

Uso de bases de datos de grafos en el rastreo de contactos de casos positivos de COVID-19

Laura Serra

Facultad de Ingeniería, Universidad de la República

Montevideo, Uruguay

laura.serra.brigante@gmail.com

Resumen—Este documento releva información sobre la posibilidad de rastrear contactos de casos positivos de COVID-19 mediante el uso de bases de datos de grafos. Se presenta la realidad de una empresa del rubro de la construcción cuyos trabajadores comparten vivienda y transporte. Por la forma en la cuál la empresa guarda la información actualmente, cuando aparece un caso positivo, el rastreo de quiénes deben ser aislados se hace a través de un proceso manual y engorroso. Se propone una solución mediante la herramienta Neo4j que permite unificar la información y así obtener resultados más eficientes.

I. INTRODUCCIÓN

Desde el pasado 11 de marzo de 2020, dónde la Organización Mundial de la Salud (OMS) declaró la enfermedad del nuevo coronavirus como una pandemia [1], ha habido más de 190 millones de casos positivos, cuatro millones de muertes a lo largo del mundo, y los números aumentan minuto a minuto.

Más allá de la llegada de la vacuna a numerosos países, muchos mantienen aún un estado de emergencia y a la fecha no existe información certera de cuándo se pondrá fin a esta pandemia.

¿En qué se diferencia el SARS-Cov-2 (virus que provoca el COVID-19) de otros virus similares que ya han azotado al mundo? La respuesta es universalmente conocida: su gran capacidad de transmisión junto con la existencia de numerosos casos que transcurren la enfermedad sin presentar síntomas. El distanciamiento y aislamiento social, el uso de tapabocas y las burbujas son conceptos que parecen haber llegado para quedarse.

Bajo esta nueva realidad y en un contexto donde la tecnología forma parte inequívoca de nuestras vidas cotidianas, resulta fundamental su uso para combatir y contener la expansión del virus. Las diferentes comunidades científicas están en una búsqueda constante para generar modelos que capturen la dinámica de la enfermedad, informando sobre su naturaleza y generando medidas de manejo exitosas [2].

Este trabajo propone el uso de bases de datos de grafos para poder detectar y notificar rápidamente a los contactos de casos positivos de COVID-19 para una empresa constructora que en la actualidad presenta un proceso manual y engorroso a la hora de detectar, notificar y aislar a sus trabajadores cuando aparecen casos positivos.

Este documento se organiza de la siguiente forma. En la Sección II se presentan trabajos relacionados y se mencionan conceptos que se utilizarán a lo largo del trabajo. Luego,

en la Sección III se desarrolla el problema, se propone la implementación de un modelo para solucionarlo y se resuelven consultas para mostrar su utilidad. Por último, en la Sección IV se presentan las conclusiones y posibles trabajos futuros.

II. TRABAJOS RELACIONADOS

En esta sección se presentan algunos trabajos que se relacionan de con el objetivo de este documento y la realidad que se pretende modelar. Asimismo se menciona la situación de Uruguay en el marco de una posible aplicación de estos trabajos.

Desde que comenzó la pandemia, a nivel mundial, se han intentado varias formas para controlar su expansión. Rastreo de contactos manual mediante llamadas telefónicas, uso de aplicaciones que envían alertas en caso de haber estado en contacto con alguien positivo, cierre de fronteras, cuarentenas obligatorias, etc.

El análisis de datos también surge como herramienta para tratar de alivianar el avance de la pandemia. Numerosos estudios buscan obtener información concluyente sobre qué medidas se deben tomar para frenar la propagación del virus.

Una corriente plantea el uso de las bases de datos de grafos para detectar de forma rápida contactos de casos positivos de COVID-19. Una base de datos de grafos es un sistema especializado para crear y manipular grafos. Los grafos contienen nodos, aristas y atributos que se utilizan para representar y almacenar datos de una forma que las bases de datos relacionales tradicionales no permiten. Uno de los sistemas de gestión de bases de datos de grafos de uso libre y generalizado es *Neo4j*. En su página web [3] se encuentra una sección exclusiva para presentar trabajos vinculados al COVID-19.

