

# Detección de Comunidades en el Mercado de Compras Estatales

Diego Mazzuco, Ana Sofia Samaniego  
*Maestría Ciencia de datos*  
*Facultad de Ingeniería, Universidad de la República*  
 Montevideo, Uruguay  
 diego.mazzuco.ortiz@gmail.com, anasofiasama@gmail.com

## Resumen

El objetivo principal del trabajo es caracterizar el Mercado de Compras y Contrataciones Estatales de Uruguay en el período entre los años 2015 y 2018, mediante la detección de Comunidades de Proveedores. Las Comunidades son conjuntos de proveedores que se relacionan a través del intercambio con una misma Unidad de Compra (UC). El vínculo prolongado entre un conjunto de proveedores reducido y una UC puede evidenciar un sesgo de parte del comprador y/o concentración de poder de mercado del proveedor. Como resultado de este trabajo se espera dimensionar la existencia de este tipo de comportamientos.

## I. INTRODUCCIÓN

El Mercado de Compras y Contrataciones Estatales refiere al ámbito en el cual se intercambia un bien o servicio entre un Proveedor y una Unidad de Compra (UC) que integra el Estado, a un precio determinado; transacción que se regula por un marco normativo específico. En Uruguay las Contrataciones Estatales se realizan mediante Licitación Pública u otro procedimiento competitivo, de acuerdo a lo previsto en la normativa vigente. Dependiendo del procedimiento utilizado, la publicación del llamado, así como su posterior adjudicación, se materializa en el “Sitio de Compras”, plataforma web gestionada por la Agencia Reguladora de Compras Estatales (ARCE) (<https://www.comprasestatales.gub.uy/>).

La Agencia de Compras y Contrataciones del Estado (actualmente ARCE, modificación realizada por la Ley N° 19.924 [1]) se crea en la Rendición de Cuentas del 2007 (Ley 18.362 [2]); sin embargo, no es hasta mediados del 2012 que comienza a ejercer funciones con la creación del Consejo Directivo Honorario. El objetivo de la Agencia es contribuir a mejorar las condiciones en que el Estado procesa sus compras, así como a desarrollar las herramientas que aseguren la mayor transparencia en la gestión de compras del Estado (entre ellas el “Sitio de Compras”), realizando para ello acciones de normalización, estandarización, planificación y seguimiento, así como la instrumentación de las herramientas tecnológicas de apoyo necesarias<sup>1</sup>.

La creación de la ACCE estuvo motivada, en parte, por la necesidad de contar con registros fiables e información accesible para toda la ciudadanía. La transparencia de los procedimientos de compra estatales es considerada un factor clave en la promoción y regulación de la libre competencia. La libre competencia es un estímulo a la eficiencia económica y la libertad e igualdad de condiciones de acceso de empresas y productos a los mercados [3]. En este sentido, la competencia incentiva a las empresas a mejorar la calidad de sus productos y servicios y a reducir sus precios [4]. Sin embargo, la concentración de las compras estatales en un grupo reducido de proveedores puede otorgar a los mismos un significativo poder de mercado, posibilitando una situación donde el proveedor pueda ejercer abuso de posición dominante, afectando a sus compradores y/o competidores.

Cabe contemplar otros dos grandes hitos orientados a la promoción de la transparencia y la libre competencia. Por un lado, la publicación de la Ley de Acceso a la Información Pública (Ley 18.381 [5]) en el año 2008, la cual en parte obliga a los Organismos Públicos a publicar en sus páginas web un link al Sitio de Compras de las convocatorias y adjudicaciones de las Licitaciones Públicas. Por otro lado, la creación de la Comisión de Promoción y Defensa de la Competencia (Ley N° 18159 [3]) como órgano desconcentrado en el ámbito del Ministerio de Economía y Finanzas, encargado de controlar y sancionar las prácticas anticompetitivas en el país (Ministerio de Economía y Finanzas).

En este trabajo se considera que el Mercado de Compras y Contrataciones Estatales (de ahora en más MCCE) oficia como una red donde el vínculo entre Proveedores y UC es la representación del intercambio de bienes y servicios. Este vínculo se puede ir transformando en el tiempo, no solamente por los involucrados en las transacciones, sino también por el monto y/u objeto de las adjudicaciones. Por consiguiente el MCCE es un sistema dinámico que varía con el paso del tiempo. El vínculo prolongado entre un conjunto de proveedores reducido y una UC puede evidenciar un sesgo de parte del comprador y/o concentración de poder de mercado del proveedor. La detección de Comunidades es una herramienta para identificar la existencia de este tipo de comportamientos. En este sentido, las Comunidades son conjuntos de proveedores que se relacionan a través del intercambio con una misma UC, por lo cual, su prolongación en el tiempo puede evidenciar la existencia de cierto

<sup>1</sup>Elaboración propia en base a información relevada de la web de ARCE; <https://www.comprasestatales.gub.uy/>.

poder de mercado en los proveedores que conforman dicha comunidad y/o la presencia de un sesgo positivo de la UC hacia dichos proveedores.

Con el objetivo de detectar la existencia de Comunidades, se proponen las siguientes preguntas de investigación e hipótesis:

- $P_1$  ¿Existen comunidades a nivel de proveedores que generen cierto poder de lobby y pueda disminuir la eficiencia en las compras del Estado?
- $H_1$  Existen ciertos grupos de proveedores (comunidades) que sistemáticamente son adjudicados por ciertas UC
- $P_2$  ¿Qué peso relativo tienen las comunidades en el total de las compras y contrataciones del Estado?
- $H_2$  El peso relativo de las comunidades no supera el 30 % del total de compras y contrataciones del Estado.

La Base de Datos (BD) a utilizar se obtiene de la página de “Cuentas Claras” (<http://cuentasclaras.uy/#/database>). En la misma se disponibilizan, en formato CSV, los datos de las Compras y Contrataciones Estatales realizadas a través del “Sitio de Compras” de la ARCE, para el período entre 2015-2018. Partiendo de estos datos se identifican 19.537 adjudicaciones en el período, las cuales involucran a 256 UC (54 Incisos), 7.013 proveedores y 15.266 productos (a nivel de Artículo en el Catálogo de bienes, servicios y obras).

