

Teoría de Algoritmos II (75.30)

Trabajo Práctico 2 – 30/10/2023 - Fecha de Entrega:
27/11/2023

1

Se quiere convocar a una elección a la que se presentan 4 candidatos (A, B, C y D). Hay 3 votantes del jurado que tienen sus siguientes rankings individuales:

Jurado 1 : $B \succ C \succ D \succ A$

Jurado 2 : $C \succ D \succ A \succ B$

Jurado 3 : $D \succ A \succ B \succ C$

Eliminación iterativa

En cada iteración se elimina el candidato con menor cantidad de votos.

Iteración 1

A : 0

B : 1

C : 1

D : 1

→ Se elimina A.

Jurado 1 : $B \succ C \succ D$

Jurado 2 : $C \succ D \succ B$

Jurado 3 : $D \succ B \succ C$

Iteración 2

→ Todos los candidatos tienen la misma cantidad de votos. No se elimina ningún candidato.

Borda Rule

Se asigna un puntaje a cada candidato según la posición en la que se encuentra en el ranking de cada votante. Asumo que son 3 puntos para el primero, 2 para el segundo, 1 para el tercero y 0 para el cuarto.

$$A = 0 + 1 + 2 = 3$$

$$B = 3 + 0 + 1 = 4$$

$$C = 2 + 3 + 0 = 5$$

$$D = 1 + 2 + 3 = 6$$

$\rightarrow \text{Gana } D.$

Buscando que gane A -> Eliminaciones sucesivas

Se eliminan los candidatos en rondas de a pares de candidatos. Asume que si un candidato gana en un enfrentamiento contra otro candidato B, ganaría también contra todos los candidatos que ganó el candidato B.

Orden de rondas: D C B A

Ronda 1: D contra C

Jurado 1 : Gana C

Jurado 2 : Gana C

Jurado 3 : Gana D

$\rightarrow \text{Gana } C.$

Ronda 2: C contra B

Jurado 1 : Gana B

Jurado 2 : Gana C

Jurado 3 : Gana B

$\rightarrow \text{Gana } B.$

Ronda 3: B contra A

Jurado 1 : Gana B

Jurado 2 : Gana A

Jurado 3 : Gana A

$\rightarrow \text{Gana } A.$

Al ganar A no se está cumpliendo la propiedad de ser Pareto-Eficiente ya que el orden relativo $D \succ A$, presente en las preferencias de cada uno de los jurados, no se respeta en el resultado.

2

Dado el modelo de cascada de información visto en clase. A partir de ahora llamamos T a la nueva tecnología. Vamos a analizar como actúan los individuos de la red ante una nueva tecnología. Cada individuo puede aceptar o rechazar T . Si acepta, recibe una ganancia g aleatoria que puede ser positiva o negativa. Si rechaza, recibe una ganancia $g = 0$.

Si $T = \text{"Mala"}$. Entonces sabemos que el promedio de ganancias entre los que aceptaron T será **negativo**. De esta forma, que el primer individuo que acepte la tecnología, seguramente tendrá una ganancia negativa. Luego, los siguientes individuos podrán aceptar T o rechazarla, conociendo las ganancias de los demás y sabiendo el promedio de las mismas. Si el promedio es negativo, entonces es más probable que un individuo a la hora de elegir, tenga preferencia por rechazar la tecnología en vez de aceptarla, lo cual "frena" la potencial formación de una cascada para que persista T .

Podemos ver por ejemplo, que si los primeros individuos que acepten tienen malas experiencias ($g < 0$), cada vez es menos probable que los siguientes acepten la tecnología, pudiendo llegar al punto de formar una cascada de Rechazo, si la cantidad de malas experiencias (o la importancia de las mismas) es suficiente.

Esto se corresponde con lo que experimentamos en la realidad. Si un solo amigo mío probó un juego nuevo y no le gustó, quizás me desanime a probarlo, pero aún es posible que le de una oportunidad. En cambio si muchos/todos mis amigos tuvieron malas experiencias, probablemente ni siquiera pruebe el juego.

