



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ADAPTIVE RUMOR SPREADING

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN
GESTIÓN DE OPERACIONES
MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL INDUSTRIAL

ALBERTO ABEL VERA AZÓCAR

PROFESOR GUÍA:
JOSÉ RAFAEL CORREA HAEUSSLER

MIEMBROS DE LA COMISIÓN:
MARCOS ABRAHAM KIWI KRAUSKOPF
NEIL KELVIN OLVER
DENIS ROLAND SAURÉ VALENZUELA

SANTIAGO DE CHILE
2015

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE: Ingeniero Civil Industrial y grado de
Magister en Gestión de Operaciones
POR: Alberto Abel Vera Azócar
FECHA: 25/06/2015
PROF. GUÍA: DR. José Rafael Correa

ADAPTIVE RUMOR SPREADING

El esparcimiento de rumores es un modelo intuitivo para la difusión de información en una red social. Una entidad que controla la red, por ejemplo el proveedor del servicio, desea acelerar el proceso de esparcimiento del rumor, de forma tal que se maximice la cantidad de información entregada. Este problema, definido a grandes rasgos, ha sido objeto de múltiples investigaciones en el último tiempo, entre otros como marketing viral y maximización de influencia. Un enfoque natural y ausente en los estudios previos es la adaptividad. En este trabajo se abordan las siguientes preguntas: ¿cómo el controlador puede usar la información del estado de la red para acelerar el proceso de rumor? y ¿cuánto beneficio se obtiene de tal conocimiento?

Un concepto novedoso es la comunicación oportunista en redes; cada agente de la red social comparte información (noticias, actualización de *software*, etc.) con otros usuarios al momento de estar momentáneamente en rango (vía *wi-fi*, *bluetooth*, etc.), de esta forma se evita la saturación de la infraestructura que soporta la red. Con esta motivación se estudia un modelo a tiempo continuo, donde cada par de nodos se comunica de acuerdo a un proceso de Poisson de cierta tasa y el rumor se transmite siempre que alguno estuviera informado. Las anteriores comunicaciones no tienen costo para el controlador, pero si éste lo desea puede informar a cualquier nodo pagando un costo unitario por ello. En vez de la usual restricción de presupuesto se fija un *deadline*, en tal tiempo todos los nodos deben estar informados, debiendo pagar el controlador un costo unitario por cada nodo que no haya obtenido el rumor antes del *deadline*.

Una estrategia no-adaptativa puede informar sólo al comienzo del periodo y cuando se cumple el *deadline*, pagando por todos aquellos nodos que no se comunicaron nunca con otro nodo informado. Una estrategia adaptativa puede intervenir la red en cualquier instante, usando toda la información disponible hasta ese entonces, en particular sabiendo cuales nodos tienen el rumor en cada momento.

El resultado principal de este trabajo es que en el caso homogéneo, donde cada par de nodos se encuentra con la misma tasa, el beneficio de la adaptividad está acotado por una constante. La demostración requiere un entendimiento profundo del proceso estocástico que domina el sistema, que se cree ya una contribución interesante. Adicionalmente, se presenta una extensión natural del caso homogéneo, donde el controlador está interesado sólo en un conjunto de nodos y no en toda la red social, se demuestra que en este escenario el beneficio de la adaptividad también está acotado por una constante. Finalmente, se muestra que, sin el supuesto de homogeneidad, el beneficio de la adaptividad puede crecer de forma no acotada.

A mi padre.

Agradecimientos

Me gustaría agradecer a mis amigos, con los que atravesé tantas horas de estudio y tareas. Gracias a mi familia, por estar ahí para mí. Gracias a mi profesor, José Correa, quien me brindó consejo y amistad. Gracias a Marcos Kiwi y Neil Olver por su paciencia y enseñanzas.

Tabla de contenido

1. Introducción	1
1.1. Problemas en redes sociales	1
1.2. Descripción del problema	3
1.3. Caso homogéneo	4
1.4. Organización de la tesis	5
2. Resumen en español	6
2.1. Discusión bibliográfica	6
2.2. Brecha de adaptividad bajo heterogeneidad	8
2.3. Resultados preliminares de No-Adaptativo	9
2.4. Caracterización de Adaptativo	12
2.4.1. Ejemplo introductorio	13
2.4.2. Caso general	14
2.5. Cálculo de la estrategia adaptativa	17
2.6. Brecha de adaptividad para τ pequeño	20
2.7. Acotando la brecha de adaptividad	22
2.8. Generalizaciones	26
2.9. Caso discreto	29
3. Conclusión	31
3.1. Trabajos futuros	31
Bibliografía	32
A. Adaptive Rumor Spreading	36
A.1. Introduction	38
A.2. Preliminaries	42
A.3. Estimates on the evolution of the non-adaptive process	47
A.4. Additive Gap for Small τ	49
A.5. Bounding the adaptivity gap	53
A.6. Final comments	60

Índice de figuras

2.1. Algoritmo adaptativo en un árbol	9
2.2. Una trayectoria de \tilde{U} y su transformación. Una activación ocurrió en tiempo t' y un empuje en tiempo \tilde{t}_{k_N+1}	25
A.1. A sample evolution of \tilde{U} , and the corresponding evolution of the transformed process H . A random activation occurred at time t' , and a push occurred at time \tilde{t}_{k_N+1}	55

Capítulo 1

Introducción

Desde el surgimiento de las redes sociales y los primeros modelos formales que intentan cuantificar fenómenos en este contexto, el estudio de la difusión de rumores ha sido una pregunta básica. Un rumor puede representar, además de información, contenido como actualización de una aplicación móvil, adopción de una nueva tecnología y por supuesto conocimiento sobre algún producto o marca. Es natural además pensar que el rumor debe conocerse antes de un cierto tiempo, por ejemplo tratándose de información sobre rutas y congestión, oferta de un producto por tiempo limitado o bien si un cambio tecnológico causaría incompatibilidades con usuarios que no tengan la última actualización de *software*, cf. [Han *et al.* \[2010\]](#); [Whitbeck & Amorim \[2011\]](#); [Keller *et al.* \[2012\]](#).

El tráfico de información se hace cada vez más intenso, pero las limitaciones tecnológicas (3G, 4G) parecen no estar siguiendo el paso a la magnitud de contenido que los usuarios desean transmitir. Supongamos que dos usuarios desean descargar un video, como es lento y costoso usar internet móvil, una solución justa es que cada uno descargue parte del contenido y luego se intercambien vía *bluetooth* (u otro similar) las mitades, así cada uno tendrá el video completo. Ahora supondremos que hay muchos usuarios participando de una red social, cuando dos teléfonos estén en cercanía (por ejemplo, caminando en la calle) intercambian información de forma rápida y gratuita.

El contexto anterior da origen al concepto de una red social oportunista, cf. [Pelusi *et al.* \[2006\]](#). Hay muchas preguntas por responder: ¿cómo se forma esta red?, ¿cuáles son sus propiedades?, ¿puede esparcir un contenido (rumor) de forma rápida? Veremos a continuación algunos trabajos que intentan responder estas preguntas, pero dejamos el contexto bibliográfico de rumores para el próximo capítulo, pues naturalmente nos extenderemos más sobre este tópico.

1.1. Problemas en redes sociales

El ejemplo prominente de redes sociales es la web. En el trabajo de [Broder *et al.* \[2000\]](#) se trata a la *World Wide Web* como un grafo dirigido, donde los nodos son las páginas y

los arcos son *hyperlinks* dirigidos. Entre otras cosas se muestra que este grafo tiene una sola gran componente conexas, es decir, la mayoría de las páginas están interconectadas y sólo un grupo muy reducido de ellas no se conecta con esta componente. El grafo obtenido por [Broder et al. \[2000\]](#) contiene 203 millones de nodos y 1500 millones de arcos, de éstos un 91 % estaban conectados a la componente gigante. Otras medidas son de interés, por ejemplo el camino promedio entre pares de nodos es de 16 considerando dirección y 7 con arcos no dirigidos. Ésto es varios órdenes de magnitud menor al número de nodos, lo que motivará a encontrar cotas logarítmicas en el diámetro del grafo. Además, el grado de las páginas parece seguir una ley de potencia, una distribución que será también un objetivo deseable al modelar redes sociales.

La estructura de una red social está caracterizada por tener una gran cantidad de usuarios. Además, por su dinamismo, se ha utilizado el enfoque de estructuras aleatorias para poder estudiar los fenómenos. En otras palabras, se trata de encontrar una distribución que genere redes parecidas a las reales. Por ejemplo, el modelo $G(n, p)$ propuesto por [Erdős & Rényi \[1960\]](#), consta de n nodos y cada par se conecta independientemente con probabilidad p . Este modelo es probablemente el más sencillo, pero falla radicalmente en modelar una red social, pues la distribución de grados resulta binomial, no de potencia.

Un modelo bastante difundido es el introducido por [Barabasi & Albert \[1999\]](#), referido en la literatura como *preferential attachment*. Los nodos no se conectan de igual forma, sino que llegan de manera secuencial y se conectan según una distribución que prefiere a los nodos con alto grado. Se demuestra que este modelo exhibe ley de potencia, además de otras propiedades deseables, como por ejemplo diámetro sublogarítmico.

Supongamos que tenemos un modelo que representa bien la estructura de una red social. Es de interés estudiar las llamadas comunidades. Una comunidad se define como una porción densa del grafo, es decir, un subconjunto de nodos que tienen muchos arcos entre ellos. Esto representa por ejemplo a una persona que tiene un grupo de amigos donde todos se relacionan entre ellos, asimismo una persona puede pertenecer a varias comunidades (amigos del colegio, de la universidad, trabajo). En el artículo de [Newman \[2006\]](#) se pueden encontrar las definiciones formales, además de un análisis de comunidades utilizando la estructura de la matriz de incidencia. Por otra parte, [Gallo et al. \[1989\]](#) tratan el problema utilizando técnicas combinatoriales y obtienen algoritmos rápidos para identificar comunidades.

Un problema relevante en marketing es el de maximización de influencia. Supongamos que existe una nueva tecnología, como una aplicación móvil o medicamentos, donde cada participante de la red social es un potencial usuario. Tomamos un grafo dirigido, donde un arco (u, v) representa que u influye a v . Hay usuarios que tienen más influencia, por ejemplo un médico respetado o un conductor de un programa de tecnología. Si los usuarios recomiendan el producto, lo que se conoce en marketing como *word-of-mouth*, entonces una estrategia razonable es dar muestras gratis del producto a personas influyentes. Una pregunta fundamental, expresada por [Domingos & Richardson \[2001\]](#), es elegir el mejor conjunto de nodos influyentes, dado que hay un presupuesto limitado para regalar productos. En su trabajo proponen un modelo probabilista y dan heurísticas que buscan resolver el problema. [Kempe et al. \[2003\]](#) son los primeros en dar algoritmos con garantías de aproximación en dos modelos de influencia, además muestran que el problema es NP-duro.

1.2. Descripción del problema

Una red social esparce rumores de forma oportunista, esto es, cada par de miembros (de ahora en adelante, nodos) intercambia información cuando se encuentra. El tiempo entre encuentros es aleatorio, pero con distribución conocida. Existe un controlador, por ejemplo la compañía que presta el servicio de internet móvil, que desea entregar un rumor a toda la red. Se dirá que un nodo está *activo* si tiene el rumor, de otra forma se dirá *inactivo*. Por entregar el rumor de forma directa a un nodo, el controlador debe pagar un costo unitario, pero es gratis si entre nodos se comparte el rumor.

El controlador busca que toda la red esté activa antes de cierto tiempo, además busca lograr esto minimizando el costo. Una forma de activar la red es entregar el rumor a ciertos nodos al comienzo, luego dejar que se transmita el rumor y cuando se cumpla el tiempo entregar el rumor a los inactivos (pagando por cada uno). Otra forma es entregar rumores gradualmente (de ser necesario), utilizando la información del proceso, es decir, para cada tiempo decidir si entregar rumores conociendo a los activos. ¿Es cierto que monitorizando la evolución del proceso de rumor se puede abaratar costo?, de ser así ¿qué tan beneficioso es utilizar esta información?

Estudiamos un conjunto V de nodos, que se consideran móviles, la situación se modela imponiendo que $u, v \in V$ se encuentran según un proceso de Poisson de tasa $\lambda_{u,v}$. Como es usual, cuando un par de nodos se encuentra comparten toda la información, entonces basta un nodo activo para que después del encuentro ambos estén activos.

El controlador del servicio desea que todos los nodos estén activos en un tiempo $\tau > 0$ dado (al que se referirá como *deadline*), pagando por esto lo menos posible. El controlador puede activar directamente a cualquier nodo pagando un costo unitario. El conjunto de nodos activos en tiempo $t \in [0, \tau]$ se denotará por $S(t) \subseteq V$. Inicialmente, el controlador activa en tiempo $t = 0$ a un conjunto $S(0) \subseteq V$. En tiempos posteriores los nodos se activan por una de las dos formas siguientes:

- **Comunicación oportunista:** si dos nodos $u, v \in V$ se encuentran en tiempo t y uno de ellos estaba activo, i.e. $u \in S(t^-)$ o bien $v \in S(t^-)$, entonces para todo $t' \geq t$ se tendrá $u \in S(t')$ y $v \in S(t')$. El controlador no paga por estas activaciones.
- **Empujes:** en cada tiempo $t \in [0, \tau]$ el controlador tiene toda la información $(S(t'))_{0 \leq t' < t}$ y en base a esto decide si activar o no algún conjunto a elección. En tiempo t se elige un conjunto $C(t) \subseteq V \setminus S(t^-)$, potencialmente vacío, tal que se activa a todos aquellos nodos, luego $S(t) = S(t^-) \cup C(t)$ y el controlador debe pagar $|C(t)|$.

