Diseño de Algoritmos

Trabajo Practico N° 2 - Parte II - NP e Intratabilidad

Manuel Latorre FAI-1931
manuel.latorre@est.fi.uncoma.edu.ar

Segundo cuatrimestre 2022

Facultad de Informática
UNIVERSIDAD NACIONAL DEL COMAHUE

UNIVERSIDAD NACIONAL DEL COMAHUE
1972

Índice

1. Punto 1

Una pequeña empresa de tecnología realiza trabajos sensibles para lo cual mantiene un sistema de computadora de alta seguridad con el objeto de garantizar que el mismo no se utilice con algún fin ilícito. Para ello cuenta con software que registra las direcciones IP de los accesos de todos sus usuarios a través del tiempo. Se asume que cada usuario accede al menos a una dirección IP en algún minuto dado; el software registra en un archivo log lo siguiente: Para cada usuario u y cada minuto m, un valor I(u,m) es igual a la dirección IP accedida, si la hubiera. Si no, se setea con el símbolo —, que significa que el usuario u no accedió durante el minuto m.

Los directivos de la empresa han detectado que ayer, el sistema fue utilizado para lanzar un ataque complejo sobre algunos sitios remotos. El ataque fue llevado a cabo a partir de acceder a t direcciones IP distintas en t minutos consecutivos: en el minuto 1, el ataque accedió a la dirección i_1*; en el minuto 2, accedió a la dirección* i_2 *; y así para la dirección* i_t *en el minuto t.*

¿Quién pudo haber sido el responsable de llevar adelante el ataque? La empresa chequea los archivos logs y encuentra para su sorpresa que no existe un único usuario u que accedió a cada una de las direcciones IP involucradas en el tiempo t apropiado; en otras palabras, no existe un u tal que $I(u,m) = i_m$ *para cada minuto m desde 1 hasta t.*

La pregunta es: ¿Existió una pequeña coalición de k usuarios que colectivamente podría haber llevado a cabo el ataque?

Se dice que un subconjunto S de usuarios es una coalición sospechosa si, para cada minuto m desde 1 hasta t, existe al menos un usuario $u \in S$ *para el cual* $I(u,m) = i_m$*. (En otras palabras, cada IP fue accedido en el momento apropiado por al menos un usuario de la coalición).*

El problema de Coalición Sospechosa pregunta: ¿Dado el conjunto de todos los valores $I(u,m)$*, y un número k, existe una coalición sospechosa de al menos k de tamaño?*

Enunciado

- *Indicar por qué es un problema NP*
- *Encontrar un problema conocido NP-completo para poder reducir Coalición Sospechosa.*
- *Describir un algoritmo que en tiempo polinomial, realiza la reducción y probar su validez.*

El problema de la coalición sospechosa es NP ya que si se tiene un conjunto de S usuarios, se puede comprobar que S tiene un tamaño máximo de k usuarios, y por cada minuto m de 1 a t, al menos uno de los usuarios de S accedió a la dirección IP i_m. Finalmente se puede plantear un algoritmo que verifique en tiempo polinomico $O(k \cdot t)$ la existencia de al menos un usuario $u \in S$ para el cual $I(u,m) = i_m$ (Osea que al menos cada IP fue accedido en un determinado momento por al menos un usuario) verificando asi que el problema es NP.

Para reducir el problema de Coalición Sospechosa se utilizara el problema NP-Completo Vertex Cover [1], esto es posible ya que se necesita explicar todos los t accesos sospechosos, y se tiene un limitado numero de usuarios (k) para hacer esto por lo que Vertex Cover se ajusta a las necesidades.

En Vertex Cover se necesita cubrir cada arista, teniéndose permitidos k nodos. En coalición sospechosa se deberá "cubrir" todos los accesos, y solo se tienen permitidos k usuarios. Este paralelismo sugiere fuertemente que, dada una instancia de Vertex Cover que consiste en un grafo G(V, E) de k nodos, se puede construir una instancia de Coalición Sospechosa en donde los usuarios representan los nodos de G y los accesos sospechosos representan los vertices.

Entonces a partir un grafo G de una instancia Vertex Cover de m aristas $e_1, \dots, e_m$, y $e_j = (v_j, w_j)$. Se construirá una instancia de Coalición Sospechosa de la siguiente manera. Para cada nodo G se define un usuario, y para cada arista $e_t = (v_t, w_t)$ se definirá un minuto t (de esta manera se tendrán m minutos en total). En el minuto t, los usuarios asociados con dos extremos finales de e_t acceden a una dirección IP i_t, y todos los otros usuarios no acceden a nada. Finalmente, el ataque consiste en accesos a direcciones $i_1, i_2, \ldots, i_m$ en minutos $1, 2, \ldots, m$ respectivamente.