## II. TRABAJOS RELACIONADOS

Como estructurador del análisis y la detección de comunidades realizada, se considera el trabajo elaborado por F. H. Leiva, R. Torres, O. Nicolis and R. Salas F., para IEEE Access, titulado “Characterization of the Chilean Public Procurement Ecosystem Using Social Network Analysis”(2020) [6]. En el mismo, los autores buscan caracterizar el ecosistema de Compras Públicas de Chile, aplicando “social network analysis” para detectar las principales comunidades de proveedores. La metodología que utilizan es, primero, representar la relación comprador-proveedor como un grafo bipartito, para luego proyectarlo a un monoptito de proveedores. Para finalizar utilizan un algoritmo de modularidad para detectar las principales comunidades de proveedores [6].

Siguiendo la metodología aplicada por los autores del trabajo de referencia, primero se crea un grafo bipartito proveedor - UC y se presentan algunas estadísticas resumen que intentan caracterizar algunos puntos claves del grafo. Luego se proyecta un grafo monopartito de proveedores considerando las adjudicaciones realizadas en el período 2015 -2018. Posteriormente con el grafo monopartito, se identifican las comunidades de proveedores utilizando tres criterios distintos y se selecciona la mejor según distintos indicadores. Con el mejor método de partición seleccionado en la etapa anterior, se procede a identificar las comunidades en cada año por separado, lo cual permitirá caracterizar la evolución de las comunidades y detectar posibles sesgos.

## III. TRATAMIENTO PRELIMINAR DE LA BASE DE DATOS

El tratamiento de los datos así como todo el procesamiento se puede encontrar en el link de Gitlab.<sup>2</sup>

La base de datos inicial se obtiene en la web de “Cuentas Claras”. En la misma se encuentran los datos estructurados en las siguientes planillas:

- *2015\_awa\_items.csv, 2016\_awa\_items.csv, 2017\_awa\_items.csv, 2018\_awa\_items.csv*, contiene información sobre los productos (items) y proveedores adjudicados en cada compra, en el año correspondiente.
- *2015\_releases.csv, 2016\_releases.csv, 2017\_releases.csv, 2018\_releases.csv* contiene los datos algunos datos de la oferta, inclusive la Unidad de Compra (UC) que publica el llamado.
- *2015\_awa\_suppliers.csv, 2016\_awa\_suppliers.csv, 2017\_awa\_suppliers.csv, 2018\_awa\_suppliers.csv* contienen información descriptiva de los proveedores adjudicados.

A su vez, se decide complementar con el nombre de los Incisos y Unidades Ejecutoras, para lo cual se utiliza como fuente la planilla ‘Organismos.xls’ que contiene los códigos con los nombres de cada Organismo y su Unidad Ejecutora. La misma se obtiene de la página web de la ARCE.<sup>3</sup>

Una vez que se cuenta con todas las bases de datos, se procede a unirlos y se eliminan las columnas que no se van a utilizar. A su vez, se verifica que todas las variables tengan el tipo de dato correcto, modificando el de la variable “fecha” a ‘DateTime’ para poder utilizar los métodos propios de este tipo de objeto. Al mismo tiempo, se incorpora una columna calculada con el año que facilitará el planteo del modelo.

Por otro lado, se calcula la columna *Val\_Tot* que refleja el valor total de los ítems a nivel de cada compra en millones de pesos uruguayos como la multiplicación de la cantidad que se adquiere de cada ítem por el precio por unidad. Esta variable luego será utilizada para dar los pesos a las aristas del grafo.

<sup>2</sup>[https://gitlab.com/diego.mazzuco.ortiz/grupo\\_6\\_bdnr/-/blob/main/47084203\\_Detecci%C3%B3n\\_de\\_Comunidades\\_en\\_el\\_Mercado\\_de\\_Compras\\_Estatales.ipynb](https://gitlab.com/diego.mazzuco.ortiz/grupo_6_bdnr/-/blob/main/47084203_Detecci%C3%B3n_de_Comunidades_en_el_Mercado_de_Compras_Estatales.ipynb)

<sup>3</sup>[https://observatorio.arce.gub.uy/portal/c/document\\_library/get\\_file?uuid=4ef05cc7-3101-4e2e-8b0f-d6eff4912d33&groupId=13310](https://observatorio.arce.gub.uy/portal/c/document_library/get_file?uuid=4ef05cc7-3101-4e2e-8b0f-d6eff4912d33&groupId=13310)

Se detecta la existencia de Missing Values en las columnas: *precio\_unidad*, *moneda*, *unidad\_medida*, *id\_oferta*, *oferta\_metodo*, *val\_tot*. Lo cual se debe en parte a que algunas adjudicaciones no cuentan con llamado, por lo cual no existió una oferta y no existe información sobre la misma, generando valores faltantes en las variables *id\_oferta* y *oferta\_metodo*. Como esto no afecta el análisis, se decide dejarlos e indicarlos con *no\_of*. Sin embargo, en los otros casos donde se presentan NA se decide eliminarlos, ya que no se cuenta con información sobre los mismos.

Al mismo tiempo, se identifica un error en el ingreso del identificador de los proveedores (RUT), por lo cual se decide eliminar el símbolo para prevenir posible duplicación. Por último, a modo de simplificar y reducir la base, así como ser coherente con el trabajo de referencia, se consideran las siguientes restricciones a nivel de compra y cantidad de proveedores adjudicados por compra:

- El Valor Total de la compra supera el primer cuartil del 25
- La compra involucra a más de un proveedor

Tal como se plantea en el trabajo de referencia, parte del comportamiento que se quiere detectar son las posibles colusiones entre proveedores, por lo que interesa contemplar aquellas compras que son adjudicadas a más de un proveedor. El primer cuartil de compras alcanza los 3.570 pesos uruguayos, siendo un total de 154.441 compras las que lo superan. De este total, 17.568 compras fueron adjudicadas a 2 o más proveedores.