En '*Extensive Representation of Contact Tracing with Neo4j in the Canton of Geneva*' [4] *Delphine Courvoisier et al.* comentan que la forma de detener la expansión de la pandemia es bloqueando las cadenas de transmisión mediante la cuarentena y el aislamiento de las personas.

Su trabajo plantea un caso de estudio para el Cantón de Ginebra, Suiza. Allí existe un organismo que recopila todos los resultados de los laboratorios que realizan pruebas de SARS-CoV-2. A todos los casos positivos se los llama y se identifica a sus contactos. De esta forma se obtiene la

información necesaria para representar en la base de datos de grafos. Los autores plantean que una visualización en forma de grafos, que muestra a las personas enfermas, con quienes estuvieron en contacto y en dónde, permite identificar los lugares con mayor riesgo de transmisión, relevar la existencia de las diferentes cadenas de contagio e incluso detectar cuales de estas son las más peligrosas y por ende, las que se deben accionar primero. Teniendo esta información disponible se pueden brindar herramientas para la toma de decisiones al gobierno.

En este tipo de trabajos se identifican al menos dos grandes problemas. En primer lugar, el carácter sensible de los datos de las personas contagiadas implica que actualmente no haya ninguna base estandarizada para su uso en investigación. A su vez, el carácter del virus en sí, con su enorme grado de transmisión y el hecho de que personas asintomáticas lo transmiten sin saberlo, implica que su avance es difícil de detener. Esto quiere decir que la pandemia avanza más rápido que la posibilidad de rastrear manualmente los contactos de los casos positivos.

En Uruguay, a partir de la detección de los primeros casos el 13 de marzo de 2020, el Ministerio de Salud Pública (MSP) contó con un área que se encargaba de llamar a los contactos de los casos positivos y notificarlos para que se aislaran e hisoparan. Esto se mantuvo mientras la situación fue controlable en el país. Pero, a partir de noviembre 2020 y con el crecimiento exponencial de los casos, el seguimiento 'caso a caso' se dejó de realizar.

Acceder a los datos de los contactos de los casos positivos no es tarea sencilla. En '*Digital Contact Tracing Based on a Graph Database Algorithm for Emergency Management During the COVID-19 Epidemic: Case Study*' [5] Zijun Mao et al. también plantean el uso de bases de datos de grafos para modelar contactos de casos positivos de COVID-19 pero la forma a la cuál acceden a la información es a través de un medio digital.

Los autores persiguen los mismos objetivos que en el caso anterior: brindar herramientas que ayuden a la contención de la pandemia y que sirvan de insumo para quienes toman decisiones. A su vez, presentan un método diferente para la obtención de los datos, buscando superar las limitaciones del rastreo manual. El trabajo plantea un caso de estudio para la provincia de Hainan, China. Allí, el gobierno implementó una plataforma centralizada que, entre otras cosas, permite el seguimiento de las personas contagiadas, de la población en general, vehículos y lugares públicos generando así una estricta vigilancia epidemiológica.

El método digital utiliza la tecnología para acceder a la información: se realiza un rastreo de la persona a través de su ubicación, del GPS de su celular, del Wi-Fi y del Bluetooth. Estos sistemas superan las limitaciones del rastreo manual en términos de errores, demoras y problemas de escalabilidad.

Los resultados de este trabajo en el seguimiento de los contactos de los casos positivos de COVID-19 y en la iden-

tificación de lugares de riesgo fueron exitosos. Los autores reconocen que la eficiencia de su modelo se debió a la gran coordinación entre diferentes organismos de la provincia, formando así una estructura organizada que apunta hacia el mismo objetivo.