Por otro lado, si $T = \text{"Buena"}$. Entonces cuando un individuo deba elegir, va a notar que en promedio, las ganancias/experiencias de los que adoptaron T antes son más positivas que negativas, lo que nos incentiva a Aceptar esta nueva tecnología. Podemos pensar también, que la primer persona que acepte T tendrá una ganancia positiva, por lo que la persona siguiente probablemente optará por aceptarla también. Y aunque algunos individuos hayan tenido malas experiencias (ganancias negativas), habrá mayor tendencia a adoptar la nueva tecnología por ser "Buena" que a rechazarla.

De esta forma, vemos que no debería/es poco probable que surja una cascada de rechazo si la tecnología es Buena.

3

Considerando la siguiente red, en donde cada nodo tiene inicialmente un comportamiento B:

Suponemos a E y F como *early adopters* del comportamiento A (marcado en violeta). Y que el resto de nodos adoptará este comportamiento sólo si al menos la mitad de sus vecinos lo adoptan.

A continuación se muestra la secuencia en la que adoptan el comportamiento A los nodos restantes.

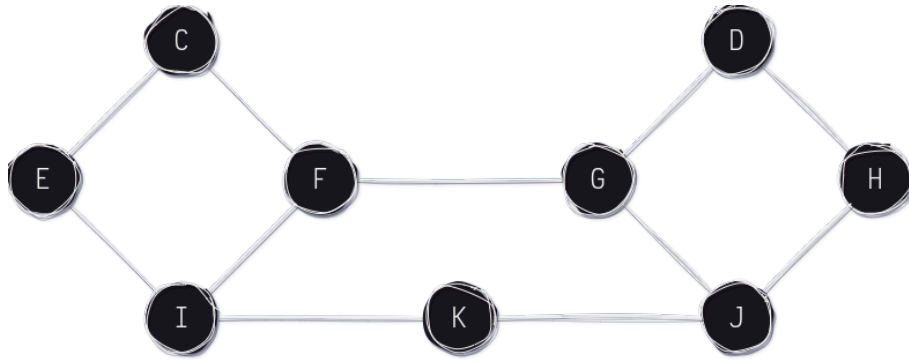


Figure 1: Grafo original

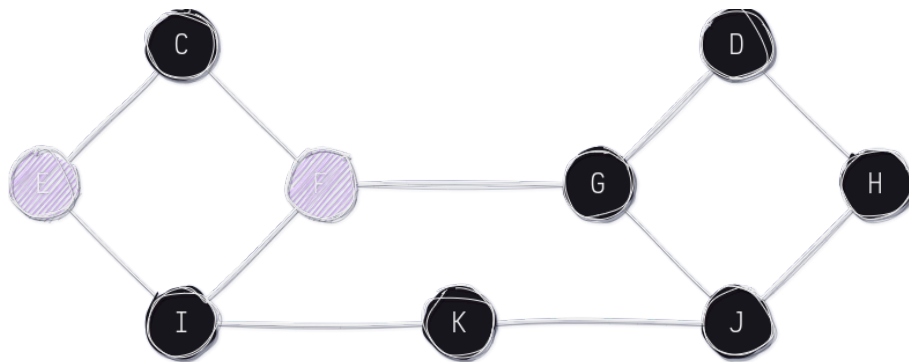


Figure 2: Grafo en etapa 1

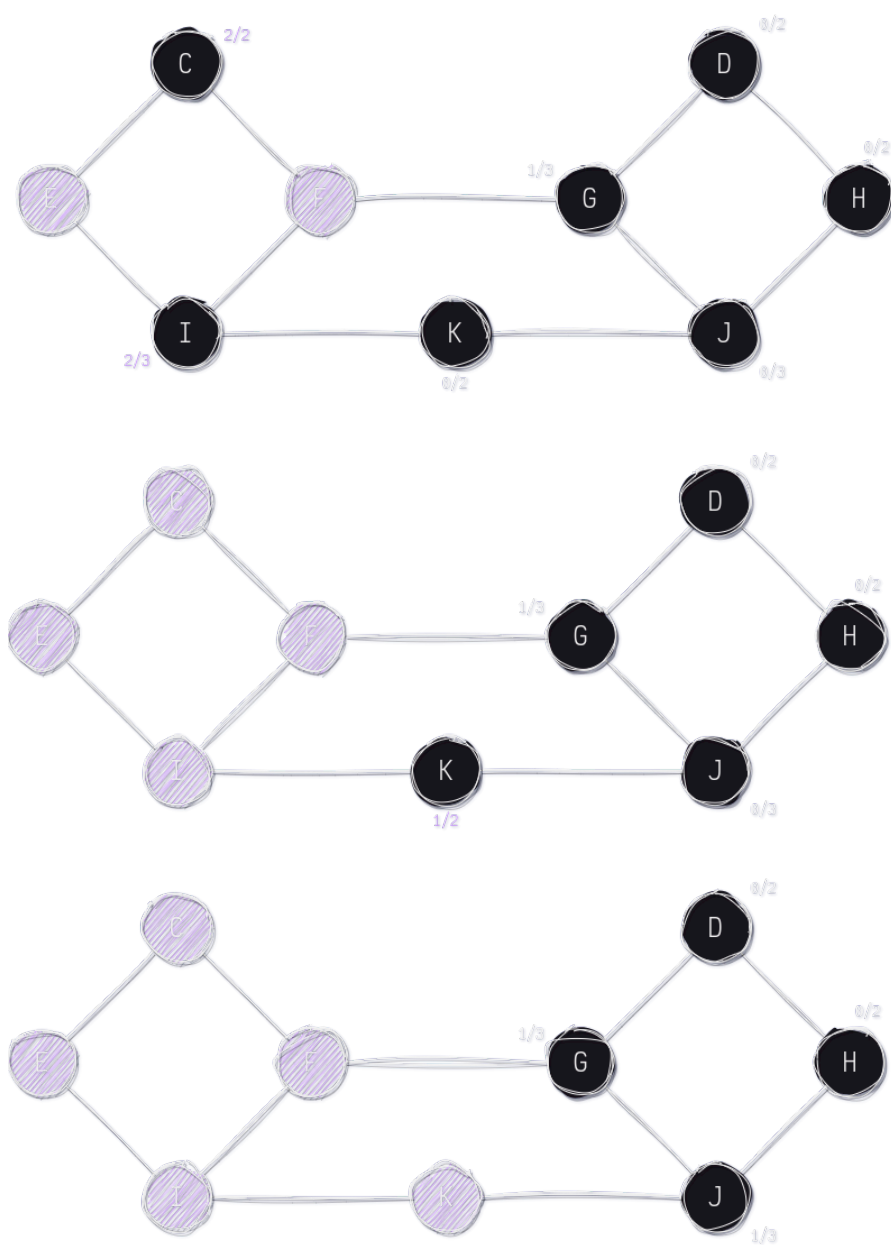


Figure 3: Grafo en etapas siguientes

Finalmente, se llega a que ni G ni J logran adoptar este nuevo comportamiento, dado que ambos tienen 3 vecinos de los cuales sólo 1 adoptó el comportamiento A. De esta forma, tampoco será posible que D o H adopten el comportamiento A, ya que ambos tienen 2 vecinos que no lo adoptaron.

El comportamiento de A no se propaga por toda la red dado que la situación inicial presentaba dos clusters, en donde sólo el cluster izquierdo poseía *early adopters*. De esta forma, la información sobre la existencia del comportamiento A llegó a algunos nodos del cluster derecho mediante puentes/lazos débiles, pero sin alcanzar la cantidad mínima de vecinos que debían adoptar el comportamiento para que estos nodos también lo adopten.

