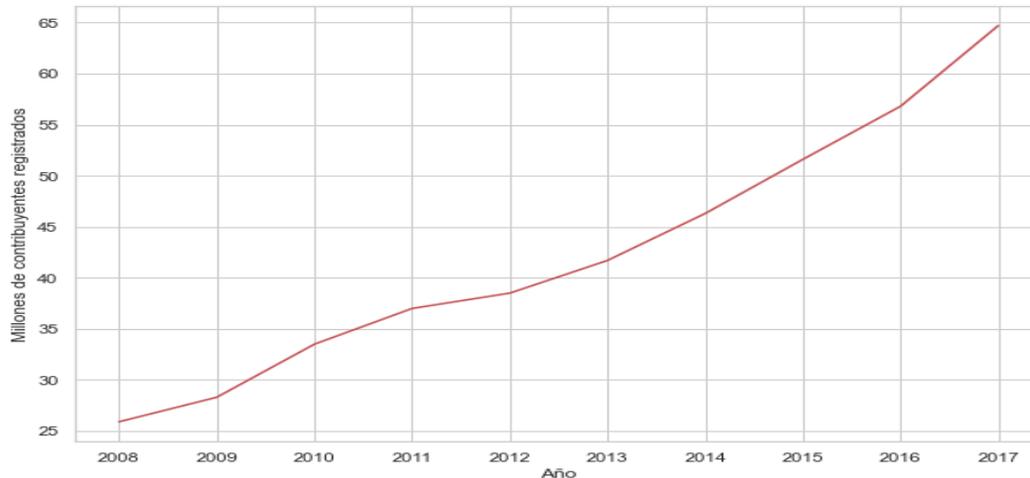


Figura 3.107: Universo de contribuyentes por año

No obstante, el monto recaudado por contribuyente pasa a presentar una tendencia no monótona dado que en varios años se presentan decrementos pero mantiene una tendencia creciente con después del 2008, aspecto que se puede observar a continuación:

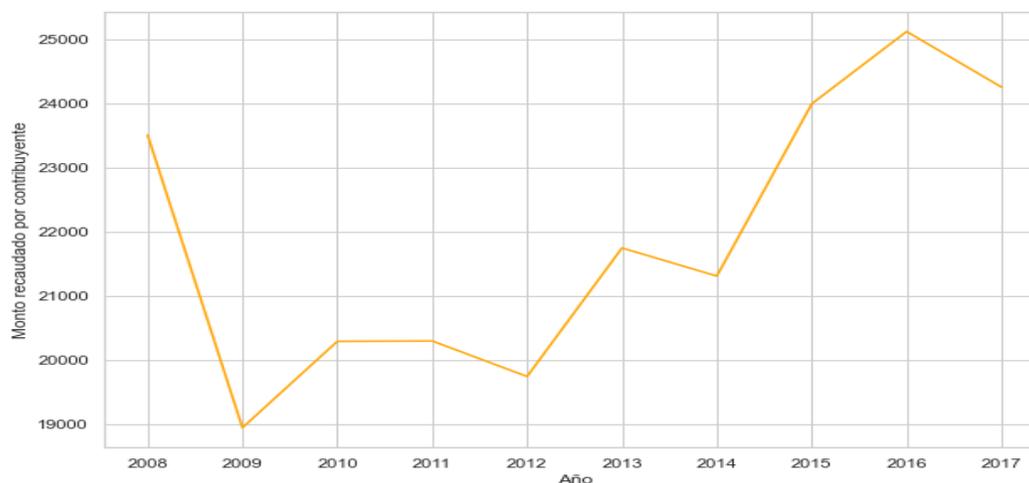


Figura 3.108: Recaudación por contribuyentes

Tras observar las gráficas anteriores, se pudo observar que la puntuación de complejidad propuesta aparentemente no presenta alguna relación significativa con el monto recaudado o el universo de contribuyentes registrado. La siguiente tabla muestra el coeficiente de correlación de Pearson y de Spearman obtenido entre la puntuación de complejidad legal y el monto recaudado así como con el universo de contribuyentes:

	Correlación de Pearson	P-Valor de Correlación de Pearson	Correlación de Spearman	P-Valor de Correlación de Spearman
Puntuación de complejidad legal y monto recaudado	-0.554825	0.095981	-0.469521	0.170966
Puntuación de complejidad legal y universo de contribuyentes	-0.514527	0.128095	-0.481716	0.158606

Figura 3.109: Correlaciones entre la puntuación de complejidad legal y el monto recaudado así como con el universo de contribuyentes

Los datos anteriores indican que ninguna correlación es significativa, lo cual afirma la hipótesis del párrafo anterior de que no presenta alguna relación significativa la puntuación de complejidad legal construida y el monto recaudado o el universo de contribuyentes registrado.