Para el caso de Uruguay existe la aplicación de celulares *CoronavirusUy* que notifica cuando una persona estuvo cerca de alguien que dio positivo por un periodo prolongado de tiempo. Esto, a diferencia del enfoque mencionado presenta un método descentralizado en el cual la persona no sabe dónde ha sido, potencialmente, expuesta al virus, y las autoridades de vigilancia epidemiológica no tienen forma de reconstruir digitalmente los movimientos de la persona infectada originalmente, ni de la persona que ha sido notificada [6]

Existen otros trabajos interesantes, que si bien no parten de casos reales como los mencionados anteriormente, si demuestran el beneficio de utilizar bases de datos de grafos a la hora de rastrear contactos de casos de COVID-19 positivos.

Uno de ellos es el realizado por *Rik Van Bruggen* en su blog [7]. Allí, el autor, partiendo de un enfoque diferente y con varios supuestos, simula los datos de personas y visitas a determinados lugares y busca entender los contactos indirectos que existen entre personas sanas y enfermos de COVID-19. A través de la herramienta *Neo4j* genera relaciones de personas que visitan determinados lugares e identifica a los contactos de casos positivos observando lugares dónde personas saludables estuvieron en el mismo rango de tiempo que personas contagiadas. A su vez, en su análisis utiliza datos de coordenadas lo que hace interesante para añadir información geoespacial sobre los lugares con más riesgo.

III. DESARROLLO

En esta sección se presenta la realidad de la empresa constructora que se busca modelar, los requerimientos para ello, el diseño que se generó y sus principales aplicaciones.

III-A. Formalización del problema

Como se mencionó anteriormente y surge a partir de los trabajos relacionados, acceder a la información para modelar casos de COVID-19 no es tarea sencilla. Existen numerosas dificultades según el método que se utilice y además, prima la protección de los datos personales. Es por ello que este trabajo no se centra en crear un modelo a nivel país o región, sino que se enfoca en algo más reducido y aplicable, modelando la realidad de una empresa en particular a la cuál se pudo acceder a los datos.

Se parte de datos de la empresa coestructura RYENA (se utiliza un nombre ficticio). RYENA se encuentra actualmente trabajando en varias obras en una ciudad del interior del país. Cuenta con aproximadamente 1300 trabajadores entre obreros y cargos técnicos.

Muchos de los trabajadores no viven en el área por lo que se les otorgan viviendas compartidas. Para el caso de los obreros, son apartamentos con capacidad de ocho personas

especialmente diseñados. A su vez, RYENA cuenta con una flota de ómnibus a disposición y se encarga de brindarle a los trabajadores el transporte hacia la obra en la mañana y desde ella en la tarde.

En el contexto que se vive actualmente resulta crucial que la información de vivienda y transporte quede almacenada, ya que al momento de aparecer un caso positivo de coronavirus se deberá aislar, no solo al enfermo, sino también a sus contactos. Actualmente RYENA maneja dicha información en numerosas planillas de Excel, a cargo de sectores diversos de la organización.

Cuando aparece un caso positivo, el proceso se maneja en varias instancias. Dicho caso puede surgir desde la obra (cuando un trabajador presenta síntomas), desde el servicio de hisopados en sitio (existe a disposición un servicio de hisopados rápido) o por un reporte del trabajador en caso de que no vaya a trabajar. El caso es manejado por el área de seguridad de la obra que articula los diferentes sectores de la organización que contribuyen a su resolución. Los prevenciónistas del área se comunican con el área de recursos humanos, quienes son responsables de determinar con quienes vive la persona que dio positivo mediante el registro de viviendas de la empresa.

A su vez, el área de seguridad se comunica con el sector de logística para determinar los contactos por viaje. Se lleva un registro de en qué asiento de qué ómnibus se sientan las personas. Tanto el ómnibus como los asientos varían según el turno (mañana o tarde) o el día. Es decir que, a diferencia de los apartamentos, los cuales los trabajadores tienen fijados previamente, los asientos y los ómnibus varían. Cuando aparece un caso positivo no se aísla a todo el ómnibus sino que se aísla a los trabajadores de los asientos de adelante, de atrás y del costado del caso confirmado, ya que se entiende que estos son los que poseen riesgo de contagio.