El problema no-adaptativo es cuando el controlador no puede empujar salvo en $t = 0$ o bien $t = \tau$, en este caso se activa a un conjunto inicial y al final del periodo se paga por los inactivos, es decir, se resuelve:

$$\min_{S(0) \subseteq V} \{|S(0)| + n - \mathbb{E}(|S(\tau)|)\}.$$

Al esquema óptimo, que resuelve el problema anterior, se le dirá *No-Adaptativo* y su costo se denota $\text{costo}_{NA}(\tau)$. Por otro lado, el problema adaptativo busca la mejor estrategia en

base a la información de $S(t)$. Para una trayectoria fija el controlador paga

$$\left| \bigcup_{0 \leq t \leq \tau} C(t) \right| + n - |S(\tau)|.$$

En general, el controlador minimiza la esperanza del costo anterior sobre todas las posibles estrategias. Al esquema que logra el óptimo lo llamamos **Adaptativo** y su costo se denota $\text{costo}_A(\tau)$, obviamente $\text{costo}_A(\tau) \leq \text{costo}_{NA}(\tau)$. Se define la **brecha de adaptividad** como $\rho(\tau) := \frac{\text{costo}_{NA}(\tau)}{\text{costo}_A(\tau)}$, no explicitamos la dependencia en n y $(\lambda_{u,v})_{u,v}$ por razones prácticas. Una pregunta natural es determinar qué tan grande es $\rho(\tau)$. Es de interés dar cotas para la brecha de adaptividad independientes de τ , este es el principal objetivo de este trabajo. Por ejemplo si $\rho(\tau)$ resulta cercano a 1 para rangos relevantes de τ (y de los otros parámetros), entonces no hay justificación en implementar una estrategia que haga uso de la información de la evolución del proceso.

1.3. Caso homogéneo

En la mayor parte del trabajo nos concentraremos en el caso homogéneo, i.e. $\lambda_{u,v} = \frac{\lambda}{n}$ para todo $u, v \in V$. Asumimos sin pérdida de generalidad (reescalando τ de ser necesario) que $\lambda = 1$. Por razones de simetría, es claro que en este contexto al controlador le basta con la información del número de nodos activos en cada momento (y no de la especificación del conjunto), además basta con decidir el número de empujes en cada tiempo.

Se define el proceso $\{K(t)\}_{0 \leq t \leq \tau}$ como el número de nodos activos en tiempo t bajo un esquema no-adaptativo, en otras palabras $K(t) = |S(t)|$. Definiendo $\lambda_k := \lambda \frac{k(n-k)}{n}$, el tiempo entre la k -ésima y la $(k+1)$ -ésima activación se distribuye $\text{Exp}(\lambda_k)$. Si T_i es el primer momento en que hay i nodos activos, por definición $T_{k_0} = 0$ y con ayuda de las variables aleatorias $X_k \sim \text{Exp}(\lambda_k)$ obtenemos las relaciones

$$T_{k+1} = T_k + X_k = \sum_{i=K(0)}^k X_i, \quad \forall k \geq K(0),$$

$$K(t) = \text{máx}\{k \in [n] : T_k \leq t\}.$$

Finalmente, enunciamos el principal resultado del trabajo: existe una cota constante para $\rho(\tau)$ en el caso homogéneo.

Teorema 1.1 *En el caso homogéneo, independiente del deadline y tamaño de la red, la brecha de adaptividad está acotada por una constante pequeña.*

Es interesante que la brecha de adaptividad acotada no se obtiene sin la hipótesis de homogeneidad, en particular se tiene el siguiente resultado que complementa al teorema.

Proposición 1.2 *Existe una familia infinita de instancias heterogéneas donde la brecha de adaptividad crece de manera arbitraria con el tamaño de la red, incluso si las tasas $\lambda_{u,v}$ se toman sólo entre dos valores posibles.*

El problema en estudio es equivalente, desde la perspectiva de optimización, al de maximización de influencia. En el contexto de marketing, se piensa que el controlador puede regalar productos. Esto le trae beneficio cero, pero aquellos nodos con productos influyen a sus amigos para que compren y esto trae beneficio unitario. El controlador maximiza el número esperado de productos no regalados, es decir, resuelve

$$\max_{k \in [n]} \{\mathbb{E}(K(\tau) - k | K(0) = k)\}.$$

Se define la brecha de adaptividad para este problema de forma análoga, pero esta vez en términos de beneficios. Demostramos un resultado más fuerte.

Proposición 1.3 *En el caso homogéneo, independiente del deadline y tamaño de la red, el problema de maximización de influencia tiene una brecha de adaptividad acotada por $1 + o(1)$.*

1.4. Organización de la tesis

Las demostraciones de los principales resultados de este trabajo pueden encontrarse en detalle en el apéndice, donde se adjunta un artículo científico. Por otro lado, el enfoque del segundo capítulo es generar intuición sobre el problema. Se presentan las demostraciones que amplían la comprensión de los resultados, omitiendo las pruebas muy técnicas.

En cuatro secciones se presentan resultados que no están en el artículo científico. En las secciones 2.4 y 2.5 se muestra un enfoque analítico para caracterizar la estrategia adaptativa óptima. En la sección 2.8 se presenta la demostración de una generalización del problema en estudio. Finalmente, en la sección 2.9 introducimos un modelo análogo, pero a tiempo discreto, que permite ampliar la discusión entre los modelos sincrónicos y asíncronos.

Capítulo 2

Resumen en español

2.1. Discusión bibliográfica

La estructura de red social se representa por un grafo $G = (V, E)$, donde los nodos son los usuarios y los arcos representan la posibilidad de encuentro, son de interés los casos con $n := |V|$ muy grande. En particular, si no existe un arco que conecte a $u, v \in V$, es porque el par u, v no se puede comunicar (no son amigos). El grafo puede ser dirigido, pues a veces las relaciones sociales no son simétricas, por ejemplo en problemas de influencia los agentes siguen a algún experto y no al revés. Además, cada arco tiene un peso que simboliza la frecuencia de contacto, cuando todos estos pesos son iguales se dice que el modelo es homogéneo.

Hay muchas líneas de investigación en el área, revisamos los resultados en base a dos criterios; el proceso de rumor puede ser a tiempo continuo o discreto, y el grafo puede ser conocido o aleatorio. Hay siempre un proceso estocástico subyacente que domina el sistema, el caso más estudiado es a tiempo discreto, también conocido en la literatura como *synchronous model*. En cada etapa de este modelo los nodos buscan comunicarse, usualmente sólo un nodo está activo al comienzo y se transmite la información según alguno de los protocolos:

- **Push**: al comienzo de cada ronda todos los nodos activos eligen, según alguna distribución, otro nodo para comunicarle el rumor.
- **Pull**: al comienzo de cada ronda todos los nodos inactivos eligen, según alguna distribución, otro nodo para preguntarle por el rumor.
- **Push-Pull**: se combinan las dos anteriores, usualmente se define que en cada ronda todos los nodos eligen, según alguna distribución, otro nodo para comunicarse e intercambiar toda la información que tengan.

Una pregunta básica es, partiendo de sólo un nodo activo, ¿en cuántas rondas (en esperanza o bien con alta probabilidad) toda la red social estará activa? Es fácil ver que en el modelo *push* se necesitan al menos $\log_2 n$ rondas, pues en cada una se duplica el número de activos en el mejor caso. Por otro lado, tal cota no es cierta en el modelo *pull*, donde en el mejor caso basta con una ronda para informar a toda la red (e.g. una estrella). Uno de los principales resultados, por [Demers et al. \[1987\]](#), es que si cada par de nodo se comunica con probabilidad

constante, los tres modelos activan a toda la red en un número esperado de $\Theta(\log n)$ etapas (acá $\log n$ denota al logaritmo en base natural). Se enfatiza la rapidez (comparada con el número de entidades en la red) en que el rumor se esparce, sobre todo porque en los modelos más estudiados de redes sociales el diámetro es justamente $\Theta(\log n)$.

Se han buscado propiedades de la red social que aseguren el esparcimiento rápido de rumores. Giakkoupis [2011b] demuestra que, con alta probabilidad, toda la red se activa en $O(\phi^{-1} \log n)$ rondas usando el protocolo *push-pull*, donde ϕ denota la conductancia del grafo, que es una medida estrechamente relacionada con las propiedades de expansión. Más aún, se muestra que la cota es ajustada si la conductancia no es muy pequeña.

Se pueden generalizar los resultados anteriores a grafos aleatorios, es de particular interés el modelo *preferential attachment* introducido por Barabasi & Albert [1999]. Con esta estructura aleatoria, es interesante que los protocolos *push* y *pull* por si solos fallan en entregar el rumor en tiempo logarítmico, mas aún lo hacen en tiempo polinomial en n , pero en el modelo *push-pull* se logra informar a la red en tiempo $O(\log^2 n)$ con alta probabilidad, cf. Chierichetti *et al.* [2011]. El resultado anterior fue luego mejorado por Doerr *et al.* [2011], donde se prueba que *push-pull* logra informar a todo el grafo en tiempo $\Theta(\log n)$, además, si en cada ronda los nodos se comunican con cualquiera excepto con el que se comunicaron en la ronda previa, entonces se muestra que el número de rondas es $\Theta(\frac{\log n}{\log \log n})$, lo que corresponde al diámetro del grafo. Otros modelos de grafos aleatorios muestran el mismo comportamiento, por ejemplo Fountoulakis & Panagiotou [2013] estudian un grafo obtenido al azar de entre todos los grafos d -regulares. Para $d \geq 3$ se obtiene que el protocolo *push* entrega el rumor en $\Theta(\log n)$ rondas con alta probabilidad.

Otro modelo de rumor menos estudiado es a tiempo continuo, también conocido en la literatura como *asynchrhonous model*. En este contexto, cada par de nodos $u, v \in V$ se comunica según un proceso de Poisson de tasa $\lambda_{u,v}$, es decir, el tiempo entre comunicaciones se distribuye $\text{Exp}(\lambda_{u,v})$. En cada comunicación, el par de nodos intercambia información según algún protocolo, generalmente *push-pull*. Aceptamos $\lambda_{u,v} = 0$ o $\lambda_{u,v} = \infty$ con las naturales extensiones; imposibilidad de contacto en el primer caso y encuentro inmediato y permanente en el segundo.

Mostramos, para generar intuición, algunos resultados en una red homogénea, donde todas las tasas son iguales y normalizadas, es decir, $\lambda_{u,v} = \frac{\lambda}{n}$ para todo $u, v \in V$. Una interpretación de la normalización es que cada nodo busca comunicarse después de un tiempo distribuido como $\text{Exp}(\lambda)$, luego elige a un nodo al azar (de entre todos los n disponibles) para intercambiar información.

El modelo bajo tiempo continuo es más realista, que el modelo a tiempo discreto, en el contexto de redes sociales, pues los integrantes no se ponen de acuerdo para interactuar en rondas sincronizadas. A su vez, el supuesto de tasas idénticas para todos los nodos debería capturar bien un escenario clave; el controlador puede conocer cuales nodos están activos en cada instante, pero obtener el detalle de las interacciones de cada nodo, de forma de estimar los parámetros $\lambda_{u,v}$, sugiere un conocimiento que podría no existir o bien ir en contra de las normas de privacidad.

El siguiente resultado muestra la propiedad básica del modelo.

Proposición 2.1 *En el caso homogéneo, comenzando con un nodo activo, si T_n es el tiempo en que toda la red está activa, entonces $\mathbb{E}(T_n) \sim \frac{2}{\lambda} \log n$ y $\text{Var}(T_n) = O(1)$.*

DEMOSTRACIÓN. Definamos T_i como el tiempo cuando exactamente i nodos están activos. Así $T_{i+1} - T_i$ es el tiempo necesario para que i nodos activos logren activar al $i + 1$, por definición este tiempo es el mínimo de $i(n - i)$ exponenciales de tasa $\frac{\lambda}{n}$, luego $T_{i+1} - T_i \sim \text{Exp}(\lambda \frac{i(n-i)}{n})$. Sigue que por linealidad de la esperanza,

$$\mathbb{E}(T_n) = \sum_{i=1}^{n-1} \mathbb{E}(T_{i+1} - T_i) = \sum_{i=1}^{n-1} \frac{n}{\lambda i(n-i)} = \frac{1}{\lambda} \sum_{i=1}^{n-1} \left(\frac{1}{i} + \frac{1}{n-i} \right) \sim \frac{2}{\lambda} \log n.$$

Para obtener la varianza es análogo, pero recordando que $\sum_{i=1}^n \frac{1}{i^2} \leq \frac{\pi^2}{6}$. □

Observación De los cálculos anteriores se desprende que si las tasas de comunicación no se normalizan, entonces $T_n \rightarrow 0$ casi seguramente, por lo que el modelo sería de escaso interés. Además, la demostración se puede adaptar para el caso heterogéneo obteniendo $\mathbb{E}(T_n) \leq \frac{2H_n}{\phi}$, con ϕ la conductancia de la matriz $\Lambda = (\lambda_{u,v})_{u,v}$.

