

Modelo de contaminación de riesgo de incumplimiento tributario

El modelo de contaminación de riesgo de incumplimiento tributario, es un modelo de difusión, donde el riesgo de incumplimiento tributario de un contribuyente es propagado hacia todos aquellos contribuyentes relacionados. Esto se basa en el hecho de que los contribuyentes, al igual que las personas en una sociedad no son entes aislados, sino que su comportamiento puede estar muchas veces inducido y/o afectado por su entorno, por ejemplo, es natural pensar que el comportamiento tributario de una empresa dependerá (en algún factor) del comportamiento tributario que tiene su dueño o sus socios, o en el caso de un contribuyente persona natural podemos pensar que su comportamiento o nivel de riesgo tributario también estará relacionado en alguna medida por el de su familia más cercana.

Similarmente a lo que sucede con el modelo general, el cual encapsula en un mismo concepto de riesgo efectos tanto deterministas como probabilísticos, al momento de propagar el riesgo, consideraremos que éste se puede propagar debido a una influencia directa de un contribuyente sobre otro, como por ejemplo, al considerar una relación patrimonial o familiar, o debido a un impacto indirecto, como al considerar una relación comercial donde por ejemplo si una de las partes en la relación no cumple de forma correcta todas sus obligaciones tributarias podría afectar a que la otra parte no cumpla o cumpla parcialmente con las suyas, luego en este caso no hay una influencia directa en el comportamiento del contribuyente contaminado, sino que la contaminación se da de forma indirecta. Así, las relaciones que se definirán para la contaminación podrán identificarse bajo uno de estos dos subtipos.

Bajo este contexto en lo que sigue se define un marco metodológico general para afrontar la propagación o difusión del riesgo de incumplimiento tributario, el cual puede ser extensible a nuevas fuentes de contaminación en el futuro, todo, sobre la base del score de riesgo general, ya explicado anteriormente.

Relaciones relevantes para el análisis de contaminación en el sistema tributario

Como se mencionó, este modelo es extensible a todas aquellas relaciones relevantes para el sistema tributario, sin embargo, para efectos de la Clasificación de Riesgo Global del año 2020 se consideran los siguientes tipos de relaciones relevantes para la propagación de riesgo:

- Relaciones Societarias
- Relaciones Comerciales (según emisión/recepción de facturas)
- Relaciones Familiares

Todas estas relaciones se trabajarán solo a un arco, sin embargo, en futuras actualizaciones del modelo se pretende abarcar más arcos incorporando factores de atenuación según el largo de las cadenas y/o sobre las iteraciones del modelo.

Marco general del modelo de contaminación

El modelo propuesto está basado en propagación de influencias en una red o grafo, y para ello definiremos una serie de conceptos relevantes hasta llegar al cálculo del score de riesgo contaminado.

Fuerza de la relación o influencia

Para ejemplificar este concepto consideremos 2 contribuyentes relacionados, de forma tal que el contribuyente A posee una influencia o podría ser capaz de generar un efecto directo o indirecto sobre el riesgo tributario del contribuyente B, como se muestra en la siguiente figura:

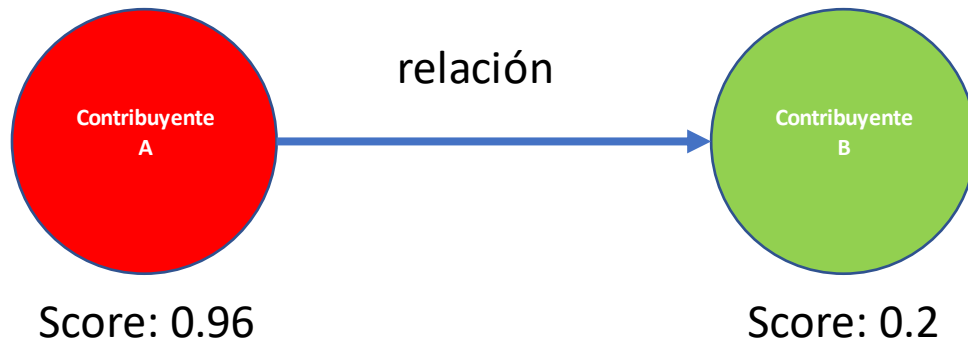


Ilustración 1: Relacion entre dos contribuyentes

Luego, nos interesa ver como el riesgo tributario del contribuyente A, el cual tiene un score inicial de 0.96, se propaga hacia el contribuyente B, el cual tiene un score inicial de 0.2. Para esto definiremos una fuerza o impacto de la relación entre A y B, por ejemplo, si consideráramos una relación patrimonial, el riesgo no debiera propagarse de la misma manera en un escenario donde el contribuyente A posee el 1% de la participación en el contribuyente B, que si tuviera el 100% de participación, sino que aquí uno esperaría que la fuerza de la relación cuando se tiene un 100% de participación fuera mucho mayor que cuando se tiene solo un 1%. Por tanto, toda relación debiera tener una magnitud específica, dependiente de las características propias de ésta y de los contribuyentes relacionados, a este valor denominaremos **fuerza de la influencia**. Para efectos del modelamiento, esta fuerza tomará valores entre 0 y 1, de forma tal que