Con toda esta información, el área de seguridad notifica a todos los trabajadores que pudieron haber tenido contacto con el caso positivo y les pide que se retiren de obra y consulten con su prestador de salud para poder hacerse un hisopado a los 7 días del último contacto. A estos trabajadores se les retira el acceso a la obra mediante la inutilización de su tarjeta de entrada y solo se les permitirá el acceso mediante un alta médica, posterior a un hisopado negativo. En caso de que el hisopado sea positivo, se repetirá el mismo proceso para la nueva persona afectada.

III-B. Implementación de base de grafos

Observando la realidad de RYENA se encuentra que si la información de los trabajadores, las viviendas y viajes que realizan estuviera almacenada en una base de datos de grafos, sería más sencillo identificar rápidamente a quienes habría que aislar cuando aparezca un caso positivo. Sería un proceso más fiable y ahorraría horas de trabajo a las diversas áreas.

Se utiliza como sistema de manejo de base de datos de grafos *Neo4j* en su versión 4.2.5 para el desarrollo del modelo. Su uso se justifica por su carácter de software libre y accesible.



Figura 1: Modelo para identificar quiénes comparten vivienda.

Los datos se almacenan en forma de nodos y aristas (relaciones). A su vez, cada nodo o arista puede contener atributos. El lenguaje de consultas que utiliza *Neo4j* se denomina *Cypher*.

Se tuvo acceso a ejemplos de las planillas que utiliza RYENA actualmente, con información de sus trabajadores, sus viviendas y sus viajes, para una obra en particular. En base a esta información y luego de realizar algunos ajustes para proteger los datos privados (utilizar un id aleatorio en vez de utilizar cédulas y asignar nuevos nombres y apellidos) se generaron archivos csv que posteriormente se cargaron a *Neo4j*. Los mismos tienen la siguiente estructura:

- **Trabajadores:** Guarda la información de los trabajadores: id, nombre, apellido, edad y sexo. Aquí también figura el apartamento en el que vive el trabajador, en caso de figurar NULL significa que el mismo vive en casa particular y no en una otorgada por la empresa.
- **Transporte:** Almacena la información general de los viajes, para cada trabajador, el día en que se realizó el viaje, en qué turno (mañana = 1 y tarde = 0), en qué ómnibus y en qué asiento se sentó el trabajador.
- **Omnibus:** Esta es una tabla auxiliar que almacena los números de coche de los ómnibus. Su utilidad es en hora de trasladar la información a *Neo4j*.

Los datos que finalmente fueron cargados a *Neo4j* a partir de los archivos mencionados, corresponden a la información de una sub-obra en particular que posee 250 trabajadores. Con respecto a los datos de transporte, se simulaban datos correspondientes a una semana laboral que va de lunes a sábado. Bajo este concepto fueron cargados 2261 registros. Cabe destacar que no todos los trabajadores utilizan el transporte de la obra todos los días.

Considerar una sub-obra dentro de las que posee RYENA se hizo para simplificar el modelo sin perder sus características