Para contrastar los modelos discreto y continuo, se menciona que [Doerr et al. \[2012\]](#) estudia el protocolo *push-pull* en el modelo de *preferential attachment*, que logra informar a casi toda la red en tiempo $O(\sqrt{\log n})$, que es considerablemente menor que lo logrado por el modelo discreto. Así pues, ambos modelos están estrechamente relacionados, por ejemplo en la proposición 2.27 demostramos una convergencia de un modelo discreto (con estrategia adaptativa) a uno continuo.

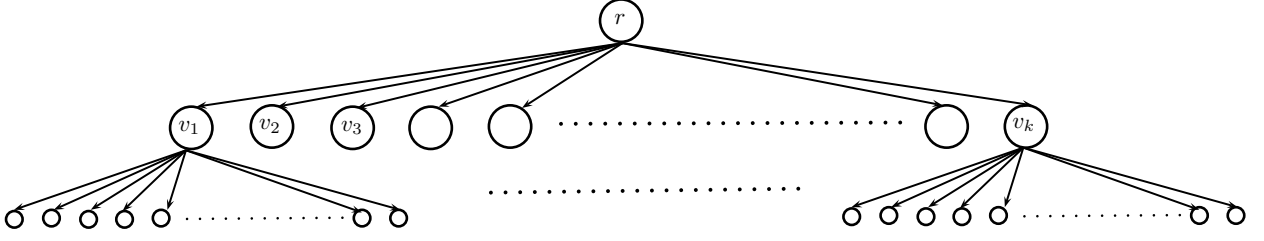
2.2. Brecha de adaptividad bajo heterogeneidad

En esta sección se desarrolla un ejemplo que muestra como construir una familia de instancias donde la brecha de adaptividad crece arbitrariamente con n , con lo que la proposición 1.2 quedaría demostrada. Más aún, se enfatiza que basta tomar dos tasas distintas; cero (imposibilidad de contacto) y tasa unitaria. Este apartado complementa a la sección A.3, donde pueden encontrarse los detalles.

Consideramos el grafo de la Figura 2.1, la raíz r tiene k hijos y cada uno de los v_i tiene a su vez k hijos, de esta forma el número de nodos es $k^2 + k + 1$. Los arcos existentes tienen una tasa normalizada, es decir, padre e hijo se comunican según un tiempo $\text{Exp}(1/k)$, luego cada v_i y la raíz buscan comunicarse con tasa unitaria.

Intuitivamente, elegiremos un *deadline* τ lo suficientemente grande como para que la raíz informe a muchos nodos intermedios, pero con alta probabilidad quedará alguno de ellos inactivo o bien se activará muy tarde. Si un nodo intermedio v_i se activa tarde tiene costo aproximadamente k , entonces No-Adaptativo está forzado a informarlos a todos, por eso resultará un costo es casi lineal en k . Por otro lado, Adaptativo empujará en un tiempo t' que es lo suficientemente grande como para que la raíz active a muchos v_i , pero aún queda

Figura 2.1: Algoritmo adaptativo en un árbol



un tiempo $\tau - t'$ necesario como para que se informen casi todos los hijos de los v_i recién activados.

No-Adaptativo informa a m de los nodos intermedios v_i y posiblemente a la raíz r , pues es claramente subóptimo informar a una hoja. Primero veremos que las únicas opciones válidas son $m = 0$ y $m = k$, i.e. sólo la raíz o bien todos los intermedios. En efecto, dado un *deadline* τ , supongamos que v_i se activa en tiempo $x \in [0, \tau]$, luego cada uno de sus hijos estará inactivo en tiempo τ de forma independiente con probabilidad $e^{-(\tau-x)/k}$. Si la raíz está activa en tiempo 0 y v_i no, el número esperado de hijos inactivos de v_i en tiempo τ es

$$U_i := k \left(\int_0^\tau \frac{1}{k} e^{-x/k} e^{-(\tau-x)/k} dx + e^{-\tau/k} \right) = e^{-\tau/k} (\tau + k).$$

Luego, informando la raíz y m nodos intermedios el costo esperado es

$$1 + m + (k - m)(U_i + e^{-\tau/k}) + mke^{-\tau/k},$$

donde la linealidad en m muestra que las soluciones posibles son aquellas extremas, $m = 0$ y $m = k$. Fijando ahora $\tau = k \log(k \log k)$, una verificación algebraica sencilla muestra que la decisión óptima es $m = k$ y No-Adaptativo paga $1 + k + \frac{k}{\log k}$.

Consideremos un esquema adaptativo, no necesariamente óptimo, que informa sólo a la raíz y luego en tiempo $t' := k \log \log k$ empuja el rumor a todos los nodos v_i inactivos. Cada v_i estará informado según un tiempo distribuido como $\min\{\text{Exp}(1/k), t'\}$ y por tanto su número esperado de hijos inactivos es

$$U_i^A = k \left(\int_0^{t'} \frac{1}{k} e^{-x/k} e^{-(\tau-x)/k} dx + e^{-t'/k} e^{-(\tau-t')/k} \right) = e^{-\tau/k} (t' + k) = \frac{\log \log k + 1}{\log k}.$$

Como a cada v_i se le empuja el rumor con probabilidad $e^{-t'/k} = \frac{1}{\log k}$, sigue que el costo esperado de esta estrategia es $1 + k \frac{\log \log k + 2}{\log k}$, con lo que la brecha de adaptividad $\rho(\tau)$ es al menos $\frac{\log k}{\log \log k}$.

2.3. Resultados preliminares de No-Adaptativo

Para poder acotar la brecha de adaptividad es necesario primero comprender el proceso sin interferencia. En esta sección se dan estimaciones de la evolución del proceso de rumor no-adaptativo, además de una caracterización del esquema óptimo. Este apartado complementa a la sección A.3, donde pueden encontrarse los detalles.

Como se comentó anteriormente, sólo se necesita describir el número $K(t)$ de nodos activos en tiempo $t \in [0, \tau]$. Como es usual, el proceso se define continuo por la derecha, es decir, $K(t^-) \leq K(t) = K(t^+)$ para todo t . Se define también

$$u_k(t) := \mathbb{E}(n - K(\tau) | K(t) = k),$$

en otras palabras, $u_k(t)$ es la cantidad esperada de nodos inactivos en tiempo τ condicional a $K(t) = k$. El número de empujes iniciales elegidos por No-Adaptativo es entonces

$$k_N(\tau) := \operatorname{argmín}\{k \in [n] : u_k(0) + k\}.$$

Proposición 2.2 *Para todo $t \in [0, \tau]$, $k \in [n - 1]$,*

$$u_k(t) = (1 + o(1)) \frac{n}{1 + \frac{k}{n-k} e^{\tau-t}} + o(1).$$

Este resultado se conoce en la literatura de modelos epidémicos, cf. [Bailey \[1950\]](#); [Bartlett \[1978\]](#). Por completitud y para generar intuición sobre el proceso damos un esquema de como obtener la aproximación. Se denotará T_i como el tiempo en que por primera vez hay i nodos activos. Hay dos regímenes a considerar; si hay muchos nodos activos e inactivos entonces la evolución es altamente concentrada, por otro lado, si hay pocos nodos inactivos o bien pocos activos entonces hay variabilidad. Para el primer caso, si por ejemplo $\frac{n}{\log n} \leq K(t) \leq n - \frac{n}{\log n}$, entonces la tasa con que un nodo adicional se activa es $\lambda_{K(t)} \geq \lambda_{n/\log n} = \Omega(\frac{n}{\log n})$, luego la varianza entre tiempos de activación es a lo sumo $O(\frac{\log^2 n}{n^2})$, implicando que $\operatorname{Var}(T_{n-n/\log n} - T_{n/\log n}) = O(\frac{\log^2 n}{n})$, es decir, toda esta “fase media” tiene muy poca aleatoriedad.

Se estudia $X(t) := \frac{K(t)}{n}$, la proporción de nodos activos, observar que considerando una pequeña variación $\frac{1}{n}$ se puede escribir, asumiendo $X(t) = \frac{k}{n}$,

$$\begin{aligned} X(t + \Delta t) &= \lambda_k \Delta t (X(t) + \frac{1}{n}) + (1 - \lambda_k \Delta t) X(t) + o(\Delta t) \\ X(t + \Delta t) - X(t) &= \Delta t X(t) (1 - X(t)) + o(\Delta t), \end{aligned}$$

se demuestra entonces que $X(t)$ sigue con alta probabilidad la curva logística de la ecuación $\frac{dx}{dt} = x(1 - x)$ y cuya solución es exactamente la mostrada en la [proposición 2.2](#).

Por otro lado, si $K(t) \leq \frac{n}{\log n}$, entonces la evolución está bien representada por un proceso lineal de nacimiento (en la literatura conocido como *pure linear birth process*, cf. [Ross \[1995\]](#)). En efecto, si $k = o(n)$, entonces $\lambda_k \approx k$. Se interpreta que, dada la alta cantidad de nodos inactivos, cada vez que un nodo activo busca comunicarse, con alta probabilidad activará otro nodo. En el caso que hay pocos nodos inactivos, e.g. $K(t) \geq n - \frac{n}{\log n}$, entonces se aproxima por un proceso lineal de muerte, ya que $\lambda_k \approx n - k$. En ambos casos existen resultados que permiten aproximar el número de nodos activos, luego juntando con la parte anterior uno obtiene la [proposición 2.2](#).

Ahora podemos obtener una aproximación de la decisión de No-Adaptativo, i.e., de $k_N(\tau)$. Las tasas λ_k son unimodales (creciendo hasta $k = n/2$ y luego decreciendo). Entonces, intuitivamente la estrategia óptima apunta a tener $n/2$ nodos activados en tiempo $\tau/2$, porque

el proceso es simétrico en ese punto, además, haciendo esto las tasas serían lo más grande posibles en esperanza. En conclusión, es natural pensar que el costo k_N pagado al comienzo es aproximadamente el mismo que el costo $\mathbb{E}(n - K(\tau))$ pagado al final; cf. [Sciancalepore et al. \[2014\]](#).

Proposición 2.3 *Dado $\tau > 0$, el esquema no-adaptativo óptimo es tal que*

$$k_N = (1 + o(1)) \frac{n}{1 + e^{\tau/2}} \quad y \quad u_{k_N}(0) = k_N(1 + o(1)).$$

Luego, el costo total esperado de la estrategia óptima es $2(1 + o(1))k_N$.

La demostración sigue de una optimización directa del estimador en la proposición 2.2.

Los resultados anteriores son suficientes para estimar el costo de No-Adaptativo, pero para estudiar la estrategia adaptativa se necesitan propiedades más precisas. Intuitivamente, veremos que Adaptativo interferirá en base a dos criterios: cuanto decrece el costo si hubo una activación en t y cuanto aumenta el costo si no han habido activaciones en un tiempo Δt . El siguiente lema muestra este comportamiento, debido a la extensión presentamos acá la demostración del primer hecho y se interpreta a continuación el segundo.

Lema 2.4 *Para todo $k < n/2$, si $\tau - t = \omega(1)$, entonces*

$$u_k(t + h) \leq u_k(t)e^h \quad \forall h \leq \tau - t \quad y \quad u_{k+1}(t) = u_k(t) \left(1 - \frac{1 + o(1)}{\lambda_k} \right).$$

DEMOSTRACIÓN. Se prueba el primer hecho por inducción reversa en k . Es conveniente extender de forma artificial las funciones $u_k(t)$ como $n - k + 1$ para todo $t > \tau$. Con esta convención, la recurrencia $u_k(t) = \mathbb{E}(u_{k+1}(t + X_k))$ está bien definida para todo $t \leq \tau$, donde $X_k \sim \text{Exp}(\lambda_k)$.

Observar que $u_{n-1}(t) = e^{-\lambda_{n-1}(\tau-t)}$ para todo $t \leq \tau$, como $\lambda_{n-1} = \frac{n-1}{n}$ se obtiene que $u_{n-1}(t + h) \leq u_{n-1}(t)e^h$ para todo $t + h \leq \tau$ y el caso base está demostrado. Finalmente, asumiendo la propiedad cierta para $k + 1$, para todo $t + h \leq \tau$,

$$u_k(t + h) = \mathbb{E}(u_{k+1}(t + h + X_k)) \leq \mathbb{E}(u_{k+1}(t + X_k)e^h) = u_k(t)e^h.$$

□

Debido a la naturaleza exponencial de los tiempos entre activación, es natural pensar que $u_k(t)$ crece exponencialmente según el tiempo que toma una nueva activación, es decir, la diferencia entre $u_k(t)$ y $u_{k+1}(t)$ es $\mathbb{E}(e^{X_k})$ y como $\mathbb{E}(e^{\text{Exp}(\lambda_k)}) = (1 - \frac{1}{\lambda_k})^{-1}$ se obtiene el segundo hecho del lema 2.4.

Terminamos esta sección con otra caracterización de $k_N(\tau)$. Se demuestra que la secuencia discreta $\{u_k(t) + k\}_{k \in [n]}$ es convexa para todo $t \in [0, \tau]$, en particular esto implica que se puede escribir $k_N(\tau) = \min\{k \in [n] : u_k(0) - u_{k+1}(0) \leq 1\}$, pues éste es el primer momento en que un nodo activo adicional reduce el costo en menos de 1 y la convexidad implica que no hay óptimos locales.