3.2.11. La complejidad legal y algunos servicios proporcionados por la PRODECON

La Procuraduría de la Defensa del Contribuyente (PRODECON), es un organismo público especializado en materia tributaria el cual proporciona de forma gratuita

servicios de orientación, asesoría, consulta, representación y defensa legal, investigación, recepción y trámite de quejas y reclamaciones contra actos u omisiones de las autoridades fiscales federales. El 1 de septiembre del 2011 abrió sus puertas al público. En la siguiente gráfica se puede visualizar el número total de servicios prestados por año por este organismo:

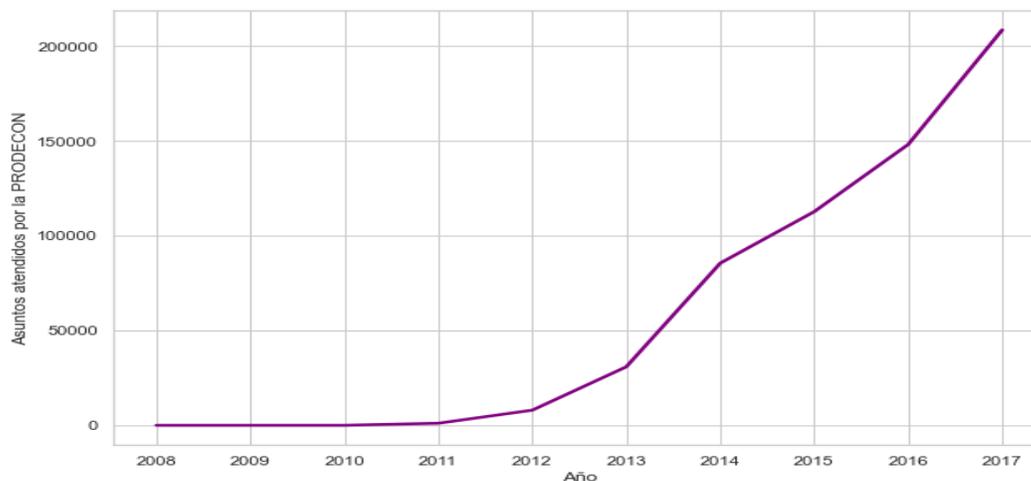


Figura 3.110: Servicios realizados por la PRODECON

Se puede observar que conforme fue pasando el tiempo los servicios brindados por este organismo fueron aumentando.

Cabe señalar que esto no necesariamente indica que la complejidad de la ley haya aumentado ya que al ser un nuevo organismo público no resulta extraño que durante los primeros años de operación sea poco conocido y por ello no hayan acudido a él las personas. La tendencia monótona creciente que se observa en los asuntos atendidos también se observa en la razón entre los asuntos atendidos y el universo de contribuyentes:

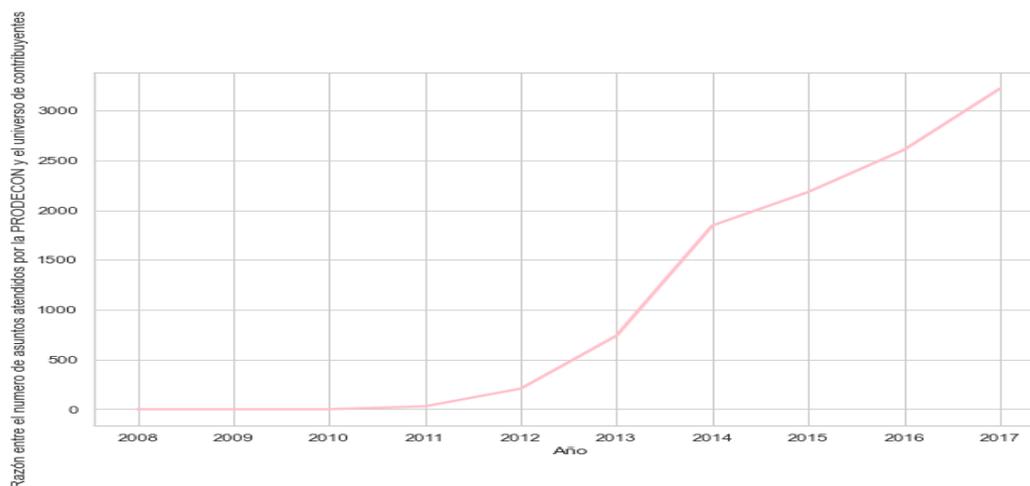


Figura 3.111: Razón entre los asuntos atendidos por la PRODECON y el universo de contribuyentes