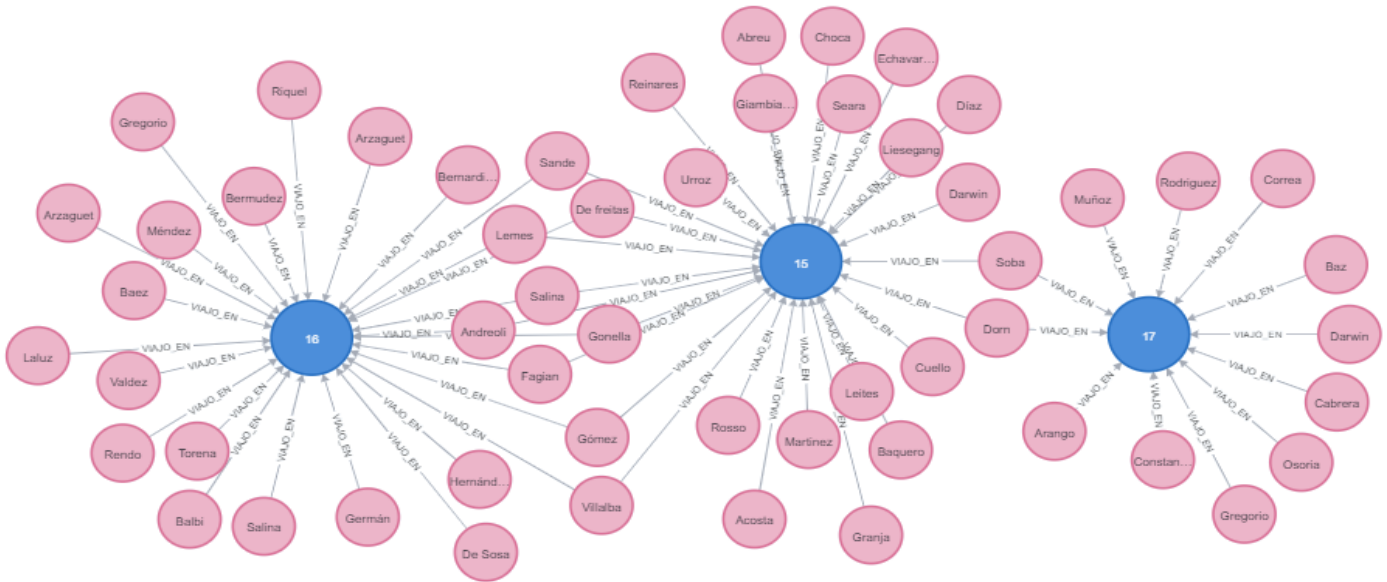


Figura 2: Modelo para identificar personas que compartieron ómnibus.

fundamentales. El mismo se puede expandir a todas las obras y para todos los días. El hecho de considerar una semana en los viajes fue principalmente debido a la importancia del dato temporal a la hora de detectar contactos: si un trabajador da positivo sin síntomas, se consideran dos días para atrás del hisopado para considerar sus contactos. Por otro lado, si un trabajador presenta síntomas, se retira de la obra y se le hace hisopado. Si el hisopado es positivo, se consideran dos días antes de presentar síntomas para considerar sus contactos.

Los nodos utilizados corresponden a los trabajadores y a los ómnibus. Por otro lado, las aristas que se crearon fueron las siguientes:

- **VIVE_CON:** Esta relación se da de trabajador a trabajador y es generada cuando los mismos tienen igual número dentro del atributo de *housing* (identificador de la vivienda). Como se mencionó anteriormente el límite de una vivienda es de ocho trabajadores, por lo tanto cada persona tendrá esta relación con otras siete. Dentro de esta relación se encuentra el atributo que identifica la vivienda.
- **VIAJO_EN:** Esta relación se da de trabajador a ómnibus y muestra en qué número de coche viajaron los distintos trabajadores. No se le exige a los trabajadores utilizar los mismos ómnibus, por lo tanto, dado un trabajador, el número de coche puede variar en el mismo día para el turno de la tarde con respecto al de la mañana. Dentro de los atributos de la relación se encuentran valores para identificar el viaje en particular, como ser el día, el turno y el asiento.

En la figura 1 se muestra, a modo de ejemplo, cómo a

través de la relación generada 'VIVE_CON' se identifica rápidamente qué trabajadores comparten vivienda. Esta relación no suele cambiar en el transcurso de la obra salvo determinadas circunstancias especiales donde se deba reasignar un trabajador. Si se quisiera identificar en qué vivienda están los ocho trabajadores, bastaría con visualizar en la relación el valor del atributo *housing*. Para el ejemplo citado es '2A1'.