Lema 2.5 Para todo $t \in [0, \tau]$ la secuencia $(k + u_k(t))_{k \in [n]}$ es convexa. En consecuencia, k_N se puede tomar como el k más pequeño tal que $u_k(0) - u_{k+1}(0) \leq 1$.

El resultado es muy intuitivo en el sentido de retornos marginales decrecientes, partiendo de sólo un nodo activo y activando uno a la vez, cada nodo adicional hará reducir el costo en una cantidad menor que el anterior. De hecho este resultado se puede generalizar a redes heterogéneas, acá no tiene sentido hablar de la secuencia convexa (pues el costo depende del conjunto activo), se muestra que en este caso $\mathbb{E}(K(\tau)|S(t))$ es una función submodular en $S(t)$, en el trabajo de [Han et al. \[2010\]](#) se pueden encontrar los detalles.

2.4. Caracterización de Adaptativo

En este apartado damos una caracterización intuitiva de la estrategia adaptativa óptima. El controlador tiene la libertad de mirar la red de forma continua, sin embargo al esquema óptimo le basta con consultar la evolución del proceso a lo más $\frac{n}{2}$ veces. En tiempo $t = 0$ se calculan los parámetros $t_1 \leq t_2 \leq \dots \leq t_{n-1}$, llamados *checkpoints*, tales que sólo en esos tiempos el controlador evaluará si empujar un rumor o no. Adaptativo empuja un rumor en tiempo $t = 0$, luego espera hasta tiempo t_1 (posiblemente 0) para empujar si no han habido activaciones adicionales. En general en tiempo t_i empuja un rumor si en $[t_{i-1}, t_i]$ no hubieron activaciones, luego espera hasta tiempo t_{i+1} y repite. El siguiente resultado resume la discusión.

Proposición 2.6 Existen valores $t_1 \leq t_2 \leq \dots \leq t_{n-1}$, dependientes en τ , tales que la estrategia óptima es empujar un rumor en tiempo 0, luego en tiempo t_i empujar un rumor si hay i o menos nodos activos. Más aun, $t_i = \tau$ para todo $i \geq \frac{n}{2}$, es decir, Adaptativo nunca interferirá si hay más de $\frac{n}{2}$ nodos activos.

A menos que τ sea muy grande, una estrategia típica es empujar en tiempo cero $k > 1$ rumores, i.e. $t_1 = t_2 = \dots = t_{k-1} = 0$ y luego el controlador revisará en los tiempos estrictamente crecientes $t_k < t_{k+1} < \dots < t_{n/2-1}$ que se cumpla la condición. La proposición 2.6 es, intuitivamente, consecuencia de la pérdida de memoria del proceso. Para formalizar y demostrar otras propiedades adoptamos un enfoque analítico.

Observación Se enfatiza que todos los *checkpoints* son calculados en $t = 0$, por lo tanto son invariables en la evolución del proceso.

Sea $K^*(t)$ el número de nodos activos bajo el esquema Adaptativo y $P(t)$ el número de empujes hechos hasta (e incluyendo) tiempo t , pero excluyendo los empujes en tiempo τ . Ambos procesos son, como es habitual, continuos por la derecha. En particular $K^*(0) = P(0)$, $P(\tau^-) = P(\tau)$ y en tiempo τ el controlador paga $n - K^*(\tau)$. Será útil definir la función $c_k(t)$, análoga a la función $u_k(t)$ de No-Adaptativo, como el costo esperado que se pagará desde tiempo t en adelante condicional en $K^*(t^-) = k$, i.e.,

$$c_k(t) := \mathbb{E}(n - K^*(\tau^-) + (P(\tau) - P(t^-)) | K(t^-) = k).$$

Observar que si $K^*(t^-) = k$, entonces se tiene la relación $c_k(t) \leq c_{k+1}(t) + 1$, de otra forma se empujaría un rumor en tiempo t . Una propiedad que será muy útil es el hecho que Adaptativo empuja cuando la diferencia de un nodo adicional, cuantificada por $c_k(t) - c_{k+1}(t)$, es grande. Como Adaptativo tiene más control sobre el proceso, se debe tener que en tal momento $u_k(t) - u_{k+1}(t)$ es más grande aún. En otras palabras, un nodo activo adicional beneficia siempre en mayor medida a No-Adaptativo.

Proposición 2.7 *Si $K^*(t^-) = k$, Adaptativo no empujará en tiempo t si se cumple $u_k(t) - u_{k+1}(t) < 1$.*

Observación El recíproco no es cierto, si $K^*(t^-) = k$ y $u_k(t) - u_{k+1} \geq 1$, eso no implica que Adaptativo empujará.

Ahora nos centramos en demostrar la proposición 2.6, para esto introducimos una caracterización general y luego vemos que es equivalente a la inducida por los *checkpoints*. Sea $\gamma_k(t)$ una función que dice el momento óptimo para empujar condicional en $K^*(t) = k$. Eligiendo estas funciones de forma óptima, la estrategia será: si $K^*(t^-) = k$, no se empujará en t si $\gamma_k(t) > t$, donde el controlador verifica “de forma continua” está función, si en algún tiempo $\gamma_k(t) = t$ o bien hubo una activación, entonces el controlador verificará la función $\gamma_{k+1}(t)$ (y pagará un costo unitario en el primer caso). Sea $\tilde{c}_k(t, t')$ el costo esperado si $K^*(t^-) = k$ y se empuja un rumor en tiempo t' , se escribe

$$\begin{aligned}\gamma_k(t) &= \operatorname{argmín}_{t' \in [t, \tau]} \tilde{c}_k(t, t'), \\ c_k(t) &= \operatorname{mín}_{t' \in [t, \tau]} \tilde{c}_k(t, t') = \tilde{c}_k(t, \gamma_k(t)).\end{aligned}$$

2.4.1. Ejemplo introductorio

Veremos los casos donde $n = 2, 3, 4$ para generar intuición sobre el problema y las definiciones presentadas. Si $n = 2$ la estrategia óptima claramente es activar sólo a un nodo en $t = 0$ y no interferir hasta el final de periodo. Si $n = 3$ también es óptimo no interferir, pero este hecho ya no es tan obvio. Veremos que $n = 4$ es el primer caso donde sí es conveniente interferir y Adaptativo sobrepasa a No-Adaptativo.

Asumamos $n \geq 3$ y τ fijo. Es evidente que $c_{n-1}(t) = e^{-\lambda_{n-1}(\tau-t)}$, pues el último nodo se informa con tasa λ_{n-1} y si sólo resta un inactivo no se harán empujes. Si hay dos inactivos, entonces podría ser óptimo empujar en algún tiempo t' , por definición se tiene la expresión

$$\tilde{c}_{n-2}(t, t') = \int_0^{t'-t} \lambda_{n-2} e^{-\lambda_{n-2}x} c_{n-1}(t+x) dx + e^{-\lambda_{n-2}(t'-t)} (1 + c_{n-1}(t')).$$

Un cálculo directo entrega $\frac{\partial}{\partial t'} \tilde{c}_{n-2}(t, t') = e^{-\lambda_{n-2}(t'-t)} (c'_{n-1}(t') - \lambda_{n-2})$. Notar que, para cada t fijo, la función $\tilde{c}_{n-2}(t, \cdot)$ es decreciente. En efecto, como $c'_{n-1}(t') = \lambda_{n-1} e^{-\lambda_{n-1}(\tau-t')} \leq \lambda_{n-1} \leq \lambda_{n-2}$, sigue que $\frac{\partial}{\partial t'} \tilde{c}_{n-2}(t, t') \leq 0$. Concluimos que si $K^*(t) = n-2$ la solución óptima siempre es $t' = \tau$, i.e., no interferir. Como $c_{n-2}(t) = \tilde{c}_{n-2}(t, \tau)$, asumimos $n > 3$ y de la

expresión integral anterior se obtiene

$$c_{n-2}(t) = \frac{\lambda_{n-2}}{\lambda_{n-2} - \lambda_{n-1}} (e^{-\lambda_{n-1}(\tau-t)} - e^{-\lambda_{n-2}(\tau-t)}) + 2e^{-\lambda_{n-2}(\tau-t)}.$$

Busquemos la estrategia si hay tres nodos inactivos. Para ésto usamos $\tilde{c}_{n-3}(t, t')$ y optimizamos sobre t' la expresión

$$c_{n-3}(t, t') = \int_0^{t'-t} \lambda_{n-3} e^{-\lambda_{n-3}x} c_{n-2}(t+x) dx + e^{-\lambda_{n-3}(t'-t)} (1 + c_{n-2}(t')).$$

Un cálculo directo entrega $\frac{\partial}{\partial t'} \tilde{c}_{n-3}(t, t') = e^{-\lambda_{n-3}(t'-t)} (c'_{n-2}(t') - \lambda_{n-3})$, además se verifica que para cada t fijo la función $\tilde{c}_{n-3}(t, \cdot)$ es convexa. Asumiendo $n = 4$ es fácil optimizar, basta con notar que $\lambda_1 = \lambda_3$ y resolver la ecuación $\frac{\partial}{\partial t'} \tilde{c}_{n-3}(t, t') = 0$, que es equivalente a

$$\frac{\lambda_{n-2}}{\lambda_{n-2} - \lambda_{n-1}} (\lambda_{n-1} e^{-\lambda_{n-1}(\tau-t')} - \lambda_{n-2} e^{-\lambda_{n-2}(\tau-t')}) + 2\lambda_{n-2} e^{-\lambda_{n-2}(\tau-t')} - \lambda_{n-1} = 0.$$

Hacemos el cambio de variables $y = e^{-\frac{\tau-t'}{4\lambda}}$, con lo que la ecuación se reduce a

$$12y^3 - 8y^4 - 3 = 0 \quad y \in [0, 1].$$

La ecuación anterior tiene sólo una raíz en $[0, 1]$, que denotamos $r \approx 0,82$. Finalmente, reemplazando el cambio de variable obtenemos $t' = \tau - \frac{4}{\lambda} \ln(1/r) \approx \tau - \frac{0,79}{\lambda}$. Concluimos que, si $\tau > \frac{4 \ln(1/r)}{\lambda}$, entonces la estrategia óptima es empujar $\frac{4 \ln(1/r)}{\lambda}$ unidades de tiempo antes del *deadline*. De otra forma, si $\tau \leq 4\lambda \ln(1/r)$, entonces la estrategia óptima es empujar inmediatamente.

Asumamos $\lambda = 1$. Un *deadline* interesante es cuando No-Adaptativo está indiferente entre empujar uno o dos rumores al comienzo. Tal tiempo es $\tau \approx 1,36$ y No-Adaptativo paga aproximadamente $\text{costo}_{NA}(\tau) \approx 2,92$. Si aplicamos la estrategia óptima, se empezaría sólo con un nodo y en tiempo $\tau - 4 \ln(1/r) \approx 0,56$ se empuja un rumor si siguen tres nodos inactivos. Adaptativo paga en este escenario $\text{costo}_A(\tau) \approx 2,88$.

2.4.2. Caso general

Ahora veremos como extender el razonamiento anterior al caso general. El siguiente resultado (casi) demuestra la caracterización de la proposición 2.6, pues se establece que las funciones $\gamma_k(t)$ son lineales.

Proposición 2.8 *Para todo $k \in [n-1]$, la función $c_k(\cdot)$ es convexa y existe $t_k \in [0, \tau]$ tal que*

$$\gamma_k(t) = \begin{cases} t_k, & \text{si } t < t_k, \\ t, & \text{si } t \geq t_k. \end{cases}$$

DEMOSTRACIÓN. Para todo $k \in [n - 1]$ se tiene la relación

$$\tilde{c}_k(t, t') = \int_t^{t'} \lambda_k e^{-\lambda_k(x-t)} c_{k+1}(x) dx + e^{-\lambda_k(t'-t)} (c_{k+1}(t') + 1),$$

luego un cálculo directo da la útil fórmula

$$\frac{\partial}{\partial t'} \tilde{c}_k(t, t') = e^{-\lambda_k(t'-t)} \left(\frac{d}{dt} c_{k+1}(t') - \lambda_k \right). \quad (2.1)$$

Ahora se procede por inducción reversa. Como $c_{n-1}(t) = e^{-\lambda_{n-1}(\tau-t)}$ el caso base es trivial fijando $t_{n-1} = \tau$. Suponiendo que $c_{k+1}(\cdot)$ es convexa, se muestra que $\gamma_k(t)$ es como se afirma, luego se muestra que $c_k(\cdot)$ es convexa completando la inducción.