#### IV. DESCRIPCIÓN DE LA BASE DE DATOS FINAL

La base de datos final cuenta con 17.568 adjudicaciones que involucran a 54 incisos, 253 unidades ejecutoras, a 4.708 proveedores y 15.260 ítems. A continuación se presenta la evolución en el período de la cantidad de ítems, de compras y el monto de las mismas.

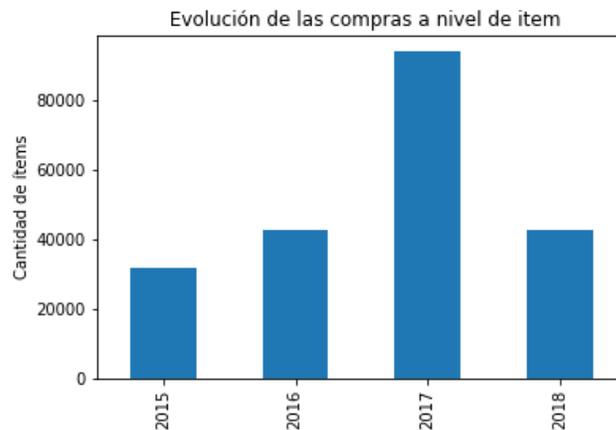


Figura 1: Evolución de las compras a nivel de ítem.

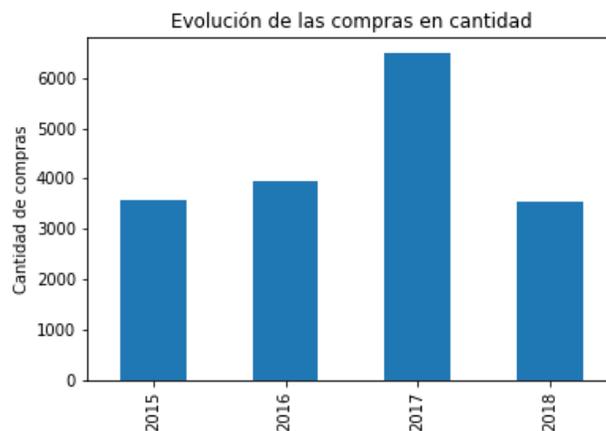


Figura 2: Evolucion de las compras en cantidad

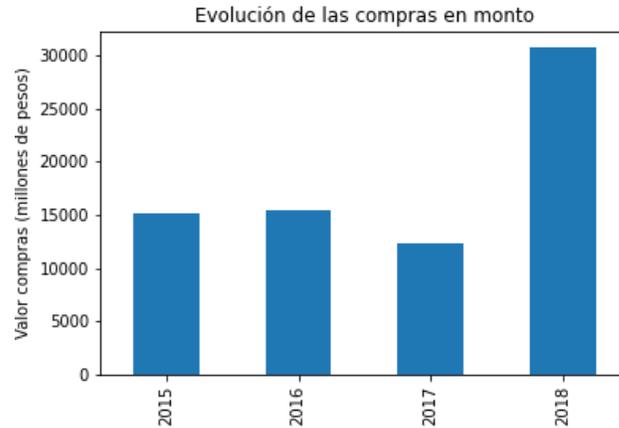


Figura 3: Evolucion de las compras en monto

Se observa que la cantidad de compras y su valor se comportan de manera opuesta: mientras que las cantidades aumentan entre 2015 y 2017, en el 2018 disminuyen; sin embargo, en términos de valor se mantienen relativamente constante entre 2015 y 2017, presentando un pico en el 2018. Esta discrepancia en 2017 y 2018 suena anti intuitiva, no obstante puede deberse a cientos de causas tener menores compras, por ejemplo un cambio en la política de las compras, un cambio los proyectos que se tienen, factores económicos, etc.

## V. CREACIÓN DE LA RED

Para crear el grafo se considera que el MCCE funciona con la siguiente lógica: un proveedor es adjudicado en una o varias compras de una UC. Por tanto, el grafo es bipartito y no dirigido contando con dos tipos de nodos: las UC y los proveedores. Se opta por plantearlo como un grafo no dirigido porque en el trabajo de referencia toman esta resolución y el abanico de funciones y métodos desarrollados para Digrafos es menor.

Al mismo tiempo, las aristas vinculan los proveedores a las UC, siendo la importancia de cada uno (peso) dada por el monto total vendido a cada UC. Además como el grafo es dinámico, se agrega el año como un atributo de los vértices. El grafo bipartito está conformado por 4.932 nodos - 4.679 proveedores y 253 UC- y por 21.258 aristas.

## VI. ANÁLISIS DE LA RED

Para tener una idea global de la red se decide analizar algunas estadísticas de conectividad y centralidad del grafo.

### VI-A. Conectividad

Un grafo se considera conectado todos los vértices son alcanzables desde cualquier otro [7]. En este caso el grafo no está conectado, sino que está dividido en componentes, es decir conjuntos de nodos relacionados entre sí por alguna ruta en el subgrafo, pero no conectados con otros nodos en el global [7]. En este sentido, se identifican 3 componentes de los cuales uno es el principal, concentrando 4.923 nodos. En la tabla I se presenta un resumen de la información de los dos nodos más pequeños que cuentan con 3 y 6 nodos.

Se observa que ambos componentes refieren a compras realizadas en el 2018, exceptuando únicamente una realizada en el 2017. En base a el resultado, no se puede afirmar que existan componentes que se prolonguen en el tiempo, es decir, no existe un conjunto de proveedores y UC que mantengan un intercambio exclusivo en el período.