En la figura 2 se observa, para un día en particular, algunos de los trabajadores que utilizaron determinados ómnibus ya sea en el turno de la mañana o en el de la tarde. Esto se muestra a través de la relación 'VIAJO_EN'. Aquí se puede visualizar cómo los trabajadores viajan en varios ómnibus indiferentemente, aunque la situación ideal sería que no variaran demasiado y así evitar tener que aislar varios trabajadores.

Cabe destacar que la visualización de ambas relaciones podría hacerse en conjunto y que en hora de identificar los contactos interesa ver ambas relaciones.

III-C. Aplicación del modelo

Una vez generado este modelo, existe la ventaja de tener toda la información unificada, esto implica que al aparecer un caso positivo de COVID-19 ya no deberían ser varias áreas las que atravesen el proceso de rastreo sino que, con una única persona que maneje la herramienta, se obtendrían los resultados rápidamente. A continuación se demuestra la utilidad del modelo generado realizando determinadas consultas en lenguaje *Cypher*.

Como se mencionó anteriormente, cuando un trabajador da positivo para COVID-19, además de aislarse él, se deben aislar a todas las personas que se alojan con él y a su vez,

a todos aquellos que hayan viajado en los asientos linderos en el periodo de tiempo que se deba considerar según las características del caso positivo (si presentó síntomas o no, etc).

Para ilustrar esto, se tiene el siguiente ejemplo, el trabajador *id* 102, que no presentaba síntomas, dio positivo en un hisopado de control (las empresas constructoras suelen hacerlos debido a la cantidad de casos asintomáticos existentes) realizado la mañana del martes 29 de junio. Bajo estas condiciones, se deberán aislar los trabajadores que vivan con él así como las personas que hayan viajado el lunes 28 en la mañana y en la tarde, y el martes mismo en la mañana. Se asume que para el viaje de la tarde el trabajador ya fue correspondientemente aislado y enviado en un transporte sanitario fuera de la obra.

Listado 1 Consulta Cypher para identificar contactos

```

1 MATCH (p:Persona {id:102})-[r:VIAJO_EN]
2 ->(o)<-[:VIAJO_EN]-(p2)
3 WHERE r.dia=t.dia= "28/6/2021"
4 AND r.turno = t.turno = 1
5 AND (CASE WHEN r.asiento%4 = 1
6 then r.asiento = (t.asiento - 2)
7 ELSE r.asiento = (t.asiento + 2) END
8 OR r.asiento = (t.asiento + 4)
9 OR r.asiento = (t.asiento - 4))
10 RETURN p2.nombre AS Nombre, p2.apellido
11 AS Apellido, p2.id AS Id, ("Transporte" + " " +
12 r.dia + " " + CASE WHEN r.turno = 1 THEN
13 "manana" ELSE "tarde" END) AS Contacto
14 UNION
15 MATCH (p:Persona {id:102})-[r:VIAJO_EN]
16 ->(o)<-[:VIAJO_EN]-(p2)
17 WHERE r.dia=t.dia= "28/6/2021"
18 AND r.turno = t.turno = 0
19 AND (CASE WHEN r.asiento%4 = 1
20 then r.asiento = (t.asiento - 2)
21 ELSE r.asiento = (t.asiento + 2) END
22 OR r.asiento = (t.asiento + 4)
23 OR r.asiento = (t.asiento - 4))
24 RETURN p2.nombre AS Nombre, p2.apellido
25 AS Apellido, p2.id AS Id, ("Transporte" + " " +
26 r.dia + " " + CASE WHEN r.turno = 1 THEN
27 "manana" ELSE "tarde" END) AS Contacto
28 UNION
29 MATCH (p:Persona {id:102})-[r:VIAJO_EN]
30 ->(o)<-[:VIAJO_EN]-(p2)
31 WHERE r.dia=t.dia= "29/6/2021"
32 AND r.turno = t.turno = 1
33 AND (CASE WHEN r.asiento%4 = 1
34 then r.asiento = (t.asiento - 2)
35 ELSE r.asiento = (t.asiento + 2) END
36 OR r.asiento = (t.asiento + 4)
37 OR r.asiento = (t.asiento - 4))
38 RETURN p2.nombre AS Nombre, p2.apellido
39 AS Apellido, p2.id AS Id, ("Transporte" + " " +
40 r.dia + " " + CASE WHEN r.turno = 1 THEN
41 "manana" ELSE "tarde" END) AS Contacto
42 UNION
43 MATCH (p:Persona {id:102}), (p2:Persona)
44 WHERE (p)-[:VIVE_CON]->(p2)
45 OR (p2)-[:VIVE_CON]->(p)
46 RETURN p2.nombre AS Nombre, p2.apellido
47 AS Apellido, p2.id AS Id, ("Vivienda" + " " + p2.housing)
48 AS Contacto