Sea $t \geq 0$ tal que $\gamma_k(t) > t$ y sea $h \in [t, \gamma_k(t)]$ cualquiera. Se mostrará que

$$\tilde{c}_k(t+h, t') = e^{\lambda_k h} (\tilde{c}_k(t, t') - \alpha_{h,t}) \quad \forall t' \in [t+h, \tau], \quad (2.2)$$

donde $\alpha_{h,t}$ depende sólo de h y t , pero no de t' , entonces tomando el mínimo en (2.2) se concluye que $\gamma_k(t+h) = \gamma_k(t)$. Para probar (2.2) se calcula directamente:

$$\begin{aligned} \tilde{c}_k(t+h, t') &= \int_{t+h}^{t'} \lambda_k e^{-\lambda_k(x-t-h)} c_{k+1}(x) dx + e^{-\lambda_k(t'-t-h)} (c_{k+1}(t') + 1) \\ &= e^{\lambda_k h} \left(\int_t^{t'} \lambda_k e^{-\lambda_k(x-t)} c_{k+1}(x) dx + e^{-\lambda_k(t'-t)} (c_{k+1}(t') + 1) - \alpha_{h,t} \right) \\ &= e^{\lambda_k h} (\tilde{c}_k(t, t') - \alpha_{h,t}). \end{aligned}$$

Es claro entonces que $\gamma_k(t) > t$ implica $\gamma_k(t+h) = \gamma_k(t)$. Para finalizar la caracterización de $\gamma_k(\cdot)$ se muestra que, si $\gamma_k(t) = t$, entonces $\gamma_k(t+h) = t+h$. Observar que esto se tendría si $\tilde{c}_k(t, \cdot)$ fuera convexa. La convexidad de $\tilde{c}_k(t, \cdot)$ es fácil, pues aplicando (2.1)

$$\frac{\partial}{\partial t'} \tilde{c}_k(t, t') = -e^{-\lambda_k(t'-t)} (\lambda_k - c'_{k+1}(t')),$$

donde $-c'_{k+1}(t')$ y $e^{-\lambda_k(t'-t)}$ son decrecientes en t' , luego $\tilde{c}_k(t, \cdot)$ es convexa.

Terminamos la demostración con la convexidad de $c_k(\cdot)$. Si $t_k = 0$ el resultado es trivial, pues $c_k(t) = c_{k+1}(t) + 1$. Asumiendo que $t_k > 0$, una inspección de (2.1) muestra que por optimalidad $c'_{k+1}(t_k) \leq \lambda_k$. Luego, un cálculo directo, usando la regla de Leibniz, muestra que

$$\begin{aligned} c_k(t) &= \int_0^{t_k-t} \lambda_k e^{-\lambda_k x} c_{k+1}(t+x) dx + e^{-\lambda_k(t_k-t)} (c_{k+1}(t_k) + 1), \\ \frac{d}{dt} c_k(t) &= \int_0^{t_k-t} \lambda_k e^{-\lambda_k x} c'_{k+1}(t+x) dx + \lambda_k e^{-\lambda_k(t_k-t)}, \\ \frac{d^2}{dt^2} c_k(t) &= \int_0^{t_k-t} \lambda_k e^{-\lambda_k x} c''_{k+1}(t+x) dx - \lambda_k e^{-\lambda_k(t_k-t)} c'_{k+1}(t_k) + \lambda_k^2 e^{-\lambda_k(t_k-t)}. \end{aligned}$$

Como se mencionó, $c'_{k+1}(t_k) \leq \lambda_k$ y $c_{k+1}(\cdot)$ se supone convexa, luego $c''_k(t) \geq 0$ y la demostración está completa. \square

Observación Si los parámetros t_k no se eligen de forma óptima, entonces las funciones pueden no ser convexas. Un caso particular de esto es si se asigna $t_k = \tau$, i.e. la estrategia no-adaptativa.

De la demostración se desprenden dos hechos importantes. Primero, la convexidad de las funciones $c_k(\cdot)$ sugiere que, a medida que avanza el tiempo, hay menos aleatoriedad en el proceso, pues el cambio del costo por no observar activaciones en $[t, t + \Delta t]$ es menor cuando t se acerca al *deadline*. Entonces, intuitivamente Adaptativo obtiene la mayor ventaja al comienzo del periodo. Segundo, se verifica fácilmente que si t_k es una solución interior, i.e. $t_k \in (0, \tau)$, entonces $c'_{k+1}(t_k) = \lambda_k$, además se verá en la proposición 2.9 que en este caso también $c'_k(t_k) = \lambda_k$. Adaptativo elige entonces t_k de forma que las tasas de activación promedio con k nodos es igual al caso con $k + 1$, lo que es bastante intuitivo.

Proposición 2.9 Para todo $k \in [n - 1]$, se tiene

$$\frac{d}{dt}c_k(t) = \begin{cases} \lambda_k(c_k(t) - c_{k+1}(t)), & t \leq t_k, \\ \frac{d}{dt}c_{k+1}(t), & t > t_k. \end{cases}$$

DEMOSTRACIÓN. Supongamos $K^*(t) = k$. Gracias a la proposición 2.8 sabemos que, si $t > t_k$ o bien $t_k = 0$, entonces Adaptativo empuja inmediatamente y $c_k(t) = c_{k+1}(t) + 1$, obteniendo el resultado en este caso. Si $t < t_k$, entonces $\gamma_k(t) = t_k$ y un cálculo directo da

$$\begin{aligned} c_k(t) &= \int_t^{t_k} \lambda_k e^{-\lambda_k(x-t)} c_{k+1}(x) dx + e^{-\lambda_k(t_k-t)} (c_{k+1}(t_k) + 1), \\ c'_k(t) &= \lambda_k \left(\int_t^{t_k} \lambda_k e^{-\lambda_k(x-t)} c_{k+1}(x) dx - c_{k+1}(t) + e^{-\lambda_k(t_k-t)} (c_{k+1}(t_k) + 1) \right) \\ &= \lambda_k (c_k(t) - c_{k+1}(t)). \end{aligned}$$

Observar que si $t_k = \tau$ el resultado es consecuencia de lo ya demostrado. Finalmente, si $t_k \in (0, \tau)$, por ser solución interior se cumple $c'_{k+1}(t_k) = \lambda_k$ y entonces (inductivamente) la función es efectivamente derivable. \square

Damos a continuación otra interpretación de la ecuación diferencial de la proposición 2.9. Estudiamos a No-Adaptativo con las funciones $u_k(t) = \mathbb{E}(K(\tau) | K(t) = k)$, ésto es más fácil pues no hay que dividir el análisis en casos. En efecto, si hay k nodos activos en tiempo t , en $[t, t + \Delta t]$ no habrá una activación con probabilidad $1 - \lambda_k \Delta t + o(\Delta t)$ y habrá una con probabilidad $\lambda_k \Delta t + o(\Delta t)$. Así, el costo evoluciona según la ecuación

$$u_k(t) = (1 - \lambda_k \Delta t) u_k(t + \Delta t) + \lambda_k \Delta t u_{k+1}(t + \Delta t),$$

obteniéndose $\frac{u_k(t+\Delta t) - u_k(t)}{\Delta t} = \lambda_k (u_k(t + \Delta t) - u_{k+1}(t + \Delta t))$, que corresponde a la misma ecuación diferencial.

Ahora estamos en condiciones de terminar la demostración de la proposición 2.6. El siguiente resultado es intuitivo, pues las tasas λ_k son unimodales (crecen con $1 \leq k \leq n/2$ y decrecen con $n/2 \leq k \leq n - 1$), entonces después de que $\frac{n}{2}$ nodos estén activos un empuje decrecería en promedio las tasas, luego se enterará menos que un nodo adicional y por tanto no es rentable interferir.

Lema 2.10 *La estrategia adaptativa óptima no interfiere si $\frac{n}{2}$ o más nodos están activos. Es decir, para todo $k \geq \frac{n}{2}$ se tiene $t_k = \tau$.*

DEMOSTRACIÓN. Se muestra que $\tilde{c}_k(t, \cdot)$ es decreciente para $k \geq \frac{n}{2}$. Gracias a (2.1), esto es equivalente a probar $c'_{k+1}(t') - \lambda_k \leq 0$ para todo $t' \geq t$. Como $k \geq \frac{n}{2}$ se tiene $\lambda_{k+1} < \lambda_k$, adicionalmente siempre es cierto que $0 \leq c_{k+1}(t) - c_{k+2}(t) \leq 1$. Por inducción, asumiendo que $\gamma_{k+1}(t) = \tau$ y usando la proposición 2.9:

$$\frac{d}{dt}c_{k+1}(t') - \lambda_k = \lambda_{k+1}(c_{k+1}(t') - c_{k+2}(t')) - \lambda_k \leq \lambda_{k+1} - \lambda_k < 0.$$

□

Terminamos la sección mostrando que, después de tiempo $t = 0$, Adaptativo nunca empuja más de un rumor a la vez. Es decir, en tiempo $t = 0$ se empujan $P(0)$ rumores, luego $t_i = 0$ para $i < P(0)$ y $t_i < t_{i+1}$ para todo $P(0) \leq i < n/2$. En particular, siempre $t_{n/2-1} < \tau$ incluso si τ es muy grande.

Lema 2.11 *Para todo $k \leq \frac{n}{2}$, $t_k < \tau$. Además, si $t_{k+1} > 0$, entonces $t_{k+1} > t_k$.*

DEMOSTRACIÓN. Usamos la ecuación diferencial de la proposición 2.9. Es fácil ver que $t_k < \tau$, pues $\frac{\partial}{\partial t'}\tilde{c}_k(t, \tau) > 0$, lo que se verifica directo de (2.1) y el hecho que $\lambda_k < \lambda_{k+1}$. Asumimos ahora $t_k > 0$. Como t_k es solución interior, i.e. $t_k \in (0, \tau)$, se obtiene $c'_{k+1}(t_k) = \lambda_k$. Para ver que $t_k < t_{k+1}$ y terminar la demostración, por convexidad basta mostrar que $\tilde{c}_k(t, t')$ es creciente en t' para $t' = t_{k+1}$. Usando (2.1) es suficiente con mostrar $c'_{k+1}(t_{k+1}) > \lambda_k$. Por inducción, ya se mostró que $c'_{k+1}(t_{k+1}) = \lambda_{k+1}$, como $k < n/2$ se tiene el resultado. □

2.5. Cálculo de la estrategia adaptativa

En esta sección se presentan resultados especialmente útiles para el cálculo numérico de la estrategia adaptativa óptima. Para la obtención de los parámetros t_k es necesario conocer y poder evaluar la función $c_{k+1}(t)$. A grandes rasgos, estas funciones dependen de un número de aproximadamente n coeficientes, que pueden tomar valores arbitrariamente altos, lo que causa inestabilidad numérica y por tanto es necesario evaluar con precisión.

El siguiente resultado muestra como calcular los costos c_k , para $k \geq \frac{n}{2}$, de forma recursiva. La demostración se omite, pues es análoga al lema 2.13, sólo que más sencilla pues hay menos parámetros. Se asume además, por simplicidad, que n es par.

Lema 2.12 *Para $k \geq \frac{n}{2}$, para todo $t \in [0, \tau]$*

$$c_k(t) = \sum_{i=1}^{n-k} \alpha_i^{(k)} e^{-\lambda_i(\tau-t)},$$

donde $\sum_{i=1}^{n-k} \alpha_i^{(k)} = n - k$ y los coeficientes están definidos por

$$\alpha_i^{(k)} = \alpha_i^{(k+1)} \frac{\lambda_k}{\lambda_k - \lambda_i}, \quad \text{para } i \in [n - k - 1],$$

$$\alpha_{n-k}^{(k)} = n - k - \sum_{i=1}^{n-k-1} \alpha_i^{(k)}.$$

Lema 2.13 Para $k \leq \frac{n}{2} - 1$ y para $t \leq t_k$,

$$c_k(t) = \sum_{i=1}^{n/2} \beta_i^{(k)} e^{\lambda_i t} + \sum_{i=k}^{n/2-1} \gamma_i^{(k)} e^{\lambda_i t},$$

donde $\beta_k^{(k)}$ está dado por la ecuación $c_k(t_k) = c_{k+1}(t_k) + 1$ y las demás constantes por

$$\beta_i^{(k)} = \frac{\lambda_k \beta_i^{(k+1)}}{\lambda_k - \lambda_i}, \quad i = 1, \dots, k-1 \quad \text{y} \quad i = \frac{n}{2},$$

$$\beta_i^{(k)} = \lambda_k \left(\frac{\beta_i^{(k+1)}}{\lambda_k - \lambda_i} + \frac{\gamma_i^{(k+1)}}{(\lambda_i - \lambda_k)^2} \right) \quad i = k+1, \dots, \frac{n}{2} - 1,$$

$$\gamma_i^{(k)} = \frac{\gamma_i^{(k+1)} \lambda_k}{\lambda_k - \lambda_i} \quad i = k+1, \dots, \frac{n}{2} - 1,$$

$$\gamma_k^{(k)} = -\lambda_k \beta_k^{(k+1)}.$$

DEMOSTRACIÓN. Procedemos por inducción reversa, el caso base ($k = n/2$) está demostrado por el lema 2.12 usando $\beta_i^{(n/2)} = \alpha_i^{(n/2)} e^{-\lambda_i \tau}$. Asumiendo cierto para $k+1$, comenzamos observando que $t_k < t_{k+1}$, luego para $x \in [0, t_k]$ vale la ecuación

$$c_{k+1}(x) = \sum_{i=1}^{n/2} \beta_i^{(k+1)} e^{\lambda_i x} + \sum_{i=k+1}^{n/2-1} \gamma_i^{(k+1)} e^{\lambda_i x}.$$

Usaremos también la identidad

$$\int_t^{t_k} e^{\lambda x} x dx = \frac{e^{\lambda t}(1 - \lambda t) - e^{\lambda t_k}(1 - \lambda t_k)}{\lambda^2}.$$