Por otro lado, la correlación de Spearman entre la puntuación de complejidad legal propuesta y el total de servicios prestados por la PRODECON es no significativa mientras que la correlación de Pearson esta muy cerca de no serlo. Esto se puede validar con los datos de la siguiente tabla:

	Correlación de Pearson	P-Valor de Correlación de Pearson	Correlación de Spearman	P-Valor de Correlación de Spearman
Puntuación de complejidad legal y servicios de la PRODECON	-0.765191	0.045010	-0.522544	0.228878

Figura 3.112: Correlación entre la puntuación de complejidad y los servicios prestados por la PRODECON

Capítulo 4

Conclusiones

Se lograron cuantificar características que comparten las leyes en México las cuales permitieron describir la ley analizada y con ello se pudo comparar múltiples versiones de esta ley. Estas comparaciones pueden realizarse analizando algunas características en particular o algunas de las puntuaciones de complejidad propuestas. Dichas puntuaciones lograron detectar eventos que influyeron en la ley analizada, como por ejemplo la derogación del IETU y el IDE.

Se observó que, con base en las puntuaciones de complejidad propuestas, las características propias de la estructura de la LISR son las características preponderantes en la puntuación de complejidad legal durante los primeros años del periodo analizado. No obstante, esto cambió durante los últimos años dado que el lenguaje y las interdependencias pasaron a ser las características preponderantes en la puntuación de complejidad legal.

Por otro lado, el número de nodos es la característica que más tiende a aumentar la puntuación de complejidad de estructura y con ello la complejidad en general por lo que es la característica que más dificulta la interpretación de la ley con la puntuación propuesta. Esto se pudo observar en el comportamiento de las puntuaciones de complejidad obtenidas y se validó con ayuda de las correlaciones obtenidas.

Respecto a la relación entre la puntuación de complejidad legal propuesta con la recaudación obtenida con base en el ISR, se observó que por si sola no tiene una relación significativa, ni siquiera con la razón entre el universo de contribuyentes y el monto recaudado. De igual manera sucedió con los servicios presentados por la PRODECON, no hubo una relación significativa con la puntuación de complejidad legal propuesta. Un par de características que compartieron el monto recaudado con base en el ISR y los servicios prestados por la PRODECON fue una tendencia monótona creciente y que no se vieron cambios significativos en estas cuando ocurrieron los eventos que si impactaron a la puntuación de complejidad legal, como la derogación del IETU. Lo anterior generó la impresión de que la puntuación de complejidad legal cumplió con su objetivo de detectar fenómenos que afectan directamente al documento legal en si, la ley, que aquellos que dependen de la ley tal como es la recaudación.

Al contemplarse la información generada por todas las características analizadas se concluyó que las reformas aplicadas a la LISR durante el periodo analizado tienden a hacerla más compleja en su estructura, debido a la acumulación de excepciones

en la aplicación de la ley. Esta complejidad en particular se vio afectada principalmente por la adición de nuevos nodos, los cuales tienden a ser nodos que contienen información cada vez más especializada de un tema en particular; su función es complementar aspectos muy específicos de temas que con el pasar de los años se volvieron más complejos o que en un inicio fueron ambiguos en la aplicación de la ley. Este comportamiento hará que esta ley tenga cada vez más casos particulares de los temas que la integran y con ello quienes apliquen esta ley tendrán que estar conscientes de información cada vez más detallada al aplicarla.

En contraste con lo anterior, son pocos los nodos nuevos cuyo contenido sea un tema que no se haya tratado anteriormente en la LISR. Estos temas nuevos serán proporcionados por cambios en los modelos de negocio, los cuales irán evolucionando conforme la tecnología y la sociedad lo hagan. Un ejemplo es la reforma a la LISR publicada en diciembre de 2019¹. Esta reforma adiciona en el título IV, capítulo II, la sección III denominada “De los ingresos por la enajenación de bienes o la prestación de servicios a través de Internet, mediante plataformas tecnológicas, aplicaciones informáticas y similares”. Esta sección contiene información de un cambio de negocio que se ha presentado durante los últimos años, los servicios que se ofrecen mediante plataformas digitales.

Se concluye que el enfoque empleado para desarrollar la puntuación de complejidad legal en el presente trabajo puede enriquecerse de ramas de las matemáticas y de las ciencias de la computación que permitan abstraer información útil para la generación de una puntuación de complejidad legal la cual permita analizar la evolución de las leyes, prevenir posibles ajustes dados los cambios que se van dando en los modelos de negocio y con ello se pueda modificar tales leyes o crear mecanismos que faciliten la comprensión de estas.