Cmp	Id Proveedor	Nombre Inc	Nombre UE	Año	Peso
2	R130066990015	Adm. de Servicios de Salud del Estado	Red de Atención Primaria de Río Negro	2018	0.980
2	R130090820019	Adm. de Servicios de Salud del Estado	Red de Atención Primaria de Río Negro	2018	0.622
3	R200078780015	Intendencia de Treinta y Tres	Junta Departamental de Treinta y Tres	2018	0.016
3	R200095280017	Intendencia de Treinta y Tres	Junta Departamental de Treinta y Tres	2017	0.001
3	R200095280017	Intendencia de Treinta y Tres	Junta Departamental de Treinta y Tres	2018	1.954
3	R200103660014	Intendencia de Treinta y Tres	Junta Departamental de Treinta y Tres	2018	0.007
3	R200138410014	Intendencia de Treinta y Tres	Junta Departamental de Treinta y Tres	2018	0.040
3	R200175900011	Intendencia de Treinta y Tres	Junta Departamental de Treinta y Tres	2018	6.848

Cuadro I: Tabla con los nodos de los componentes 2 y 3

## VI-B. Centralidad

A continuación se realizan algunas medidas de centralidad para dimensionar la importancia de los distintos nodos e identificar posibles comportamientos en torno a la Red.

*VI-B1. Distribución del grado:* Al ser un grafo bipartito se analiza la distribución de cada tipo de nodo por separado.

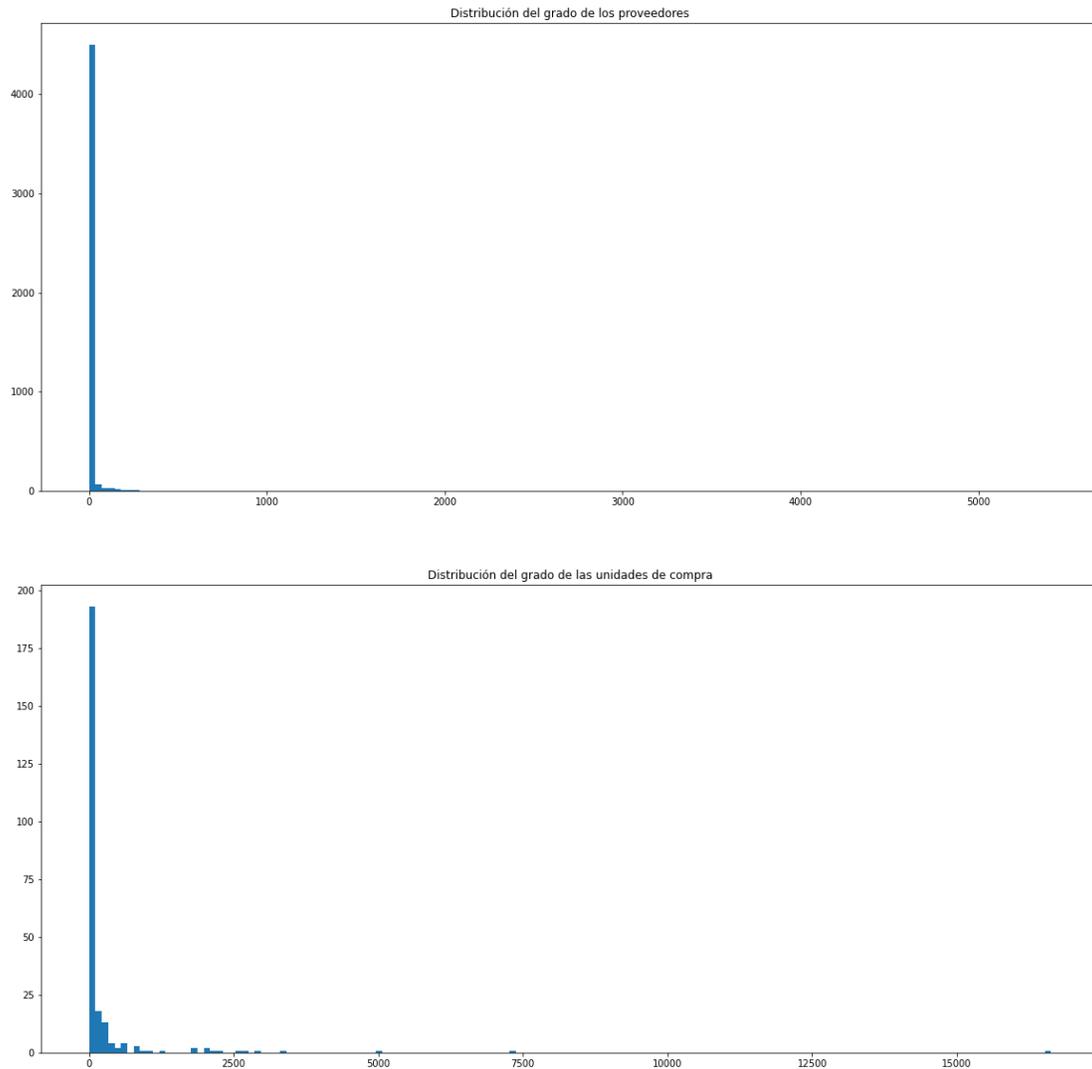


Figura 4: Distribucion de los grados segun unidad de compra y proveedores.

```

Cantidad de nodos de tipo proveedor por rango de grados

(0, 1]          3328
(1, 5]          704
(5, 10]         181
(10, 20]        148
(20, 50]        140
(50, 200]       119
(200, 500]      35
(500, 15000]    24
dtype: int64

Cantidad de nodos de tipo UC por rango de grados

(0, 1]          44
(1, 5]          40
(5, 10]         27
(10, 20]        24
(20, 50]        27
(50, 200]       47
(200, 500]      21
(500, 15000]    22
(15000, 20000]  1
dtype: int64

```

Figura 5: Cantidad de nodos por rango de grados.