Sabemos que $c_k(t) = \lambda_k \left(\int_t^{t_k} e^{-\lambda_k(x-t)} c_{k+1}(x) dx \right) + e^{-\lambda_k(t_k-t)} (c_{k+1}(t_k) + 1)$, entonces cal-

culamos el primer término, pero agrupando todos los términos que acompañan a $e^{\lambda_k t}$,

$$\begin{aligned}
& \int_t^{t_k} e^{-\lambda_k(x-t)} c_{k+1}(x) dx \\
&= \int_t^{t_k} e^{-\lambda_k(x-t)} \left(\sum_{i=1}^{n/2} \beta_i^{(k+1)} e^{\lambda_i x} + \sum_{i=k+1}^{n/2-1} \gamma_i^{(k+1)} e^{\lambda_i x} \right) dx \\
&= \sum_{\substack{i=1 \\ i \neq k}}^{n/2} \beta_i^{(k+1)} e^{\lambda_k t} \int_t^{t_k} e^{(\lambda_i - \lambda_k)x} dx + \beta_k^{(k+1)} e^{\lambda_k t} \int_t^{t_k} 1 dx \\
&\quad + \sum_{i=k+1}^{n/2-1} \gamma_i^{(k+1)} e^{\lambda_k t} \int_t^{t_k} e^{(\lambda_i - \lambda_k)x} x dx \\
&= \sum_{\substack{i=1 \\ i \neq k}}^{n/2} \frac{\beta_i^{(k+1)}}{\lambda_i - \lambda_k} e^{\lambda_k t} (e^{(\lambda_i - \lambda_k)t_k} - e^{(\lambda_i - \lambda_k)t}) + \beta_k^{(k+1)} e^{\lambda_k t} (t_k - t) \\
&\quad + \sum_{i=k+1}^{n/2-1} \gamma_i^{(k+1)} e^{\lambda_k t} \frac{e^{(\lambda_i - \lambda_k)t} (1 - (\lambda_i - \lambda_k)t) - e^{(\lambda_i - \lambda_k)t_k} (1 - (\lambda_i - \lambda_k)t_k)}{(\lambda_i - \lambda_k)^2} \\
&= - \sum_{\substack{i=1 \\ i \neq k}}^{n/2} \frac{\beta_i^{(k+1)}}{\lambda_i - \lambda_k} e^{\lambda_i t} - \beta_k^{(k+1)} e^{\lambda_k t} t + \sum_{i=k+1}^{n/2-1} \gamma_i^{(k+1)} e^{\lambda_i t} \frac{1 - (\lambda_i - \lambda_k)t}{(\lambda_i - \lambda_k)^2} + \beta_k^{(k)} e^{\lambda_k t}.
\end{aligned}$$

Por inspección se concluye el resultado. □

La siguiente proposición muestra como calcular numéricamente de forma eficiente t_k conociendo c_{k+1} y c_{k+2} , entonces con ayuda de los dos resultados anteriores se puede calcular numéricamente la estrategia adaptativa.

Lema 2.14 *Para todo $k \leq \frac{n}{2} - 1$ se tiene que t_k cumple la ecuación*

$$\lambda_k = \lambda_{k+1}(c_{k+1}(t_k) - c_{k+2}(t_k)).$$

DEMOSTRACIÓN. Aplicando la proposición 2.9 y continuidad se obtiene $c'_{k+1}(t_k) = \lambda_k$. Usando la ecuación diferencial $c'_{k+1}(t) = \lambda_{k+1}(c_{k+1}(t) - c_{k+2}(t))$ se obtiene el resultado. □

Si bien es posible encontrar de forma numérica los parámetros que caracterizan la estrategia óptima, no esperamos que se puedan dar fórmulas cerradas para esto. Además, los parámetros t_k se obtienen de ecuaciones polinómicas de alto grado, por lo que es difícil dar caracterizaciones más precisas. Es por esto que en la sección siguiente tomamos otro enfoque, nos concentramos en la caracterización probabilística y aproximamos la estrategia óptima de forma de poder demostrar una brecha constante.

2.6. Brecha de adaptividad para τ pequeño

En esta sección consideramos el caso donde τ es pequeño, más precisamente asumimos $\tau \leq 2 \log \log n$. Se muestra que, en este caso, los costos de Adaptativo y No-Adaptativo difieren en un factor aditivo $O(\sqrt{n} \log^2 n)$, es decir, $\text{costo}_{NA}(\tau) - \text{costo}_A(\tau) \leq O(\sqrt{n} \log^2 n)$. Observar que, gracias a la proposición 2.3, No-Adaptativo empuja inicialmente un número de $k_N(\tau) = (1 + o(1)) \frac{n}{1 + e^{\tau/2}} = \Omega(\frac{n}{\log n})$ rumores iniciales, por lo tanto la diferencia aditiva implica una brecha $1 + o(1)$ para el caso τ pequeño. Este apartado complementa a la sección A.4, donde pueden encontrarse los detalles.

Ya que $k_N(\tau) = \Omega(\frac{n}{\log n})$, los tiempos de activación T_i son altamente concentrados, por ejemplo se demuestra $\mathbb{P}(|T_i - \mathbb{E}(T_i)| \geq \frac{\log^2 n}{\sqrt{n}}) \leq n^{-2}$. Una interpretación ingenua de este hecho es que, al ser los T_i muy concentrados, Adaptativo no puede explotar la información para obtener beneficio. Sin embargo, la concentración de $K(t)$ no es igual a la de los tiempos, pues una pequeña variación en T_i repercute de forma exponencial en $K(t)$. Para poder demostrar el resultado es necesario entonces lidiar con la concentración de $K(t)$, se usa un argumento de acoplamiento (conocido en la literatura como *coupling*) para hacer esto preciso.

Sea S una colección infinita de puntos en \mathbb{R}_+ , con al menos un punto en 0. Aceptamos múltiples puntos en una misma posición. Dado S definimos el análogo a tiempos entre activaciones y construimos un proceso de rumor. Denotamos S_i la posición del i -ésimo punto, sean $X_i^S := \frac{S_{i+1} - S_i}{\lambda_i}$ y $T_i^S := \sum_{j=1}^{i-1} X_j^S$ los tiempos entre activaciones y tiempos de activación respectivamente. Como exigimos que S tenga un punto en cero se tiene siempre $T_1^S = 0$. Sea $(K^S(t))_{0 \leq t \leq \tau}$ definido como sigue: para $i \leq n - 1$ se fija $K^S(t) = i$ para todo $t \in [T_i^S, T_{i+1}^S) \cap [0, \tau]$ y se fija $K^S(t) = n$ para $t \in [T_n^S, \tau]$, con la convención $[a, b] = \emptyset$ si $a > b$.

En lo que sigue N denotará un proceso de Poisson de tasa unitaria. Sea N' definido como N , pero con k_N puntos adicionales en 0. De esta forma $T_i^{N'} = 0$ para todo $i \leq k_N$. Para $i \geq k_N$ se tiene por definición $N'_{i+1} - N'_i \sim \text{Exp}(1)$, además es trivial verificar que $\lambda^{-1} \cdot \text{Exp}(1) \sim \text{Exp}(\lambda)$ para todo $\lambda > 0$, por tanto $X_i^{N'}$ distribuyen como los tiempos originales X_i . Se concluye que la ley de $K^{N'(t)}$ es igual a la del proceso no-adaptativo con k_N empujes iniciales.

Mediante esta transformación del espacio, se considera que el controlador observa el proceso S y puede decidir si adicionar un punto en la posición $s \in \mathbb{R}_+$ basándose sólo en la información $S \cap [0, s)$. Si S_i es el mayor punto en $[0, s)$, adicionar un punto en s corresponde a empujar un rumor en tiempo $T_i + \frac{s - S_i}{\lambda_i}$ cuando habían i nodos activos. Formalizamos definiendo la estrategia adaptativa como una función φ , que a una colección de puntos S le asigna $\varphi(S) \supset S$ correspondiente a los puntos originales más los empujes. Se debe imponer que $\varphi(S) \cap [0, s]$ dependa sólo de $S \cap [0, s)$ para todo s .

Es fácil verificar que el proceso resultante $K^{\varphi(S)}$ tiene la ley correcta. También se nota que cualquier estrategia adaptativa se puede describir de esta forma. En particular, la estrategia óptima se define en función de *checkpoints* s_i que imitan los t_i definidos en las secciones previas. Se hace hincapié en que $\varphi(S) \setminus S$ corresponden a empujes y por tanto para encontrar

la estrategia óptima se resuelve

$$\min_{\varphi} \left\{ \mathbb{E}(|\varphi(S) \setminus S|) + \mathbb{E}(n - K^{\varphi(S)}(\tau)) \right\}.$$

La principal dificultad del problema anterior es que φ puede depender sólo del pasado, atributo que justamente lo hace estocástico. Dada la concentración se intuye que, aun relajando la condición anterior, Adaptativo no podría obtener mucho beneficio. Es así que se define una estrategia **clarividente** como aquella que empuja conociendo toda la realización, i.e., los puntos agregados en $[0, s]$ pueden depender de todo S . Se denota ξ a la estrategia clarividente, así la relajación permite que $\xi(S) \cap [0, s]$ dependa de todo S . Es evidente que el costo de la mejor estrategia clarividente (que escribimos simplemente Clarividente) es una cota inferior al costo de Adaptativo. Dada una realización S , Clarividente resuelve el siguiente problema

$$\min_{\xi} \left\{ |\xi(S) \setminus S| + (n - K^{\xi(S)}(\tau)) \right\} = \min_{\xi} \left\{ |\xi(S) \setminus S| + |\{j \in [n] : T_j^{\xi(S)} > \tau\}| \right\}.$$

Observación Se podría definir una estrategia clarividente análoga, que decida cuando empujar conociendo todas las realizaciones $X_i \sim \text{Exp}(\lambda_i)$ de antemano. Sin embargo, esta estrategia **no** es equivalente a ξ , sino que tiene más poder. La principal característica de Clarividente es que, cada vez que hay un empuje, todos los tiempos posteriores al empuje son recalculados.

El siguiente resultado ilustra el “poco” control que tiene Clarividente. Como no puede eliminar los tiempos largos, sino que mitigarlos, resulta que Clarividente sólo empujará en $t = 0$.

Lema 2.15 *Hay una estrategia clarividente óptima que sólo agrega puntos en el origen, es decir, todos los empujes ocurren en $t = 0$ o bien $t = \tau$.*

DEMOSTRACIÓN. Se fija la colección S de puntos. Supongamos que Clarividente decide $\xi(S) = R$, i.e., agrega los puntos $R \setminus S$ y paga $|R \setminus S| + |\{j \in [n] : T_j^R > \tau\}|$. Denotamos R_i como la posición del i -ésimo punto.

Sea R_j un punto de $R \setminus S$ agregado tal que no está en el origen. Si $j \leq n/2$ afirmamos que moviendo el punto R_j al origen disminuye el costo. Para ver lo anterior, se razona que colocando R_j sobre R_{j-1} el costo disminuye, por lo que inductivamente se tiene la propiedad. Sea R' la colección obtenida de R , pero desplazando R_j sobre R_{j-1} . Es claro que $X_i^R = X_i^{R'}$ para todo $i < j - 1$ y para todo $i > j$. Además $X_{j-1}^{R'} = 0$, con lo que usando $\lambda_j < \lambda_{j+1}$

$$X_{j-1}^R + X_j^R = \frac{R_j - R_{j-1}}{\lambda_{j-1}} + \frac{R_{j+1} - R_j}{\lambda_j} \geq \frac{R_{j+1} - R_{j-1}}{\lambda_j} = X_j^{R'} = X_{j-1}^{R'} + X_j^{R'}.$$

Por lo tanto, todos los tiempos entre activación decrecen y es óptimo colocar R_j en el origen. Si el punto R_j agregado es con $j > n/2$, un argumento análogo muestra que removiendo el punto se mejora el costo, en este caso usando $\lambda_j > \lambda_{j+1}$. \square

Nos podemos concentrar en un esquema Clarividente que sólo agrega puntos en el origen y luego paga en tiempo τ por aquellos no activos. Se reitera que, la diferencia entre este esquema

y No-Adaptativo, es que el primero elige el mejor número inicial de empujes conociendo toda la realización de puntos. Dada la realización S de puntos, Clarividente resuelve el siguiente problema

$$\min_k \left\{ k + |\{j \in [n] : T_j^{\xi(S)} > \tau\}| \right\},$$

es decir, el número de empujes iniciales más los nodos inactivos en tiempo τ , donde $T_j^{\xi(S)} = 0$ para $j \leq k$ y definido por los demás puntos (no agregados) para $j > k$.

Para finalizar el análisis es necesario un argumento de concentración. Se muestra que para cualquier $k \neq k_N(\tau) \pm o(\sqrt{n})$, comenzando con k nodos activos se debe pagar un costo final mayor a $k_N(\tau) + o(\sqrt{n})$ con probabilidad $1 - n^{-3}$. Por lo tanto, usando una cota de unión, se concluye que ningún k logra un costo menor a $2k_N(\tau) - o(\sqrt{n})$ con probabilidad $1 - n^{-2}$. Finalmente, como $k_N = \Omega(\frac{n}{\log n})$, se obtiene la cota $\rho(\tau) \leq (1 + o(1))$ para $\tau \leq 2 \log \log n$. El siguiente resultado formaliza lo anterior.