4.1. Siguiendo pasos

Algunas áreas de oportunidad, de las que se tiene consciencia, que mejorarían el presente trabajo son las siguientes:

- Estructura: se podrían aplicar otras propiedades de teoría de gráficas y de estructuras de datos que ayuden a contemplar de una mejor manera la relación jerárquica que existe entre los distintos tipos de nodo contemplados en el presente trabajo o trabajar la estructura de las leyes con alguna estructura de datos que sea más útil para la abstracción de métricas que ayuden a medir la complejidad de las leyes.
- Lenguaje: se puede hacer un uso adecuado del procesamiento natural del lenguaje, también conocido como NLP por sus siglas en inglés, para mejorar la detección de referencias tanto internas como externas. Además, también se podría

¹Esta reforma publicada en el diario oficial de la federación se puede consultar en la siguiente página de internet: https://www.diputados.gob.mx/LeyesBiblio/ref/lisr/LISR_ref03_09dic19.pdf

emplear para la aplicación de pruebas similares como la prueba de legibilidad de Flesch pero enfocadas en las leyes. Así como, en la aplicación de funciones que complementen a la entropía en el análisis del texto que integra a la LISR.

- Interdependencia: se encontraron diversos algoritmos cuyo objetivo es medir las relaciones de dependencia que existen entre varios elementos que se citan. Podría existir un algoritmo que pondere de mejor manera las citas que existen en las leyes ya que en el presente trabajo sólo se contempló la existencia de dichas citas más no el contenido de los objetos citados y las características de estos objetos. De analizarse lo anterior se podría abstraer información de estas citas que permita cuantificar con mayor detalle la contribución que éstas tienen a la complejidad legal de la ley que las contiene o del marco jurídico mexicano en si. También puede ser que existe un enfoque distinto al que se contempló en el presente trabajo para abordar de una mejor manera dichas citas.

El presente trabajo se hizo desde una perspectiva matemática que buscaba abstraer características de las leyes que contribuyan a cuantificar la complejidad de estas. No obstante, se considera que este tema, la complejidad legal, al igual que otros temas creados por elementos meramente humanos, como es el caso del marco jurídico mexicano, se podrá analizar desde más de una perspectiva y si dichas perspectivas trabajan en conjunto se podrá genera un trabajo más completo.

Anexos A

Definiciones

A.1. Gráfica

En teoría de gráficas, una gráfica G se define como:

$$G = (V, E)$$

Donde $E \subseteq [V]^2$. De este modo, los elementos de E son subconjuntos de dos elementos de V .

A.2. Nodos en una gráfica

Dada una gráfica G , los elementos del conjunto V se denominan nodos, vértices o puntos. Usualmente, los nodos de una gráfica se visualizan como puntos.

A.3. Aristas en una gráfica

Dada una gráfica G , los elementos del conjunto E se denominan aristas o líneas. Usualmente, las aristas de una gráfica se visualizan como líneas que unen dos puntos.

A.4. Media geométrica

La media geométrica G de n números reales positivos x_1, x_2, \dots, x_n está definida por la fórmula:

$$G = (x_1 x_2 \dots x_n)^{\frac{1}{n}}$$

A.5. Media ponderada

Si w_1, w_2, \dots, w_n son n números no negativos, llamados pesos, no todos ceros, la media aritmética ponderada \bar{a} de a_1, a_2, \dots, a_n se define mediante la fórmula:

$$\bar{a} = \frac{\sum_{k=1}^n w_k a_k}{\sum_{k=1}^n w_k}$$

Cuando los pesos son todos iguales, este valor coincide con la media aritmética ordinaria.

A.6. Correlación de Pearson

El coeficiente de correlación muestral de Pearson se define como:

$$r_{xy} = \frac{\sum (x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum (x_t - \bar{x})^2 (y_t - \bar{y})^2}}$$

Donde \bar{x} y \bar{y} son las medias muestrales de x y y , respectivamente. Mide la relación lineal entre x y y y satisface que $-1 \leq r_{xy} \leq 1$, siendo -1 y 1 los valores que indican una relación lineal perfecta entre x y y . Por otro lado, si el valor es igual a 0 entonces la relación entre x y y no es lineal.