De la distribución de grado se observan dos características en particular. La primera es que una proporción significativa de los nodos de ambos tipos tienen un grado menor a 1. En este sentido, los nodos de tipo proveedor con grado menor a 1 son 3.328 de 4.708, mientras que los de tipo UC alcanzan los 44 de 253. La segunda es que existen algunos nodos que tienen grados ponderado relativamente alto, es así que si se consideran los nodos de tipo proveedor que corresponden al 5 % mayor en grado ponderado - superan los 30,78 millones de pesos - se identifican 236, mientras que si se consideran los nodos de tipo UC - superan los 1.465,23 millones de pesos uruguayos - se obtienen los siguientes 13 nodos:

	UC	Grado	Ponderado	id_inc	nombre_inc	id_ue	nombre_ue
7	9-3	16632	9	Ministerio de Turismo y Deporte	3	Dirección Nacional de Turismo	
6	16-101	7371	16	Poder Judicial	101	Poder Judicial	
3	6-1	5036	6	Ministerio de Relaciones Exteriores	1	Ministerio de Relaciones Exteriores	
5	12-103	3344	12	Ministerio de Salud Pública	103	Dirección General de la Salud	
11	3-4	2913	3	Ministerio de Defensa Nacional	4	Comando General del Ejército	
8	3-33	2717	3	Ministerio de Defensa Nacional	33	Dirección Nacional de Sanidad de las Fuerzas A...	
9	25-2	2618	25	Administración Nacional de Educación Pública	2	Consejo de Educación Primaria	
4	61-1	2216	61	Administración Nacional de Usinas y Trasmision...	1	Administración Nacional de Usinas y Trasmision...	
12	27-1	2147	27	Instituto del Niño y Adolescente del Uruguay	1	Instituto del Niño y Adolescente del Uruguay	
0	64-1	2025	64	Administración Nacional de Puertos	1	Administración Nacional de Puertos	
2	29-18	1986	29	Administración de Servicios de Salud del Estado	18	Centro Departamental de Salud Pública de Colonia	
1	83-1	1815	83	Intendencia de Colonia	1	Intendencia Municipal de Colonia	
10	84-1	1793	84	Intendencia de Durazno	1	Intendencia Municipal de Durazno	

Figura 6: Principales nodos de tipo UC según grado.

Se observa que el nodo con mayor grado ponderado en el período es el correspondiente al Inciso “Ministerio de Turismo y Deporte” y Unidad Ejecutora “Dirección Nacional de Turismo”, el mismo alcanza a 16.632 millones de pesos uruguayos. Como el grado no captura la correlación entre los grados [8], se complementa con el cálculo del ‘PageRank’ considerando que este refleja la importancia de las UC en función a la cantidad de empresas que la proveen (cuán diversificadas tienen las compras en términos de proveedores) y el poder de dichos proveedores (ventas a distintas UC ponderadas por el peso). Los resultados obtenidos para el top 10 de UC se presentan a continuación.

UC	PageRank	id_inc	nombre_inc	id_ue	nombre_ue	
4	3-4	0.0205	3	Ministerio de Defensa Nacional	4	Comando General del Ejército
2	16-101	0.0165	16	Poder Judicial	101	Poder Judicial
8	3-33	0.0155	3	Ministerio de Defensa Nacional	33	Dirección Nacional de Sanidad de las Fuerzas A...
5	27-1	0.0130	27	Instituto del Niño y Adolescente del Uruguay	1	Instituto del Niño y Adolescente del Uruguay
6	9-3	0.0127	9	Ministerio de Turismo y Deporte	3	Dirección Nacional de Turismo
1	25-2	0.0126	25	Administración Nacional de Educación Pública	2	Consejo de Educación Primaria
3	3-18	0.0123	3	Ministerio de Defensa Nacional	18	Comando General de la Armada
7	98-1	0.0107	98	Intendencia de Montevideo	1	Intendencia Municipal de Montevideo
9	25-4	0.0100	25	Administración Nacional de Educación Pública	4	Consejo de Educación Técnico-Profesional
0	15-1	0.0094	15	Ministerio de Desarrollo Social	1	Ex-Adm. Gral- Dir. General Secret.

Figura 7: Page Rank top 10 de UC.

Considerando el indicador ‘PageRank’ las principales UC son el “Ministerio de Defensa Nacional - Comando General del Ejército” y el “Poder Judicial”. Según el indicador ambas UC compran a proveedores que son relevantes en función de los montos vendidos en el período. Considerando el resultado anterior con respecto al grado ponderado, se observa que el “Poder Judicial” es la segunda UC en compras y además, es de las que más compra a proveedores importantes en ventas.

*VI-B2. Distribución de intermediación:* Indica en qué proporción un vértice se encuentra entre otro par de vértices. La importancia de un vértice depende de su ubicación en la ruta del grafo, ya que considerando que se traspasa información entre nodos, los que se encuentren entre la mayor cantidad de nodos, son los más críticos para la comunicación de la información [7].

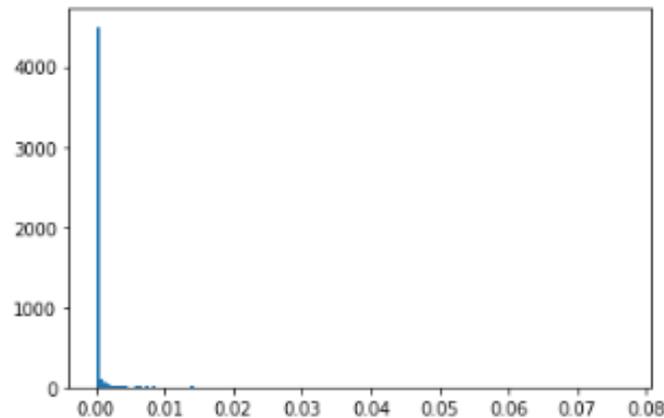


Figura 8: Distribución del indicador de intermediación.

```
Cantidad de nodos por rango de intermediación
(0.0, 0.005]      2120
(0.005, 0.01]    55
(0.01, 0.03]     64
(0.03, 0.1]      13
dtype: int64
```

Figura 9: Cantidad de nodos por rango de intermediación

A continuación se identifican los 13 nodos que se encuentran en el rango más alto.