$$F_{A \rightarrow B} = \text{Fuerza de la influencia de A sobre B} = F_{A \rightarrow B}(\vec{x}) \in [0, 1]$$

Donde \vec{x} es un vector de las características de la relación entre A y B, y donde $F(\vec{x}) \in [0, 1]$ denotará la intensidad o fuerza de esta relación, siendo un valor 0 indicador que la relación es nula o no significativa, y donde un valor 1 indicaría una influencia total. Así, en el ejemplo anterior, si la única

característica relevante de la relación patrimonial fuera el porcentaje de participación, entonces en el primer escenario donde el contribuyente A tiene un 1% de participación sobre el contribuyente B se esperaría una fuerza de la influencia cercana a 0 y en el caso que la participación fuera del 100% se esperaría una fuerza de la influencia cercana a 1.

Así uno de los mayores desafíos del modelamiento de la propagación del riesgo, es identificar aquellas características relevantes en una relación y modelar esta fuerza de la influencia entre dos contribuyentes.

Por ejemplo, para el caso de relaciones patrimoniales, se ha definido la fuerza de la influencia de un socio sobre otro contribuyente persona jurídica de la siguiente manera:

$$F_{A \rightarrow B}(\vec{x}) = \beta * (PctParticipacion^{\alpha_1} * RatioVCSocioSociedad^{\alpha_2} * RatioVCSocios^{\alpha_3} * InfluenciaSocioSociedad)^{\frac{1}{\alpha_1 + \alpha_2 + \alpha_3}}$$

Donde:

- ***PctParticipacion***: Porcentaje de participación del contribuyente A en el contribuyente B.
- ***RatioVCSocioSociedad***: Indicador de la relación entre el Valor del Contribuyente del contribuyente A y el Valor del Contribuyente del contribuyente B.
- ***RatioVCSocios***: Indicador del Valor del Contribuyente del contribuyente A relativo al Valor del Contribuyente de todos los socios del Contribuyente B.
- ***InfluenciaSocioSociedad***: Indicatriz que toma el valor 1 si $Score_A > Score_B$ y 0 si no.

Todos estos indicadores tienen valores en el rango $[0, 1]$, siendo un mayor valor un indicador de una mayor influencia en esa característica. Por su parte β , α_1 , α_2 y α_3 son parámetros de ajuste del modelo que dan cuenta del peso de cada factor para determinar la fuerza de la influencia total entre A y B.

La indicatriz *InfluenciaSocioSociedad* tiene el efecto de solo considerar la “contaminación negativa”, esto es, cuando los relacionados tiene un mayor riesgo que el contribuyente contaminado, de esta manera este modelo solo propaga riesgo, y no permite disminuir el riesgo base de un contribuyente.

Notar que la fórmula para $F_{A \rightarrow B}(\vec{x})$ es una media geométrica ponderada, y se utiliza esto, ya que lo que se pretende es comparar los distintos aspectos de una relación, los cuales son medidos con diferentes indicadores, además, la media geométrica tiene dos propiedades deseables para el análisis de la contaminación:

- La media geométrica Vale 0, si alguno de sus componentes vale 0, lo cual en nuestro análisis es algo deseable, ya que si por ejemplo, en nuestro caso anterior, el *PctParticipacion* = 0, significa que el contribuyente A no es socio de la empresa B, y por tanto es correcto que la contaminación societaria de A sobre B sea 0 (porque no existe tal relación), esto no se obtendría si en vez de utilizar una media geométrica se usara una media aritmética.

- Y lo segundo es que la media aritmética va a valer 1, si solo si todas sus componentes valen 1, es decir, la fuerza de la influencia alcanza su máximo solo cuando lo alcanza en todas las componentes en que se mida.

Con la fuerza de la influencia ya calculada, el siguiente paso es combinarlas cuando el contribuyente B posee más de un relacionado, es decir, la fuerza de la influencia solo considera el efecto entre dos individuos, A y B, pero no toma en consideración todos los efectos a los que está expuesto el contribuyente B, luego, el siguiente concepto se encarga de ajustar estos valores de fuerza de influencia en un escenario con múltiples agentes contaminadores.