A.7. Correlación de Spearman

Sea $r(X_t)$ que denota el rango de X_t y $r(Y_t)$ que denota el rango de Y_t , con X_t y Y_t con el mismo número de observaciones. Si se usan pares de dichos rangos, la correlación muestral ordinal se calcula como:

$$S = \frac{\sum [r(X_t) - \bar{r}(X)][r(Y_t) - \bar{r}(Y)]}{\sqrt{\sum [r(X_t) - \bar{r}(X)]^2 [r(Y_t) - \bar{r}(Y)]^2}}$$

Donde:

- $\bar{r}(X) = \sum r(X_t)/n$
- $\bar{r}(Y) = \sum r(Y_t)/n$

Apéndice A

Proceso de ingesta de datos

Las versiones de la LISR se extrajeron de la página de internet¹ de normatividad del SAT el día 5 de febrero de 2021. Se descargó el PDF que contiene la LISR, se extrajo la información para vaciarla en un archivo Excel. Una vez que se creó dicho archivo con la información de todo el periodo analizado se importó su información a la herramienta SAS Studio utilizando el encoding UTF-8. Se validó que con dicho encoding se conservaran todos los caracteres, incluyendo palabras con acentos y aquellas que contienen la letra “ ñ ”, así como los signos de puntuación.

Por otro lado, la información de la recaudación obtenida mediante el ISR se obtuvo de la página de internet² de datos abiertos del SAT el día 5 de febrero de 2021. Se extrajeron los datos de recaudación de los reportes anuales del periodo analizado que se encuentran en dicha página de internet. Dichos datos se depositaron en un archivo Excel para importarlo en la herramienta SAS Studio.

Finalmente, las cifras de los servicios realizados por la PRODECON se obtuvieron de la página de internet³ de estadísticas de este organismo público. Se depositaron dichas cifras en un archivo Excel para importarlo en la herramienta SAS Studio.

¹<https://www.sat.gob.mx/personas/normatividad>

²http://omawww.sat.gob.mx/cifras_sat/Paginas/inicio.html

³<https://prodecon.gob.mx/gobmx/estadisticas/estadisticas.html>

Apéndice B

Procesamiento de datos

El software utilizado en el presente trabajo para realizar el procesamiento del texto contenido en la LISR, el proceso de parseo de dicho texto y el análisis de este fue SAS Studio. Dado que esta herramienta no realiza el proceso de parseo de algunas palabras en español de manera correcta, se revisaron las listas de palabras generadas por dicho proceso y se corrigieron los errores detectados en el parseo mediante el uso de diccionarios de sinónimos y de multi-términos.

Por otro lado, la detección de las referencias internas y referencias externas también se realizó en esta herramienta de SAS mediante la implementación de funciones con expresiones regulares.

Por su parte, el software empleado para el cálculo de las puntuaciones Authority y Hub fue la librería networkx del lenguaje de programación Python, el cual se trabajó desde la herramienta Jupyter Notebook.

Por último, dada la variedad de gráficos que se pueden crear con distintas librerías de Python, también se optó por la creación de todo el material gráfico del presente trabajo en dicho lenguaje. Las librerías que se emplearon para esto fueron principalmente seaborn, matplotlib y altair.

Apéndice C

Puntuaciones de complejidad con base al año y el título

Con excepción del año 2011, número de títulos no cambió durante el periodo de tiempo contemplado por lo que se pudo analizar la evolución de estos. Los títulos y el tema en el que se enfocan son los siguientes:

- Título I. Disposiciones generales.
- Título II. De las personas morales.
- Título III. Del régimen de las personas morales con fines no lucrativos.
- Título IV. De las personas físicas.
- Título V. De los residentes en el extranjero con ingresos. provenientes de fuente de riqueza ubicada en territorio nacional.
- Título VI. De los regímenes fiscales preferentes y de las empresas multinacionales.
- Título VII. De los estímulos fiscales.

El título extra que tiene la LISR del año 2011 es el título VIII, “Del fomento al primer empleo”, el cual no se consideró en el siguiente análisis dado que sólo se presenta en un año.

C.1. Puntuación de complejidad de la estructura de la LISR

En la siguiente gráfica se muestran las puntuaciones de complejidad de la estructura calculadas para cada título:

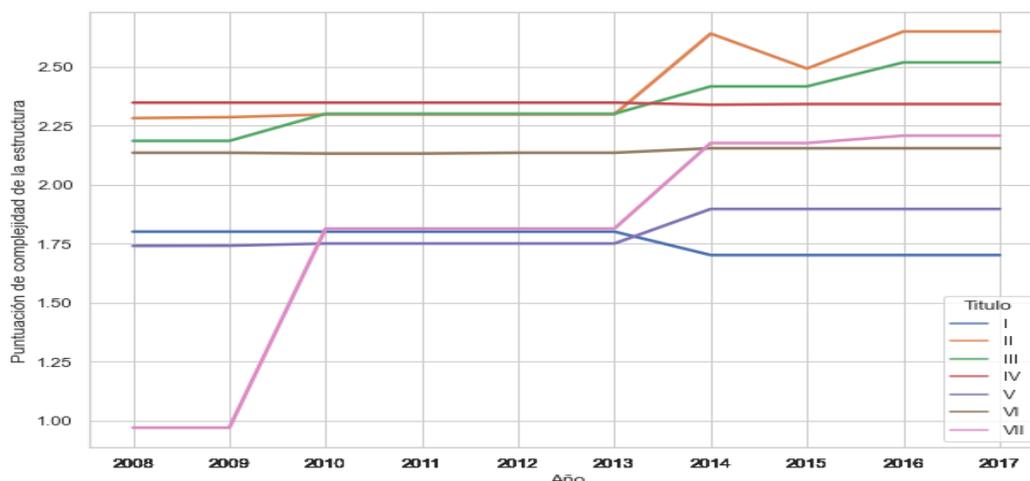


Figura C.1: Puntuaciones de complejidad de la estructura de los títulos de la LISR por año

Se pudo observar que la mayoría de los títulos no tuvieron cambios notorios en su puntuación con excepción de los títulos II y VII. Contrario a estos, los títulos que menos cambiaron fueron los títulos IV y VI. Cabe señalar que los títulos con la estructura más compleja durante todo el periodo de tiempo son los títulos II, III y IV. En este nivel de análisis se pudo ver el impacto que tuvieron los cambios del año 2014 ya que a todos los títulos les afectó, a algunos el cambio fue mayor que a otros, y en la mayoría fue un cambio que aumentó la puntuación de complejidad de estructura pese a que el número de nodos disminuyó de 2013 a 2014.

C.2. Puntuación de complejidad del lenguaje de la LISR

En esta puntuación sobresale la complejidad del título I dado que en la puntuación anterior es de los títulos con menor puntuación de complejidad de estructura pero es de los títulos con mayor puntuación de complejidad del lenguaje. Lo anterior ocurre de forma inversa con el título II, de ser uno de los títulos con mayor complejidad de estructura es también uno de los títulos con menor complejidad de lenguaje. Esto se pudo observar en la siguiente gráfica:

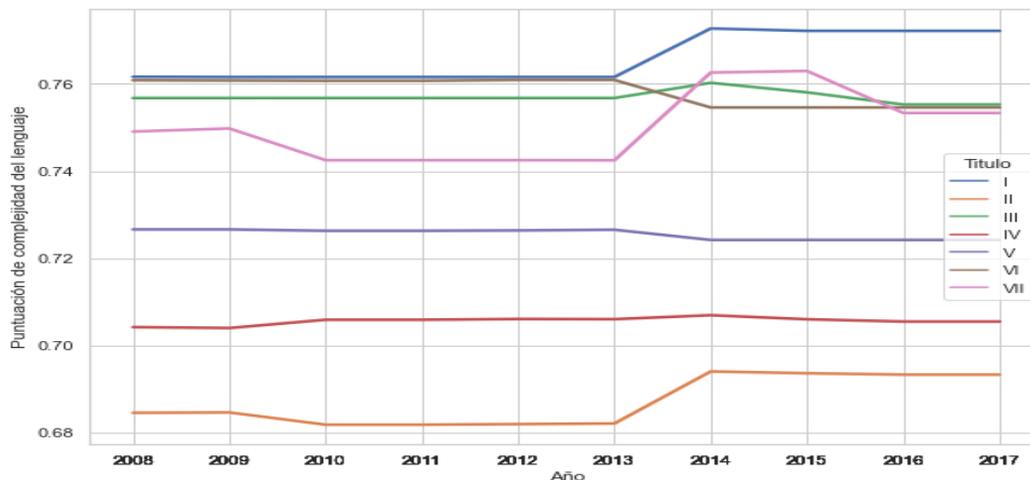


Figura C.2: Puntuaciones de complejidad del lenguaje en los títulos de la LISR por año

El efecto del cambio en el 2014 también se pudo observar en esta puntuación dado que es único año donde todos los títulos tuvieron un cambio que sobresale respecto a los demás años. Un título que sobresale por mantenerse como uno de los más complejos tanto en la puntuación de complejidad estructura como en la puntuación de complejidad del lenguaje es el título III. Lo mencionado anteriormente indica que existen títulos cuya estructura puede ser compleja, ya sea por la profundidad de sus nodos o el grado de salida de estos, pero su contenido no lo es tanto comparado contra otros cuya estructura es menos compleja dado que no varían tanto las palabras que contienen.