Lo planteado entonces establece que Vertex Cover es reducible en tiempo polinómico y por lo tanto finaliza la prueba de que Coalición sospechosa es NP-Completo.

2. Punto 2

Se le ha encomendado organizar un seminario para estudiantes ingresantes que se reunirá una vez por semana durante el próximo semestre. El plan es tener la primera porción del semestre consistente de una secuencia de l conferencias invitadas por oradores externos, y la segunda porción del semestre dedicada a la realización de una secuencia de p proyectos prácticos que los estudiantes desarrollarán.

Hay n opciones para los oradores en general, y en la semana número i (para $i = 1, 2, \ldots, l$)un subconjunto de Li de estos oradores está disponible para dar la conferencia.

Por otro lado, cada proyecto requiere que los estudiantes hayan visto cierto material de base para que puedan completar el proyecto exitosamente. En particular, para cada proyecto j (para $j = 1, 2, \ldots, p$), existe un subconjunto Pj con oradores relevantes por lo que el estudiante necesita haber ido al menos una vez, a la conferencia de uno de esos oradores en el conjunto Pj para completar el proyecto. Dados estos conjuntos, ¿se puede seleccionar exactamente un orador para cada una de las l primeras semanas del seminario, de manera de elegir oradores que estén disponibles en la semana designada, y para que, para cada proyecto j, los estudiantes hayan visto al menos uno de los oradores en el conjunto de Pj relevante? Se denomina este enunciado como el Problema de la Planificación de Conferencia.

Considerar la siguiente instancia de ejemplo. Se supone que l=2, y p=3. Y que existen n=4 oradores denotados A, B, C, D. La disponibilidad de los oradores está dada por los conjuntos $L_1 = A, B, C$ y $L_2 = A, D$. Los oradores relevantes para cada proyecto están dados por los conjuntos $P_1 = B, C$, $P_2 = A, B, D$ y $P_3 = C, D$. Para esta instancia la respuesta es SÍ, ya que se puede elegir al orador B en la primera semana y al orador D en la segunda; de esta manera, para cada uno de los tres proyectos, los estudiantes habrá visto a los oradores relevantes.

Enunciado

- *Indicar por qué es un problema NP*
- *Encontrar un problema conocido NP-completo para poder reducir Coalición Sospechosa.*
- *Describir un algoritmo que en tiempo polinomial, realiza la reducción y probar su validez.*

El problema es NP porque, dado una secuencia de oradores, se puede verificar en tiempo polinómico que

- Todos los oradores están disponibles en las semanas en las cuales fueron programados
- Para cada proyecto, al menos uno o mas oradores relevantes fueron programados

Al igual que en el primer punto, para el problema de la Planificación de la Conferencia se puede realizar una reducción a partir del problema Vertex Cover [1]. Esto es debido a que se puede ver a Vertex Cover como una estructura similar de dos fases: primero se deben elegir un conjunto de k nodos del grafo de entrada, y entonces se debe verificar que para cada uno de dichos nodos se cubran todas sus aristas a las cuales están conectados.

Dada una entrada de Vertex Cover que consiste en un grafo $G = (V, E)$ y un numero k, se creara un orador z_v para cada nodo v. Se seteara $l = k$, y se define $L_1 = L_2 = \ldots =$

$L_k = \{z_v : v \in V\}$. En otras palabras, para las primeras k semanas, todos los oradores estan disponibles. Despues de esto, se creara un projecto j para cada arista $e_j = (v, w)$, con un conjunto $P_j = \{z_v, z_w\}$

Ahora, si existe un Vertex Cover S de a lo sumo k nodos, entonces se considerara el conjunto de oradores $Z_S = \{z_v : v \in S\}$. Para cada projecto P_j, al menos un orador relevante pertenecera a Z_s, donde S cubre todas las aristas de G. Es mas, se pueden programar todos los oradores en Z_S durante las primeras k semanas. De este modo se plantea una solucion factible a una instancia del problema de la Planificación de Conferencia

Por el contrario si se supone que existe una solucion factible para una instancia de Planificación de conferencia, y sea T el conjunto de todos los oradores que hablan en las primeras k semanas. Sea X el conjunto de nodos en G que corresponden a oradores en T. Para cada proyecto P_j, al menos uno de los dos oradores relevantes aparecen en T, y por lo tanto al menos un extremo de cada arista e_j esta en el conjunto X. De este modo X es un Vertex Cover con al menos k nodos. Y asi concluye la prueba de que Vertex Cover es reducible en tiempo polinómico al problema de la Planificación de Conferencia

Referencias

[1] Boštjan Brešar et al. «Minimum k-path vertex cover». En: *Discrete Applied Mathematics* 159.12 (2011), págs. 1189-1195. ISSN: 0166-218X. DOI: https://doi.org/10.1016/j.dam.2011.04.008. URL: https://www.sciencedirect.com/science/article/pii/S0166218X11001387.