Factor de influencia

Consideremos el siguiente escenario, donde el contribuyente B se ve influenciado, o está relacionado a los contribuyentes A_1, A_2, \dots, A_n , con fuerzas de influencia F_1, F_2, \dots, F_n .

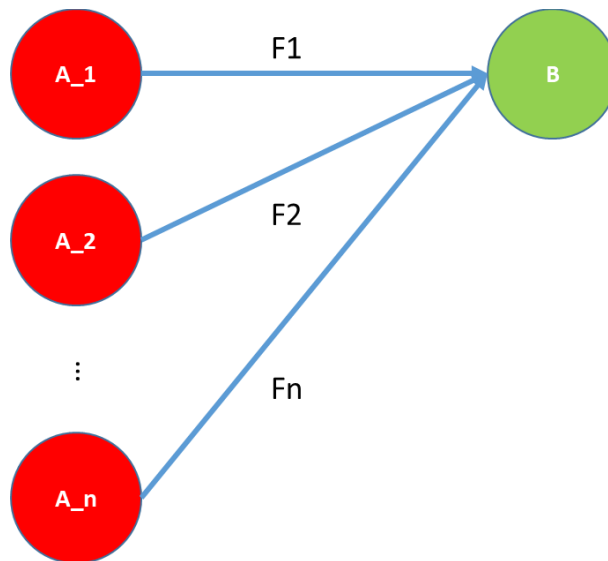


Ilustración 2: Diagrama con múltiples relaciones

Luego, el factor de influencia asociado a cada relación, es una normalización de las Fuerzas de las influencias ejercidas sobre el contribuyente B, de manera que se satisfagan algunos criterios deseables del modelo.

Sea f_i el factor de influencia asociado a F_i , definiremos las siguientes propiedades sobre los f_i :

1. Si $n = 1$, es decir solo existe una relación afectando a B, entonces $f_1 = F_1$.
2. Si $F_i = 0 \Rightarrow f_i = 0$
3. $\sum_{i=1}^n f_i \leq 1$
4. Sea $f_0 = 1 - \sum_{i=1}^n f_i$, luego $\sum_{i=0}^n f_i = 1$, y $f_0 = 0 \Leftrightarrow F_i = 1$ para algun $i \in \{1, \dots, n\}$
5. Si $F_i = F_j$ entonces $f_i = f_j$ (simetría)

De esta manera, en base al punto 4), el factor de influencia se puede interpretar como la probabilidad de que el contribuyente B sea afectado o influenciado por el contribuyente A_i , y donde f_0 lo asociamos a la probabilidad de que el contribuyente B no sea afectado o influenciado por ninguno de sus relacionados.

Luego, si los f_i fueran probabilidades, y los f_i $i \geq 1$ fueran todos independientes entonces podríamos calcular f_0 de la siguiente manera

$$\begin{aligned} f_0 &= \mathbb{P}(\text{No es influenciado por ninguno de los } A_i) \\ f_0 &= \mathbb{P}\left(\bigcap (\text{No es influenciado por } A_i)\right) \\ f_0 &= \prod \mathbb{P}(\text{No es influenciado por } A_i) \\ f_0 &= \prod (1 - \mathbb{P}(\text{Es influenciado por } A_i)) \\ f_0 &= \prod (1 - f_i) \end{aligned}$$

Ahora, si bien los f_i no son exactamente probabilidades, ni son eventos independientes, de igual manera podemos para nuestro propósito, definir f_0 inspirándonos en el resultado anterior, ya que representaría en cierta medida el hecho de “no ser influenciado”. Así utilizando la ecuación anterior y ajustándola a los valores conocidos de F_i definiremos f_0 como:

$$f_0 = \prod_{i=1}^n (1 - F_i)$$

Notemos que f_0 definido de esta manera no entra en conflicto con las propiedades 1 a 5, y por lo tanto podemos establecer la siguiente fórmula para el resto de los f_i de manera que se siguen cumpliendo todas las propiedades antes declaradas

$$f_i = (1 - f_0) * \left(\frac{F_i}{\sum_j F_j} \right)$$

Finalmente, con los f_i definidos, ya podemos combinar los scores de lo relacionados, en un score contaminado.

Score de Riesgo Contaminado

Consideremos la Ilustración 2, luego el score de riesgo contaminado lo calcularemos como una suma ponderada del score de riesgo de los contribuyentes relacionados, utilizando como ponderador los valores de los factores de influencia (de ahí la importancia de la propiedad 4). Esto es

$$SC_B = \sum_{\{i \geq 0\}} f_i * S_i$$

Donde S_i es el Score de riesgo general del Contribuyente A_i para $i = 1, \dots, n$. Y para $i = 0$, S_0 es el score de riesgo general del contribuyente B.