C.3. Puntuación de complejidad de las interdependencias en la LISR

Al igual que en la puntuación anterior, sobresale el título I dado que es el título con la mayor puntuación de complejidad de interdependencias mientras que el título III aún figura como uno de los títulos más complejos. Lo anterior se pudo notar en la siguiente gráfica:

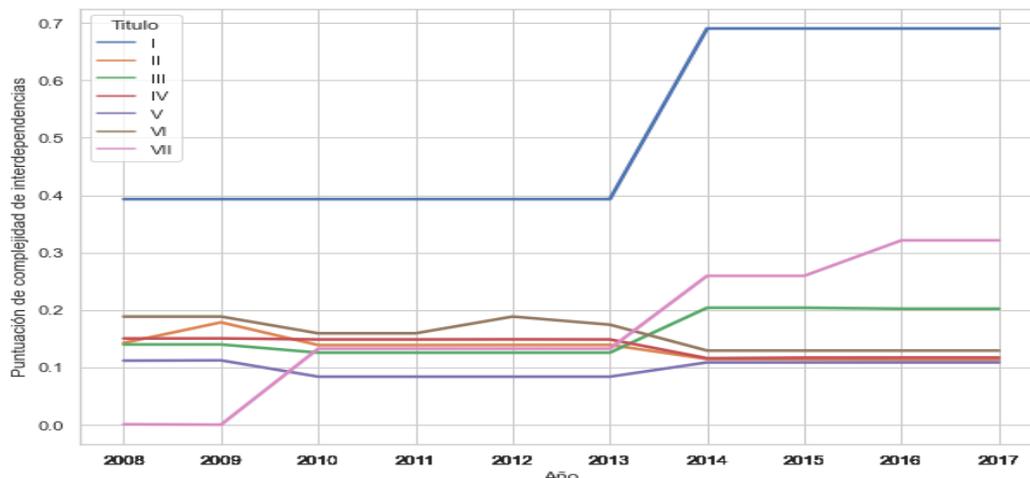


Figura C.3: Puntuaciones de complejidad de las interdependencias en los títulos de la LISR por año

Si bien se vieron cambios en las puntuaciones del 2014, estos no son tan sobresalientes como en las puntuaciones de complejidad anteriores, con excepción del título I. También sobresale el título VII ya que, al igual que en la puntuación de complejidad de estructura, pasó de ser de uno de los títulos con menor puntuación de complejidad a uno de los que tiene mayor puntuación de complejidad en los años más recientes. Dicha característica la comparte con el título III.

C.4. Puntuación de complejidad legal de la LISR

Finalmente, las puntuaciones de complejidad legal calculadas fueron las siguientes:

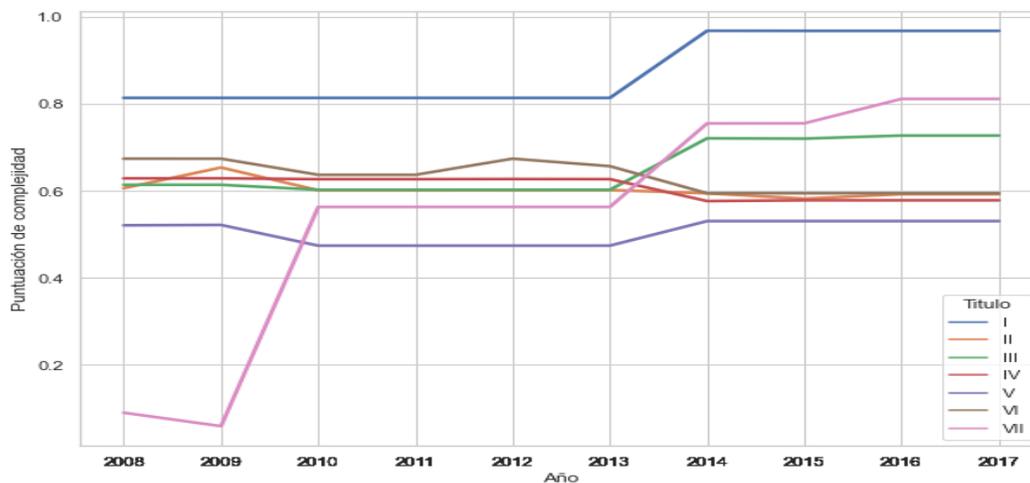


Figura C.4: Puntuaciones de complejidad legal de los títulos de la LISR por año

No es de extrañar que el título I sea el título con mayor puntuación de complejidad legal dado que lo fue en dos de 3 puntuaciones. Así mismo, el título VII mantiene su comportamiento de ser uno de los títulos con mayor puntuación de complejidad legal en los últimos años, característica que comparte junto con el título III. Resultó interesante observar que los demás títulos tienen puntuaciones de complejidad legal muy similares después del 2014.