Lema 2.16 *Sea $\frac{\log^2 n}{\sqrt{n}} < \tau \leq 2 \log \log n$. Dado un proceso de Poisson N , el costo esperado de la estrategia Clarividente es $2k_N(\tau) - O(\sqrt{n} \text{polylog}(n))$.*

Para usar el resultado de concentración es necesario que τ no sea muy pequeño. De otra forma, si $\tau \leq \frac{\log^2 n}{\sqrt{n}}$, es fácil ver que $k_N = n/2 - o(n)$ con lo que las tasas λ_i no cambian durante el proceso. Se nota también que toda la construcción es (casi) independiente de como se genera la colección S de puntos, sólo se usa que el proceso es Poisson para obtener las cotas de concentración, se espera entonces que los mismos argumentos se puedan generalizar para una clase más amplia de procesos de propagación de rumores.

Consecuencias para la versión de maximización de beneficio

Este resultado garantiza que la versión de maximización de beneficio del problema tiene una brecha de adaptividad $1 + o(1)$. En efecto, si $\tau \geq 2 \log \log n$ entonces $k_N = o(n)$ y el beneficio de No-Adaptativo será $n - o(n)$, mientras que Adaptativo tiene un beneficio de a lo más n . Por otra parte, si τ es menor que $2 \log \log n$, pero al menos una constante, el lema 2.16 garantiza que Adaptativo y No-Adaptativo difieren en $o(\sqrt{n})$ empujes y como el beneficio de ambos es $n - o(n)$ la brecha también es $1 + o(1)$. Finalmente, cuando $\tau = o(1)$, como $\lambda_i = \frac{4}{n+o(n)}$, Adaptativo no puede empujar el rumor y cambiar significativamente el costo, por lo que la brecha $1 + o(1)$ queda demostrada.

2.7. Acotando la brecha de adaptividad

Se estudia ahora el caso donde $\tau \geq 2 \log \log n$. Aquí no es suficiente sólo con explotar la concentración del proceso, pues cuando éste comienza hay suficiente aleatoriedad. A grandes rasgos, se puede dividir el proceso en dos etapas; si $K^*(t) = o(n)$ entonces Adaptativo puede utilizar la información del proceso a su favor, pero si $K^*(t) = \Omega(n)$, entonces se ve que la concentración impide a Adaptativo obtener beneficio (tal como en la sección anterior).

Para ser más concretos, cuando hayan menos de $n/4$ nodos activos se acotará el proceso por otro más fácil y se muestra que Adaptativo no interviene mucho. Cuando hay más de $n/4$ nodos activos un argumento de concentración finaliza mostrando que Adaptativo no afecta el proceso de manera significativa. Este apartado complementa a la sección A.5, donde pueden encontrarse los detalles.

La estrategia óptima es difícil de estudiar, pues está definida por ciertos parámetros t_k , los cuales se obtienen como soluciones recursivas de ciertos polinomios. Es por esto que se define otro Adaptativo (tal como en la sección anterior), más fácil de manejar y con mejor costo, tal que entregue una cota superior a la brecha de adaptividad.

Se vio en la proposición 2.7 que, si $K^*(t^-) = k$, entonces Adaptativo no empuja en tiempo t si $u_k(t) - u_{k+1}(t) < 1$. Con esta motivación definimos otro esquema, llamado **adaptativo modificado**, dotado de la capacidad de empujar gratis, pero que sólo puede empujar si hay $k < n/2$ nodos activos en tiempo t^- y $u_k(t) - u_{k+1}(t) \geq 1$. Como Adaptativo siempre cumple la condición anterior, entonces el mejor adaptativo modificado (llamado simplemente Modificado) tendrá un mejor costo.

Denotamos $\tilde{K}(t)$ el número de nodos activos bajo el esquema modificado óptimo y $\tilde{P}(t)$ como el número de empujes hechos hasta (e incluyendo) tiempo t . Describir a Modificado es muy fácil; como no paga por empujar, lo hará tan pronto como se cumpla la condición. Se tiene que, si $\tilde{K}(t^-) = k$, entonces $u_k(t) - u_{k+1}(t) \leq 1$ y si se tuviera la igualdad entonces habría un empuje en tiempo t y $\tilde{K}(t) = k+1$. Lo anterior sugiere que, al igual que Adaptativo, Modificado revisará la red a lo más n veces para ver si se cumple alguna condición.

Se definen parámetros análogos, llamados **checkpoints modificados**, como $\tilde{t}_k := \inf\{t \in [0, \tau] : u_k(t) - u_{k+1}(t) \geq 1\}$, o bien $\tilde{t}_k = \tau$ en el caso que el ínfimo se tome sobre el vacío. En otras palabras, \tilde{t}_k es el primer instante donde conviene empujar si hay k nodos activos. Gracias al lema 2.5 se sabe que $\tilde{t}_k = 0$ para todo $k < k_N(\tau)$, además $\tilde{t}_k < \tilde{t}_{k+1}$ para $k_N(\tau) \leq k < n/2$ y $\tilde{t}_k = \tau$ para $k \geq n/2$. Así, modificado revisa en tiempo \tilde{t}_k si hay más de k nodos activos y de haber k empuja un rumor (por construcción siempre habrán al menos k).

Otra interpretación, que se explotará mucho, es definiendo $\phi_k := u_k(\tilde{t}_k)$ y observando que, condicional a $\tilde{K}(t^-) = k$, Modificado no puede empujar si $u_k(t) < \phi_k$. Los valores ϕ_k se referirán como **umbrales**, pues Modificado sólo puede empujar si el proceso $\tilde{U}(t) := u_{\tilde{K}(t)}(t)$ alcanza el valor $\phi_{\tilde{K}(t)}$. Por otra parte, si $\tilde{U}(t^-) = \phi_{\tilde{K}(t^-)}$, entonces Modificado realizó un empuje en tiempo t y el proceso \tilde{U} decreció en exactamente una unidad. En resumen, $\tilde{U}(t) \leq \phi_{\tilde{K}(t)}$ para todo $t \in [0, \tau]$ y sólo en los tiempos \tilde{t}_k^- puede darse la igualdad, en cuyo caso Modificado empuja.

De la descripción anterior se desprende que Modificado interfiere poco. En efecto, el proceso $u_{K(t)}(t) = \mathbb{E}(n - K(\tau) | K(t))$ es una martingala de Doob (cf. Ross [1995]; Norris [1998] para detalles sobre martingalas), por lo que se espera que fluctúe alrededor de la media $u_{k_N}(0)$. Por otra parte, como $\tilde{K}(t) \succ K(t)$, se tiene que $\tilde{U}(t)$ es una supermartingala, evolucionando en promedio debajo del valor inicial $\tilde{U}(0) = u_{k_N}(0)$. Sin embargo el proceso de los umbrales $\phi_{\tilde{K}(t)}$ es intuitivamente creciente. $\tilde{U}(t)$ debe alcanzar los umbrales para que Modificado pueda interferir, pero estos procesos tienen crecimientos opuestos, por lo tanto habrán muy pocas

intervenciones.

En lo que sigue se formaliza la intuición anterior. Adicional a esto, el resultado siguiente muestra que cada empuje de Modificado (i.e. cada vez que se alcanzó un umbral) reporta un beneficio unitario comparado con No-Adaptativo, por lo que de haber un número constante de intervenciones se podría concluir el resultado de brecha constante.

Proposición 2.17 *Modificado ahorra una unidad de costo respecto de No-Adaptativo por cada empuje después de $t = 0$, i.e.*

$$\mathbb{E}(n - \tilde{K}(\tau)) = u_{k_N}(0) - \mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0)).$$

DEMOSTRACIÓN. Mostramos que $\tilde{U}(t) + \tilde{P}(t)$ es una martingala. Entre empujes, digamos para tiempos $t \in (\tilde{t}_k, \tilde{t}_{k+1})$, la evolución de $\tilde{U}(t)$ es idéntica a la de $u_{K(t)}(t)$ que es una martingala de Doob. Por otro lado, si un empuje ocurre en tiempo t , entonces

$$\tilde{U}(t) + \tilde{P}(t) = \tilde{U}(t^-) - 1 + \tilde{P}(t^-) + 1.$$

Finalmente, se puede usar el teorema de parada para martingalas (ver por ejemplo el libro de [Norris \[1998\]](#)) y concluir

$$\mathbb{E}(\tilde{U}(\tau) + \tilde{P}(\tau)) = \mathbb{E}(\tilde{U}(0) + \tilde{P}(0)) = u_{k_N}(0) + \tilde{P}(0).$$

Como $\tilde{U}(\tau) = n - \tilde{K}(\tau)$ la demostración queda completa. \square

Se usa el resultado anterior como sigue. Sabemos de la proposición 2.3 que $\text{costo}_{NA}(\tau) = 2(1+o(1))u_{k_N}(0)$. También es claro que $\rho(\tau) \leq \frac{\text{costo}_{NA}(\tau)}{\mathbb{E}(n - \tilde{K}(\tau))}$, suponiendo que $\mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0)) \leq C$, se usa la proposición 2.17 para obtener

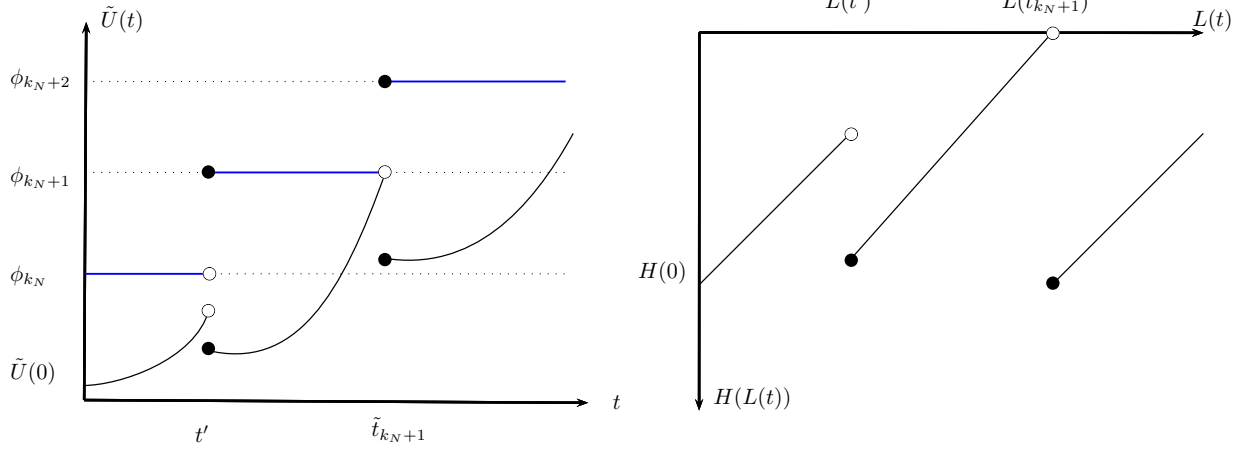
$$\rho(\tau) \leq \frac{\text{costo}_{NA}(\tau)}{u_{k_N}(0) - \mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0))} = 2(1 + o(1)) \left(1 + \frac{1}{u_{k_N}(0)/C - 1} \right).$$

Observación Esencialmente la brecha está acotada por $\rho(\tau) \leq 2(1 + o(1))$. Por otra parte, se observa mediante experimentos computacionales que es posible obtener brechas con valor de 2, pero estos ejemplos aparecen cuando k_N es muy pequeño.

El objetivo del resto de la sección es acotar $\mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0))$ por una constante. Sacamos provecho de la última caracterización de Modificado, para esto cuantificamos el crecimiento de los umbrales y el decrecimiento de $\tilde{U}(t)$. Del lema 2.4 se tiene $u_{k+1}(t) = u_k(t)(1 - \frac{1+o(1)}{\lambda_k})$ y $u_k(t+h) \leq u_k(t)e^h$ para todo $t+h \leq \tau$. En términos del proceso $\tilde{U}(t)$ esto se interpreta fácilmente; entre activaciones $\tilde{U}(t)$ crece de forma subexponencial, pero si en tiempo t hubo una activación (o un empuje) entonces $\tilde{U}(t)$ se multiplica aproximadamente por el factor $1 - \frac{1}{\lambda_{\tilde{K}(t^-)}}$.

Lo anterior sugiere dos transformaciones para obtener un proceso sencillo. Tomando logaritmo se tendrá un proceso con crecimiento sublineal, a esto lo referimos como un cambio espacial. Cuando hay una activación o empuje se multiplica el costo por aproximadamente

Figura 2.2: Una trayectoria de \tilde{U} y su transformación. Una activación ocurrió en tiempo t' y un empuje en tiempo \tilde{t}_{k_N+1} .



$1 - \frac{1}{\lambda_{\tilde{K}(t^-)}}$, mediante un mapeo $L(t)$ apropiado, referido como cambio temporal, se consigue que el factor multiplicativo sea siempre el mismo. Formalmente, se definen

$$H(L(t)) := \lambda_{\tilde{K}(t)} (\log \tilde{U}(t) - \log \phi_{\tilde{K}(t)}), \quad \text{donde} \quad L(t) := \int_0^t \lambda_{\tilde{K}(x)} dx.$$

Como $\tilde{U}(t) \leq \phi_{\tilde{K}(t)}$ sigue que H es siempre no positivo, además si alcanza el valor cero es porque Modificado empujó un rumor. En la Figura 2.2 se muestra una realización de los procesos. Observar como luego de cada activación o empuje los umbrales (lineas sólidas) crecen, mientras que \tilde{U} decrece.