Ahora a modo de ejemplo, consideremos el siguiente escenario, y calculemos el Score contaminado con los siguientes Scores y Fuerzas de influencia.

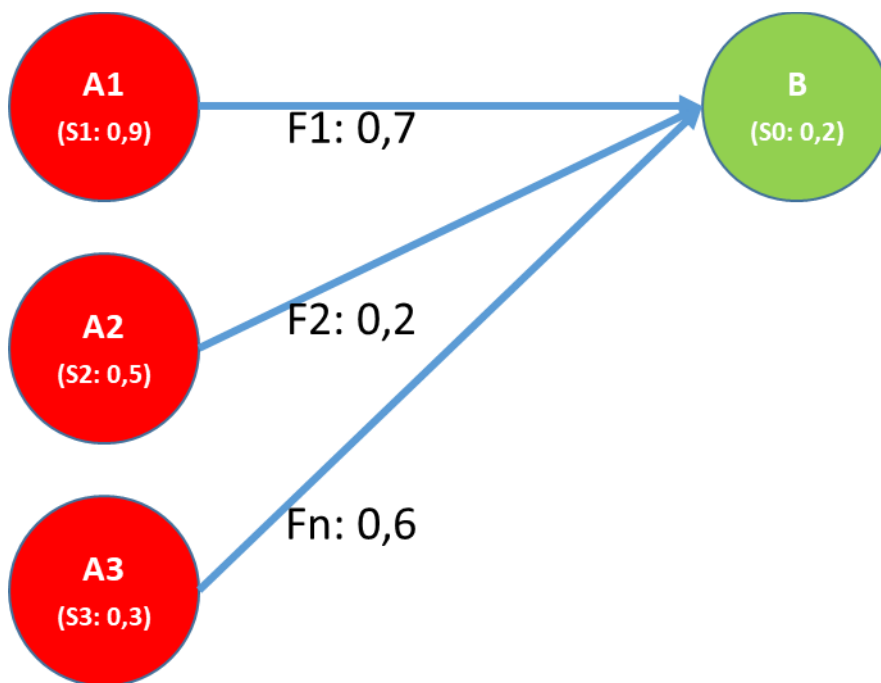


Ilustración 3: Escenario con 3 relaciones

Luego, utilizando las formulas antes descritas tendremos que:

$$f_0 = (1 - 0,7) * (1 - 0,2) * (1 - 0,6) \\ = 0,096$$

$$f_1 = (1 - 0,096) * \left(\frac{0,7}{1,5}\right) \\ = 0,422$$

$$f_2 = (1 - 0,096) * \left(\frac{0,2}{1,5}\right) \\ = 0,121$$

$$f_3 = (1 - 0,096) * \left(\frac{0,6}{1,5}\right) \\ = 0,362$$

Con

lo

que

$$SC_B = 0,096 * 0,2 + 0,422 * 0,9 + 0,121 * 0,5 + 0,362 * 0,3 \\ = 0,568$$

De esta manera, con la metodología propuesta hemos determinado un score de contaminación de 0,568.

Esta forma de calcular el score de contaminación se puede interpretar como una suma ponderada entre el riesgo de los agentes contaminadores ($A_i, i = 1, \dots, n$) y el impacto o influencia que tienen cada uno de ellos en el contribuyente B.

Notemos por último algunas de las propiedades de esta modelo:

- Como todos los $f_i \in [0, 1]$ y los $S_i \in [0, 1]$ la suma ponderada siempre estará bien definida en el rango $[0, 1]$.
- El modelo es creciente en los S_i es decir, si alguno de los scores de riesgo originales aumenta, también lo hace el Score contaminado.
- El modelo es simétrico, es decir no depende del orden que se les dé a los A_i ($SC_{\{B|A_1, A_2\}} = SC_{\{B|A_2, A_1\}}$)
- Como el modelo se definió sólo de contaminación negativa, el valor del Score Contaminado estará entre $[S_0, 1]$ siendo S_0 el score de riesgo general del contribuyente B, es decir, el Score Contaminado va a ser siempre mayor o igual que el Score inicial.
- El Score contaminado es cercano a 1, solo si existen contribuyentes con riesgo cercano a 1 cuya Fuerza de la influencia también sea cercana a 1, es decir, se conjuguen al mismo tiempo un riesgo alto, con una relación fuerte.
- El Score contaminado es cercano a 0, solo si el mismo contribuyente tiene un score de riesgo cercano a 0 y todos los contribuyentes relacionados tienen un score de riesgo cercano a 0 o una Fuerza de la influencia cercana a 0.