Dado lo anterior, los títulos que tendieron a ser más complejos durante el periodo contemplado fueron los siguientes:

- Título I. Disposiciones generales.
- Título III. Del régimen de las personas morales con fines no lucrativos.
- Título VII. De los estímulos fiscales.

Bibliografía

- [1] Michael J. Bommarito y Daniel M. Katz. «A mathematical approach to the study of the United States Code». En: *Physica A: Statistical Mechanics and its Applications* 389.19 (oct. de 2010), págs. 4195-4200. ISSN: 03784371. DOI: 10.1016/j.physa.2010.05.057. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0378437110004875>.
- [2] Danièle Bourcier y Pierre Mazzega. «Toward measures of complexity in legal systems». En: *Proceedings of the 11th international conference on Artificial intelligence and law - ICAIL '07*. New York, New York, USA: ACM Press, 2007, pág. 211. ISBN: 9781595936806. DOI: 10.1145/1276318.1276359. URL: <http://portal.acm.org/citation.cfm?doid=1276318.1276359>.
- [3] Gerardo Febres, Klaus Jaffé y Carlos Gershenson. «Complexity measurement of natural and artificial languages». En: *Complexity* 20.6 (jul. de 2015), págs. 25-48. ISSN: 10762787. DOI: 10.1002/cplx.21529. URL: <https://onlinelibrary.wiley.com/doi/10.1002/cplx.21529>.
- [4] R FLESCHE. «A new readability yardstick.» En: *The Journal of applied psychology* 32.3 (jun. de 1948), págs. 221-33. ISSN: 0021-9010. DOI: 10.1037/h0057532. URL: <http://www.ncbi.nlm.nih.gov/pubmed/18867058>.
- [5] Eric Kades. *The Laws of Complexity & the Complexity of Laws: The Implications of Computational Complexity Theory for the Law THE LAWS OF COMPLEXITY AND THE COMPLEXITY OF LAWS: THE IMPLICATIONS OF COMPUTATIONAL COMPLEXITY THEORY FOR THE LAW*. Inf. téc. 1997. URL: <https://scholarship.law.wm.edu/facpubs/646>.
- [6] Daniel Martin Katz y M. J. Bommarito. «Measuring the complexity of the law: the United States Code». En: *Artificial Intelligence and Law* 22.4 (dic. de 2014), págs. 337-374. ISSN: 0924-8463. DOI: 10.1007/s10506-014-9160-8. URL: <http://link.springer.com/10.1007/s10506-014-9160-8>.
- [7] Rick Krever y Peter Mellor. «Where's the Complexity in Tax Law ?» En: (2015).
- [8] Eliseo Muro Ruíz. *Algunos elementos de técnica legislativa*. Universidad Nacional Autónoma de México, 2006. ISBN: 970-32-3453-4. URL: <http://ru.juridicas.unam.mx:80/xmlui/handle/123456789/11063>.

- [9] Peter H. Schuck. «Legal Complexity: Some Causes, Consequences, and Cures». En: *Duke Law Journal* 42.1 (oct. de 1992), pág. 1. ISSN: 00127086. DOI: 10.2307/1372753. URL: <https://www.jstor.org/stable/1372753?origin=crossref>.
- [10] C. E. Shannon. «A Mathematical Theory of Communication». En: *Bell System Technical Journal* 27.3 (jul. de 1948), págs. 379-423. ISSN: 00058580. DOI: 10.1002/j.1538-7305.1948.tb01338.x. URL: <https://ieeexplore.ieee.org/document/6773024>.
- [11] Bernhard Wautl y Florian Matthes. «Towards Measures of Complexity: Applying Structural and Linguistic Metrics to German Laws». En: *Frontiers in Artificial Intelligence and Applications* 271 (2014), págs. 153-162. ISSN: 09226389. DOI: 10.3233/978-1-61499-468-8-153.
- [12] Ernesto Zavala. «Trayectoria del Impuesto Sobre la Renta en México». En: *Revista de la Facultad de Derecho de México* 1.99-100 (1975), págs. 627-662. URL: <https://revistas-colaboracion.juridicas.unam.mx/index.php/rev-facultad-derecho-mx/article/view/26984/24334>.