De esta forma, para que el comportamiento A se expanda por el resto de la red, otro *early adopter* debería estar presente en el cluster derecho para que el comportamiento A se propague por toda la red. Más aún, el *early adopter* podría ser cualquier nodo del cluster derecho.

4

Los grafos que cumplen con la Ley de Potencias (como los del modelo de mundo pequeño) tienen diámetro bajo y coeficiente de Clustering alto. Lo que indica que a un virus le tomaría pocos “saltos” o iteraciones llegar a todos los nodos del grafo, siendo más probable que llegue a infectar a todos los nodos rápidamente y de esta forma que crezca la cantidad de nodos infectados en simultáneo.

Por otro lado, en un grafo aleatorio de ER, el coeficiente de Clustering esperado es aproximadamente:

$$E[C] \approx \frac{\hat{k}}{n}.$$

De modo que el mismo depende del grado promedio y la cantidad de nodos, pudiendo ser más alto o bajo dependiendo del valor de probabilidad con el que se generó el grafo aleatorio. Por lo tanto, no tenemos la certeza de que G_1 tenga un coeficiente de Clustering alto o bajo.

De esta forma, si tomáramos un nodo aleatorio, podemos concluir que es más probable que ocurra una epidemia en G_2 , ya que desde cualquier nodo del mismo, sería “fácil” llegar a todos los demás.

Por otro lado, si la propagación del virus comenzara en el vértice de mayor grado la idea que se nos ocurre es que en los grafos que cumplen ley de potencias, al tener la característica de la “Heavy Tail”, siempre tenemos probabilidad de que el nodo con mayor grado del grafo sea efectivamente un nodo con un grado muy alto, por lo tanto seguiría siendo más probable que la epidemia ocurra en este tipo de grafos con más facilidad que en un grafo aleatorio de Erdős-Rényi.

En cuanto a las comunidades, podemos decir con un alto grado de certeza que las mismas benefician a la expansión del virus, ya que una vez infectado uno o algunos pocos de los miembros de la comunidad, existen múltiples “camino” que el virus puede tomar para llegar a todos los demás miembros de la misma,

facilitando la ocurrencia de una epidemia dentro de la comunidad. Otra forma de verlos es que dentro de las comunidades, es altamente probable que se cumpla la formación de triángulos dentro de la comunidad (Triadic Closure), por lo tanto los nodos infectados dentro de una gran comunidad tendrán probabilidad mayor de tener más aristas que los nodos que no pertenecen a una, facilitando la expansión del virus.

Cabe destacar, que los valores de β y δ del modelo **SIR** no deberían influir en estos resultados, ya ambos grafos deben ser sometidos a las mismas instancias del virus.

Finalmente, en cuanto a las simulaciones para contrastar y analizar los resultados, implementamos un modelo para realizar las simulaciones de propagación de un virus SIR en grafos de networkx en el siguiente collab. Sin embargo, luego de varias ejecuciones con distintos parámetros, no logramos visualizar diferencias notables entre un modelo y otro.

5

Teniendo un set de datos de reviews de Amazon, se analizó el comportamiento de los usuarios y se calcularon las métricas de justicia (de un usuario), valor (de un producto) y de confianza (de una review a un producto de determinado usuario) usando el algoritmo REV2.

Como resultado se obtuvo que los usuarios maliciosos con al menos 5 reviews son:
- 'A2G4GBZBU0191J' (fairness = 0.21) - 'A3ECP9FPY96ST2' (fairness = 0.22)
- 'A135DQ1W2SWG7R' (fairness = 0.22) - 'A3VQ1D0OILACKN' (fairness = 0.23) - 'A3VNX3A8V805AL' (fairness = 0.24) - 'A1ILBK17KOOOYN' (fairness = 0.25)

Respecto a los extremadamente justos (nodos con al menos 0.9 de fairness), se obtuvo que el 0.59% de los usuarios son extremadamente justos (375 usuarios, siendo un total de 63175 usuarios en la red).

> Tanto la implementación del algoritmo, como los resultados obtenidos, se encuentran en el notebook REV2 - Amazon reviews.