nodo	intermediación	id_inc	nombre_inc	id_ue	nombre_ue	
10	98-1	0.077	98.0	Intendencia de Montevideo	1.0	Intendencia Municipal de Montevideo
6	3-18	0.060	3.0	Ministerio de Defensa Nacional	18.0	Comando General de la Armada
1	5-3	0.048	5.0	Ministerio de Economía y Finanzas	3.0	Auditoría Interna de la Nación
4	16-101	0.047	16.0	Poder Judicial	101.0	Poder Judicial
11	3-33	0.046	3.0	Ministerio de Defensa Nacional	33.0	Dirección Nacional de Sanidad de las Fuerzas A...
12	R214854290012	0.041	NaN	NaN	NaN	NaN
9	29-27	0.040	29.0	Administración de Servicios de Salud del Estado	27.0	Centro Departamental de Rocha
7	3-4	0.038	3.0	Ministerio de Defensa Nacional	4.0	Comando General del Ejército
2	15-1	0.037	15.0	Ministerio de Desarrollo Social	1.0	Ex-Adm. Gral- Dir. General Secret.
3	25-2	0.036	25.0	Administración Nacional de Educación Pública	2.0	Consejo de Educación Primaria
8	4-26	0.036	4.0	Ministerio del Interior	26.0	Dir. Nac. Cárceles, Penitenciarias y Centros d...
0	4-1	0.034	4.0	Ministerio del Interior	1.0	Secretaría del Ministerio del Interior
5	11-3	0.033	11.0	Ministerio de Educación y Cultura	3.0	Dirección Nacional de Cultura

Figura 10: Principales nodos según intermediación

De los 13 nodos con mayor centralidad de intermediación, 12 son de tipo UC y 1 es de tipo proveedor, además los nodos que parecen ser más centrales con este indicador son la “Intendencia de Montevideo” y el “Ministerio de Defensa Nacional - Comando General de la Armada”. Se puede interpretar que estas UC son las que diversifican más sus compras y adjudican a mayor cantidad de proveedores.

*VI-B3. Cliques:* Cliques son subgrafos completos, es decir, todos los nodos que lo conforman están relacionados por vértices [7].

En el grafo UC- Proveedor, por definición no existen cliques, ya que la relación siempre es de un proveedor a un comprador, no existe relación ni entre proveedores ni entre compradores. Sin embargo, si no se considera la dirección de los vértices, los cliques serían todas las duplas de nodos que se relacionan a través de los mismos. En este caso existirían cliques como vértices tiene el grafo.

## VII. REDUCCIÓN DE LA RED

Como resultado del análisis anterior se identifica que el 71.1 % de los nodos de tipo proveedor tienen un grado ponderado menor a 1, mientras que en los nodos del tipo UC esta característica se presenta en un casi 17.4 %. Un grado ponderado menor a 1 quiere decir que la suma ponderada de las aristas incidentes por su peso no supera el uno, por lo que si existe más de una arista, el valor de las ventas/compras que se realizan en la totalidad del período no supera el millón de pesos uruguayos.

Considerando lo anterior se decide eliminar los nodos que tienen esta característica ya que, por un lado, no son significativos en valor de compras por lo que no aportan al objetivo del estudio y, por otro, se logra reducir la dimensionalidad del grafo facilitando su interpretación. Una vez reducido el grafo, el mismo cuenta con 1.560 nodos - 1351 de tipo Proveedor y 209 de tipo UC - y 13.703 aristas.

## VIII. PROYECCIÓN A UN GRAFO MONOPARTITO

Siguiendo el trabajo de referencia, se realiza la proyección del grafo bipartito a uno monopartito donde se vinculan solamente los proveedores en función de si son adjudicados a compras realizadas por la misma UC.

La proyección se realiza tanto a nivel de cada subgrafo anual, como del grafo compilado para todo el período. El procedimiento es el siguiente:

- Se proyecta un subgrafo por año.
- Se agrega el atributo del año a cada vértice de cada subgrafo proyectado según corresponda.
- Se compilan todos los subgrafos y se obtiene el grafo del período.

	Año	Nodos de tipo Proveedor	Nodos de tipo UC	Nodos total proyeccion	Aristas proyeccion
0	2015	713	17	713	95480
1	2016	774	47	774	106169
2	2017	923	88	923	130974
3	2018	831	75	831	79781
4	Total	1351	209	1351	232949

Figura 11: Cantidad de nodos por rango de intermediación

La cantidad total de nodos del grafo proyectado compilado es 1351 que corresponde a la cantidad de nodos de tipo de proveedor de la red original. La diferencia en la suma de nodos de cada grafo proyectado anual y el compilado se debe a que los proveedores varían anualmente, por lo que algunos se repiten entre años mientras que otros ingresan o salen del mercado.

En este momento teniendo el grafo reducido y proyectado se decidió guardar el mismo en neo4j. De forma poder buscar fácilmente relaciones entre diferentes proveedores y facilitar estudios futuros.

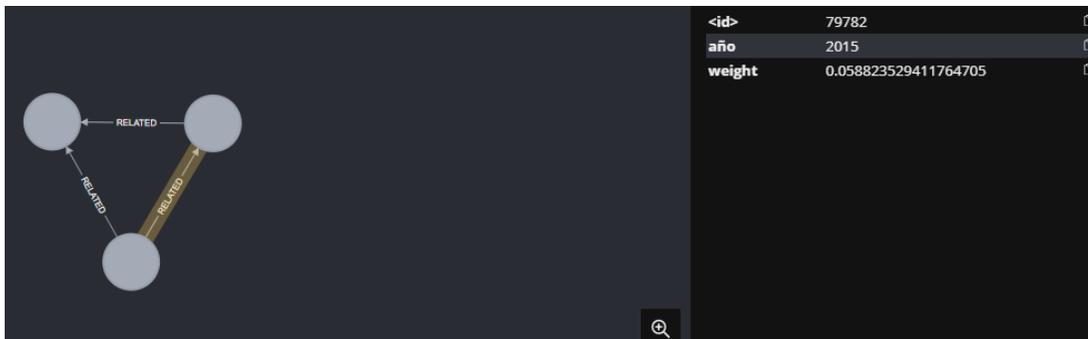


Figura 12: Tres proveedores con sus relaciones

## IX. DETECCIÓN DE COMUNIDADES

La detección de comunidades se realiza en una primera instancia para el grafo compilado del período total y luego para cada año por separado para detectar continuidad en algún patrón. La detección de comunidades en el grafo compilado se realiza por tres métodos distintos: Louvain, Clauset-Newman-Moore greedy modularity maximization y Girvan Newman.