```

El resultado de esta consulta mostrará todos los trabajadores que deban ser aislados por haber estado en contacto con el trabajador 102, ya sea por vivir con él o por haber compartido ómnibus en los asientos de al lado el día 28/6/2021 (en cualquiera de los turnos) o el día 29/6/2021 (en el turno de la mañana).

Cuadro I: Ejemplo de resultados de la consulta

Nombre	Apellido	Id	Contacto
Sofilio	Darwin	101	Transporte 28/6/2021 mañana
Ulises	Mesa	94	Transporte 28/6/2021 tarde
Rudy	Muniz	164	Transporte 29/6/2021 mañana
Ignacio	Cuadro	196	Vivienda 5C3

En el Cuadro 1 se muestra, a modo de ejemplo, cómo se ve la salida de la consulta *Cypher*. Se seleccionaron casos de diferentes características para demostrar la utilidad y no se muestra toda la salida ya que la misma era demasiado amplia. Esta consulta tiene la ventaja de mostrar el tipo de contacto que tuvo la persona con el caso positivo, ya que, cuando se tenga que notificar al trabajador que debe aislarse es importante poder explicar la razón, es decir dónde y en qué día fue que tuvo el contacto con el trabajador enfermo.

En este ejemplo, se visualiza uno de los peores escenarios en el cuál el trabajador 102 siempre viajó al lado de personas diferentes, y no vive con ninguna de ellas, entonces se tienen un total de 16 trabajadores que se deberán aislar: los siete que viven con él más los nueve que estuvieron en los asientos linderos en esas fechas. Esta situación no es la que suele ocurrir. En la práctica las diversas cuadrillas se suelen liberar a la misma hora para intentar que los trabajadores viajen con las mismas personas.

IV. CONCLUSIONES Y TRABAJO FUTURO

Realizar un correcto y veloz seguimiento y aislamiento de los contactos de los casos positivos de COVID-19 es una forma eficaz de frenar el avance de la pandemia. Desde los gobiernos, se han buscado numerosas herramientas para lograr este objetivo, algunas con mayores resultados que otras.

El uso de la tecnología llegó para quedarse y por ello, se propone (siempre que sea posible) el uso de una base de datos de grafos para detectar de forma más rápida y eficiente los contactos de casos positivos.

Para la realidad de RYENA se ahorrarían muchos recursos realizando el seguimiento de esta manera. Siempre que se haga un rastreo manual entre numerosas planillas hay lugar a errores. Sobre todo porque es un proceso que se maneja desde áreas diferentes. Alguien que deba chequear entre varios días, con varios números de asiento y de ómnibus puede fácilmente equivocarse. Utilizando el modelo que se propone, se obtienen resultados más fiables y también, se tiene la información unificada, evitando así demoras propias características de un trabajo manual y de la coordinación entre diferentes áreas.

Como trabajo futuro, este modelo tiene potenciales mejoras. A modo de ejemplo, para este documento la consulta que se generó en *Cypher* se realizó para un caso en particular.

Con mínimos cambios se podría ajustar la misma para que considere determinados atributos como parámetros y así lograr una consulta generalizada.

Por otro lado, se debe trasladar el modelo a todas las obras que RYENA posee. Muchas veces, trabajadores de distintas obras comparten transporte, por lo que, al incluir todos los trabajadores la situación sería más parecida a la realidad.

Asimismo, se debe identificar la forma de trasladar definitivamente la información a una base de grafos, dejando atrás las planillas de Excel. Es decir se debe trasladar el modelo de RYENA hacia una base de datos no relacional. Esto no es tarea sencilla, generar un esquema nuevo de base de datos es un proceso que demanda recursos y tiempo.

El uso de bases de datos de grafos permite, además, realizar otro tipo de análisis interesante como identificar hilos epidemiológicos y determinar características de ellos en base a la edad y el sexo de los contagiados. Se podría pensar en incorporar al modelo otros atributos interesantes como el hecho de si el trabajador está vacunado o no y empezar a sacar conclusiones en base a esto.

También se podrían realizar predicciones, por ejemplo, identificando qué porcentaje de los contactos del caso positivo se contagian, y a cuántas personas contagian ellos, y así estimar cuantas bajas de trabajadores se tendrán y saber de antemano si deben contratarse suplentes para no detener el avance de la obra.

En conclusión, los beneficios que presenta el uso de una base de datos de grafos en el rastreo de contactos de casos positivos de COVID-19 son enormes. Si la vacuna es el mejor aliado en la prevención del coronavirus, el uso de tecnología es el mejor aliado en la detección eficaz de los contactos. Aún así, tanto en este trabajo, como en el análisis de datos en general, se aprecia que la utilidad práctica es fuertemente dependiente de la integridad y disponibilidad de los datos. Esto es que la información se cargue en forma oportuna, y sea correcta y completa.

REFERENCIAS

- [1] Organización Mundial de la Salud *Tuit de declaración de pandemia*, Consultado 10-07-21 <https://twitter.com/WHO/status/1237777021742338049>.
- [2] Matías Arim, Daniel Herrera, Ignacio Sanguinetti, Javier Pintos, Alicia Alemán, Héctor Romero, María Inés Fariello y Horacio Botti *Nota 1: Ciencia y Datos para el análisis de la epidemia de COVID-19*, https://guiad-covid.github.io/publication/nota1/Nota_1_Ciencia%20y%20Datos_GUIAD-Covid19.pdf.
- [3] Neo4j *Página web de Neo4j con artículos relacionados al COVID-19*, Consultado 12-07-21 <https://neo4j.com/graphs4good/covid-19/>.
- [4] Delphine Courvoisier, Paul-Olivier Dehayé, Olivier Lange y Jean-Luc Falcone *Extensive Representation of Contact Tracing with Neo4j in the Canton of Geneva*, Consultado 13-07-21 <https://neo4j.com/blog/extensive-representation-of-contact-tracing-with-neo4j-in-the-canton-of-geneva/>.
- [5] Zijun Mao, Hong Yao, Qi Zou, Weiting Zhang y Ying Dong *Digital Contact Tracing Based on a Graph Database Algorithm for Emergency Management During the COVID-19 Epidemic: Case Study*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7837510/>.

- [6] Fabrizio Scrollini, Javier Baliosian, Lorena Etcheverry y Guillermo Moncecchi *Nota 6: Aplicaciones de seguimiento de contactos en la emergencia sanitaria: aportes para el diseño y evaluación de política pública*, https://guiad-covid.github.io/publication/nota6/Nota_6_Seguimiento_de_contactos_GUIAD-Covid19.pdf.
- [7] Rik Van Bruggen *Covid-19. Contact tracing: an amazing graph problem and rabbit hole*, Consultado 13-07-21 <http://blog.bruggen.com/2020/04/covid-19-contact-tracing-amazing-graph.html>.