Louvain es un algoritmo jerárquico aglomerativo basado en optimización por modularidad. V. Tunali (2021) en su trabajo titulado “Large-Scale Network Community Detection Using Similarity-Guided Merge and Refinement” precisa que el método primero crea comunidades compuestas por un único nodo y luego agrupando y cambiando nodos entre comunidades controla cual agrupación maximiza la modularidad. El resultado de cada etapa es una partición jerárquica [9].

Por otro lado, V. Tunali (2021) presenta al Método Clauset-Newman-Moore como uno de los primeros intentos para encontrar comunidades en redes grandes. Este método es similar al anterior pero a diferencia del Louvain utiliza un algoritmo aglomerativo mediante una optimización “greedy” de la modularidad [9].

El último método es Girvan Newman denominado en alusión a sus creadores Mark Newman y Michelle Girvan quienes propusieron un algoritmo para identificar comunidades basado también en la medida de modularidad (F. H. Leiva et al., 2020) pero, a diferencia de los anteriores, el método es divisivo eliminando progresivamente aristas del grafo original. La selección de las aristas la hace en función a la centralidad de intermediación de las aristas [9].

Una vez que se tienen detectadas las comunidades en función a cada método, se procede a compararlas utilizando una “función de calidad”. Según S. Fortunato (2009) una función de calidad asigna un valor a cada partición de un grafo, permitiendo ordenarlas en función al puntaje obtenido [9]. En este sentido, se decide utilizar tres indicadores para comparar la calidad de las particiones, los mismos son: performance, coverage y modularity.

La “performance” cuenta la cantidad de pares de vértices que están bien clasificados, es decir si dos vértices pertenecen a una misma comunidad deben estar conectados por una arista mientras que si no pertenecen a la misma comunidad no deben estar conectados; su valor se encuentra entre 0 y 1. La “coverage” es el ratio de las aristas intra-comunidad en el total de aristas del grafo. Por último, la “modularity” se define como la fracción de aristas dentro de una comunidad del grafo menos la fracción esperada de aristas en un grafo aleatorio. Este último indicador es menor que 1 y puede llegar a ser negativo [10]. Según S.

Fortunato (2009) el indicador de “modularity” no debe ser utilizado para comparar grafos que difieren significativamente de tamaño, en este caso como el grafo es el mismo, lo que difieren son las comunidades, no debería existir ningún problema [10]. Los resultados de cada método de detección de comunidades para el grafo compilado se presenta a continuación:

	0	1	2	3	4	5	6
Metodo							
Louvain	163	463	299	192.0	93.0	76.0	63.0
Clauset-Newman-Moore	715	412	134	88.0	2.0	0.0	0.0
Girvan_Newman	1348	2	1	0.0	0.0	0.0	0.0

Figura 13: Las comunidades detectadas por los metodos con su cantidad de nodos

Al controlar el desempeño de cada método con los indicadores se obtiene lo siguiente:

	Coverage	Performance	Modularity
Metodo			
Louvain	0.32	0.70	0.16
CNM	0.53	0.63	0.11
GN	1.00	0.26	0.00

Figura 14: Desempeño de los diferentes algoritmos para detectar comunidades

De lo anterior se obtiene que el método que divide mejor el grafo en comunidades es Louvain, en este sentido es el que obtiene mayor indicador de “modularity” por lo que logra contemplar la mayor cantidad de conexiones dentro de la comunidad y dejar muy pocas entre distintas comunidades. Al mismo tiempo, obtiene la mejor “performance” por lo que agrupa mejor a los vértices que están conectados y deja aquellos que no por fuera. Sin embargo, no es el que obtiene mayor “coverage”, no considera algunas aristas intra-comunidad.

## X. EVOLUCIÓN DE LAS COMUNIDADES

En base a lo anterior se selecciona el Método Louvain como el que mejor permite detectar comunidades en función al indicador de modularidad y de performance. Ahora se utilizará este método para observar cómo se comportan las comunidades a lo largo del período, analizando su evolución año a año.

```
La cantidad de comunidades identificadas con el método Louvain son:
2015: 6 {0: 159, 1: 198, 2: 185, 4: 132, 3: 17, 5: 22}
2016: 3 {0: 298, 1: 288, 2: 188}
2017: 4 {1: 317, 2: 307, 0: 207, 3: 92}
2018: 7 {0: 82, 1: 219, 2: 205, 3: 110, 4: 37, 5: 176, 6: 2}
```

Figura 15: Comunidades identificadas con el método Louvain

A continuación, se estudian distintas características de los subgrafos anuales para identificar patrones en la evolución de las comunidades.

En primer lugar, se compara cada comunidad de cada año con el año posterior a ver si existe isomorfismo entre comunidades. Para ello se crea un subgrafo de cada comunidad y luego se itera entre ellas comparando una a una entre años. Como resultado no se encuentra ningún indicio de isomorfismo entre las comunidades de cada año.

En segundo lugar, se compara la Modularidad de cada año que permitirá observar si existe algún año donde la detección de comunidades sea significativamente mejor o peor.

## Indicador de modularidad

```

2015: 0.17
2016: 0.15
2017: 0.15
2018: 0.16

```

Figura 16: Indicador de Modularidad

Se observa que se mantiene aproximadamente el mismo nivel de modularidad en el período por lo que no se identifica una diferencia significativa en la detección de comunidades entre años.

En tercer lugar, se decide observar alguna medida de centralidad con respecto a cada comunidad para dimensionar la distribución de la importancia de los nodos. Aquellas comunidades que cuenten con una medida de centralidad mayor, se puede considerar que son las más relevantes ya que significa que dichos proveedores presentan gran interrelación. Si se traduce lo anterior al grafo bipartito indicaría que existen algunas UC que interactúan con muchos proveedores en común. Se selecciona el grado promedio como medida de centralidad.

```

El grado promedio de cada comunidad es:
2015: {0: 245.62, 1: 316.99, 2: 303.16, 4: 227.63, 3: 36.82, 5: 108.45}
2016: {0: 229.41, 1: 327.81, 2: 263.64}
2017: {1: 323.12, 2: 277.78, 0: 294.66, 3: 143.98}
2018: {0: 80.77, 1: 210.57, 2: 279.42, 3: 268.38, 4: 72.89, 5: 98.42, 6: 1}

```

Figura 17: Grado promedio de cada comunidad por año

En función a la distribución del grado promedio de los nodos se observa que en el año 2015 la comunidad con el grado promedio mayor fue la 1, en el 2016 la 1, en el 2017 la 1 y en el 2018 la 2.

A continuación se sustrae solamente estas comunidades y se comparan para detectar si los nodos involucrados son iguales. Como resultado se obtiene que, existe un único nodo que integra las comunidades más importantes en términos de ventas promedio anuales, a lo largo del período considerado. En la siguiente tabla se presenta dicho nodo junto a la cantidad de nodos vecinos de tipo UC pertenecientes al grafo bipartito.

Cantidad de vecinos	
R216558290018	12

Figura 18: Nodo persistente con cantidad de nodos vecinos

Se observa que la cantidad de UC que adjudican compras a dicho proveedor es relativamente alta, lo que no parecería indicar un sesgo positivo de alguna UC hacia un proveedor particular, sino que el hecho de que varias UC los prefieran a lo largo del período señalaría una posible ventaja comparativa de los mismos frente a la competencia que inclina a elegirlos a ellos.

Por último, se calcula el peso relativo de cada comunidad en las ventas totales por año en función al grado total de cada una. Esto permitirá responder la segunda pregunta de investigación.

```
Peso relativo de cada comunidad:
2015: {0: 20, 1: 33, 2: 29, 4: 16, 3: 0, 5: 1}
2016: {0: 32, 1: 44, 2: 23}
2017: {1: 39, 2: 33, 0: 23, 3: 5}
2018: {0: 4, 1: 29, 2: 36, 3: 19, 4: 2, 5: 11, 6: 0}
```

Figura 19: Peso relativo de cada comunidad.

Se observa que la segunda hipótesis no se cumple, debido a que, en todos los años existen comunidades que alcanzan un peso relativo en las ventas/compras anuales mayor al 30

## XI. CONCLUSIONES

Utilizando el algoritmo de detección de comunidades desarrollado por Louvain, aplicado al grafo monopartito generado a partir de la proyección del grafo bipartito que representa las transacciones entre las UC y los Proveedores en el MCCE en el período 2015-2018, se logran responder las preguntas de investigación y contrastar las hipótesis relacionadas.

Por un lado, se pudo comprobar que si bien existe un único nodo que ha conformado las comunidades más importantes en término de ventas promedio (grado promedio) en el transcurso de todo el período, no se puede confirmar que existan UC que sistemáticamente opten por adjudicar a dicho proveedor.

Por otro lado, considerando el peso relativo de las comunidades a lo largo del período, se comprueba que existen comunidades con un peso relativo anual mayor al 30 % planteado como hipótesis, por lo cual la misma no se cumple.

Por último, queda pendiente del actual trabajo ampliar el estudio a los ítems y utilizar la información sectorial de los proveedores para detectar comunidades específicas por sector al cual pertenece dicho proveedor.

## REFERENCIAS

- [1] (2008) Ley N° 19.924 “Presupuesto Nacional de Sueldos Gastos e Inversiones. Ejercicio 2020-2024”. [Online]. Available: <https://www.impo.com.uy/bases/leyes/19924-2020>
- [2] (2008) Ley N° 18.362 “Aprobación de la Rendición de Cuentas y Balance de Ejecución Presupuestal. Ejercicio 2007”. [Online]. Available: <https://www.impo.com.uy/bases/leyes/18362-2008>
- [3] (2007) Ley N° 18.159 “DEFENSA DE LA LIBRE COMPETENCIA EN EL COMERCIO”. [Online]. Available: <https://www.impo.com.uy/bases/leyes/18159-2007>
- [4] Ministerio de Economía y Finanzas. La competencia como un elemento sustancial del sano funcionamiento de los mercados. [Online]. Available: <https://www.gub.uy/ministerio-economia-finanzas/competencia-elemento-sustancial-del-sano-funcionamiento-mercados>
- [5] (2008) Ley N° 18.381 “Ley sobre el Derecho de Acceso a la Información Pública”. [Online]. Available: <https://www.impo.com.uy/bases/leyes/18381-2008>
- [6] F. H. Leiva, R. Torres, O. Nicolis, and R. Salas F., “Characterization of the chilean public procurement ecosystem using social network analysis,” *IEEE Access*, vol. 8, pp. 138 846–138 858, 2020. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9149588>
- [7] E. D. Kolaczyk and G. Csárdi, *Statistical analysis of network data with R*. Springer, 2014, vol. 65.
- [8] N. DQ. The degree distribution of a network. [Online]. Available: [http://mathinsight.org/degree\\_distribution](http://mathinsight.org/degree_distribution)
- [9] V. Tunali, “Large-scale network community detection using similarity-guided merge and refinement,” *IEEE Access*, vol. 9, pp. 78 538–78 552, 2021.
- [10] S. Fortunato, “Community detection in graphs,” *Physics reports*, vol. 486, no. 3-5, pp. 75–174, 2010.