Veremos ahora que H puede ser dominado por un proceso sencillo. Seguimos denotando por N un proceso de Poisson de tasa unitaria y \tilde{T}_i se define como el tiempo en que Modificado activó por primera vez a i nodos.

Proposición 2.18 *Para cualquier constante $0 < c < 1/2$, existe $c' = (1 + o(1))c$ tal que durante la evolución $s \in [0, L(\tilde{T}_{cn})]$, se cumple $H(s) \prec s - 2(1 - c')N(s) - 2(1 - c')(\tilde{P}(s) - \tilde{P}(0))$.*

La proposición sigue de las aseveraciones anteriores. El proceso que domina es lineal pues \tilde{U} crecía de forma subexponencial y se tomó logaritmo. Se verifica que por cada activación el costo decrece a lo menos en $2(1 - c')$ y, como se hizo una transformación temporal, las activaciones siguen un proceso de Poisson unitario, por lo tanto $-2(1 - c')N(s)$ corresponde a como decrece el proceso gracias a su componente aleatoria. Finalmente, se muestra que cuando hay un empuje el costo también decrece a lo menos en $2(1 - c')$, lo que explica el último término.

El siguiente resultado muestra como acotar el número de veces que H alcanza el valor cero, i.e., el número de empujes de Modificado. Lo anterior unido al rápido descenso de H permite (casi) concluir que el número de empujes es constante.

Lema 2.19 *Para todo $t \in [0, \tilde{T}_{cn}]$, el número de empujes después de tiempo t se puede*

dominar por

$$\tilde{P}(t) - \tilde{P}(0) \prec 1 + \frac{1}{2(1-c')} \max\left\{0, \sup_{0 \leq r \leq L(t)} r - 2(1-c')N(r)\right\}.$$

Se toma el máximo en caso que el supremo resulte negativo, lo que sucede cuando no hay ningún empuje. Vemos ahora la dominancia. En cada empuje H toma el valor cero, definiendo θ como el último tiempo en que se observó un empuje, se obtiene $H(L(\theta^-)) = 0$. Gracias a la proposición 2.18 sabemos que $H(s) + 2(1-c')(\tilde{P}(s) - \tilde{P}(0)) \prec s - 2(1-c')N(s)$. Ahora denotamos $s = L(\theta)$ y se obtiene de la expresión anterior $2(1-c')(\tilde{P}(s^-) - \tilde{P}(0)) \prec s - 2(1-c')N(s^-)$. El resultado es claro, pues $\tilde{P}(s^-) = \tilde{P}(s) - 1$ y $s - 2(1-c')N(s^-) \leq \sup_{0 \leq r \leq L(t)} r - 2(1-c')N(r)$.

Considerando Modificado entre $t = 0$ y el tiempo en que se activan $n/4$ nodos, se puede dominar H por un proceso $s - \frac{3}{2}N(s)$, que tiene pendiente $-\frac{1}{2}$ y por lo tanto es fácil mostrar que su valor máximo en esperanza es constante. Luego, una vez que hay $n/4$ nodos activos se usa un argumento de concentración para concluir que el proceso sigue negativo con alta probabilidad, dominando esta vez por $s - N(s)$. Juntando ambas partes se concluye que el número de empujes es constante y por tanto la brecha de adaptividad también lo es.

2.8. Generalizaciones

En una red social, donde todos los agentes comparten información de manera oportunista, hay muchos mensajes siendo transmitidos, pero un nodo puede no estar interesado en todos. Por ejemplo, si los usuarios transmiten noticias vía una aplicación móvil, el controlador puede estar interesado que una noticia le llegue a los usuarios que tienen ciertas características. Otro ejemplo son redes sociales donde se transmiten cupones de descuento, en este caso el objetivo claramente no es dar a toda la red un rumor (algún cupón en particular), sino que al segmento objetivo. Asumimos que todos los integrantes de la red participan del proceso de rumor, ayudando al esparcimiento a los nodos objetivos.

Es en este contexto que generalizamos el modelo fijando un subconjunto de nodos, llamado *conjunto objetivo*. El controlador desea informar a todos estos nodos, pagando costo unitario por empujar rumores. Podemos pensar que el conjunto objetivo tiene $m \leq n$ nodos y que, cuando un *nodo objetivo* (perteneciente al conjunto) se activa, entonces es retirado del conjunto. Adelantamos a continuación el resultado de esta sección.

Proposición 2.20 *Para todo tamaño m de conjunto objetivo, la brecha de adaptividad es constante.*

El caso de τ pequeño sigue de argumentos de concentración análogos al problema original. Vemos acá como extender los argumentos del caso $\tau \geq 2 \log \log n$. Siguiendo con la notación, $K(t)$ denota el número total de nodos activos e introducimos $M(t)$ como el número de nodos objetivos desinformados, ambos bajo una estrategia no-adaptativa. Se define análogamente

$$u_{m,k}(t) := \mathbb{E}(M(\tau) | M(t) = m, K(t) = k),$$

es decir, el costo esperado que se pagará en tiempo τ . Procedemos como en el problema original, estudiamos la evolución de las funciones $u_{m,k}(t)$ y una estrategia adaptativa modificada que explota este comportamiento. El siguiente resultado facilita el análisis, pues conecta las funciones $u_{m,k}(t)$ con las ya conocidas $u_k(t)$.

Lema 2.21 *Para todo $k \in [n-1]$, $m \in [n-k]$, $t \in [0, \tau]$,*

$$u_{m,k}(t) = \frac{m}{n-k} u_k(t).$$

DEMOSTRACIÓN. Recordar que $u_k(t) = \mathbb{E}(n - K(\tau) | K(t) = k)$, entonces por linealidad de la esperanza se puede escribir como $u_k(t) = (n-k)p_{t,k}$, donde $p_{t,k}$ es la probabilidad de que un nodo en particular permanezca inactivo en tiempo τ . De manera similar, $u_{m,k}(t) = mp_{t,k}$ y reemplazando se obtiene el resultado. \square

Es evidente que, siempre que Adaptativo empuja, le da el rumor a un nodo objetivo. Es directo probar también que Adaptativo no empuja si $u_{m-1,k+1}(t) + 1 > u_{m,k}(t)$, i.e. cuando No-Adaptativo no obtiene beneficio de un nodo objetivo adicional activo. Definimos a Modificado como el esquema que empuja gratis, pero sólo cuando se cumple la condición $u_{m-1,k+1}(t) + 1 \leq u_{m,k}(t)$. En la evolución del proceso se puede activar un nodo objetivo o bien uno no objetivo, el siguiente lema resume como afecta esto el costo.

Lema 2.22 *Para todo $k \in [n-2]$, $m \in [n-k]$, si $\tau - t = w(1)$ entonces*

$$\begin{aligned} u_{m,k+1}(t) &= u_{m,k}(t) \left(1 + \frac{1}{n-(k+1)}\right) \left(1 - \frac{1+o(1)}{\lambda_k}\right), \\ u_{m-1,k+1}(t) &= u_{m,k}(t) \left(1 + \frac{1}{n-(k+1)}\right) \left(1 - \frac{1}{m}\right) \left(1 - \frac{1+o(1)}{\lambda_k}\right), \\ \tilde{u}_{m,k}(t+h) &\leq u_{m,k}(t)e^h \quad \forall t+h \leq \tau. \end{aligned}$$

DEMOSTRACIÓN. Manipulación algebraica simple luego de una aplicación directa de los lemas 2.21 y 2.4. \square

Sea $\tilde{U}(t) := u_{\tilde{M}(t), \tilde{K}(t)}(t)$ el proceso que controla Modificado. Definiendo los umbrales $\phi_{m,k}$ por la relación

$$\phi_{m,k} := \min\{u_{m,k}(t) : u_{m-1,k+1}(t) + 1 \leq u_{m,k}(t)\},$$

se tiene que $\tilde{U}(t) \leq \phi_{\tilde{M}(t), \tilde{K}(t)}$ para todo $t \in [0, \tau]$ y que si la igualdad se alcanza es por que se empujó un rumor. Se observa que Modificado escoge ciertos *checkpoints*, como en el caso original, tal que sólo en estos tiempos se verifica el estado de la red. Sea $\tilde{t}_{m,k} := \inf\{t \in [0, \tau] : u_{m-1,k+1}(t) + 1 \leq u_{m,k}(t)\}$, o bien τ en el caso que el ínfimo se tome sobre el conjunto vacío. En particular $\tilde{t}_{1,k} = \tau$ para todo k . Así, en tiempo $\tilde{t}_{m,k}^-$, si se verifica que hay k nodos activos y m nodos objetivos inactivos, entonces Modificado empuja un rumor.

Se define también una transformación equivalente,

$$H(L(t)) := \lambda_{\tilde{K}(t)} (\log \tilde{U}(t) - \log \phi_{\tilde{M}(t), \tilde{K}(t)}), \quad \text{donde} \quad L(t) := \int_0^t \lambda_{\tilde{K}(x)} dx.$$

Finalmente, el siguiente lema muestra que el proceso H aquí definido decrece de igual forma que el caso original, por lo que en particular la cantidad de empujes es constante y se tiene entonces también una brecha constante en este caso.

Lema 2.23 *Para cualquier constante $0 < c < 1/2$, existe $c' = (1 + o(1))c$ tal que durante la evolución $s \in [0, L(\tilde{T}_{cn})]$, se cumple $H(s) \prec s - 2(1 - c')N(s) - 2(1 - c')(\tilde{P}(s) - \tilde{P}(0))$.*

DEMOSTRACIÓN. Condicionamos en $\tilde{M}(t^-) = m$ y $\tilde{K}(t^-) = k$, por construcción $t \leq \tilde{t}_{m,k}$. Vemos primero el caso $t < t_{m,k}$, tomamos Δt pequeño tal que $t + \Delta t < t_{m,k}$, entonces en el intervalo $[t, t + \Delta t]$ no se empujarán rumores. Si no hay activaciones oportunistas (lo que ocurre con probabilidad $1 - \lambda_k \Delta t + o(\Delta t)$), como $\tilde{U}(t)$ crece de forma subexponencial se verifica fácilmente que la dominancia se tiene.

Con probabilidad $\lambda_k \Delta t + o(\Delta t)$ hubo una activación. Es necesario distinguir entre dos casos, si se activó un nodo objetivo (con probabilidad $p := \frac{m}{n-k}$) o si fue un nodo no objetivo (con probabilidad $1 - p$). Usando el lema 2.22 acotamos la variación,

$$\begin{aligned} H(L(t + \Delta t)) &= \lambda_{k+1} \log\left(\frac{u_{\tilde{M}(t+\Delta t), k+1}}{\phi_{\tilde{M}(t+\Delta t), k+1}}\right) \\ \frac{H(L(t + \Delta t))}{\lambda_{k+1}} &= p \log\left(\frac{u_{m-1, k+1}(t + \Delta t)}{\phi_{m-1, k+1}}\right) + (1 - p) \log\left(\frac{u_{m, k+1}(t + \Delta t)}{\phi_{m, k+1}}\right) \\ &\leq p \log\left(\frac{u_{m, k}(t) \left(1 + \frac{1}{n-(k+1)}\right) \left(1 - \frac{1+o(1)}{\lambda_k}\right) e^{\Delta t}}{\phi_{m-1, k+1}}\right) \\ &\quad + (1 - p) \log\left(\frac{u_{m, k}(t) \left(1 + \frac{1}{n-(k+1)}\right) \left(1 - \frac{1}{m}\right) \left(1 - \frac{1+o(1)}{\lambda_k}\right) e^{\Delta t}}{\phi_{m, k+1}}\right). \end{aligned}$$

Usando la desigualdad $\log(1 + x) \leq x$ para todo $x > -1$, la identidad $\Delta s = \lambda_k \Delta t$ y agrupando términos se tiene

$$\frac{H(L(t + \Delta t))}{\lambda_{k+1}} \leq \log\left(\frac{u_{m, k}(t)}{\phi_{m, k}}\right) + \frac{\Delta s}{\lambda_k} + A - \frac{1 + o(1)}{\lambda_k} + \frac{1}{n - (k + 1)} - \frac{1 - p}{m},$$

donde $A = \log(\phi_{m, k}) - p \log(\phi_{m-1, k+1}) - (1 - p) \log(\phi_{m, k+1})$. Observar que se reconoce ya el término lineal, en efecto, como $\log\left(\frac{u_{m, k}(t)}{\phi_{m, k}}\right) < 0$, entonces

$$\lambda_{k+1} \log\left(\frac{u_{m, k}(t)}{\phi_{m, k}}\right) + \lambda_{k+1} \frac{\Delta s}{\lambda_k} \leq H(s) + \Delta s + o(\Delta s).$$

Para acotar A , notamos que por definición $\phi_{m, k} = u_{m, k}(\tilde{t}_{m, k}) = u_{m-1, k+1}(\tilde{t}_{m, k}) + 1$, además gracias al lema 2.22 se puede relacionar $u_{m, k}(t)$ con $u_{m-1, k+1}(t)$ para todo t . Juntando las dos ecuaciones se verifica que para todo m, k

$$\phi_{m, k} = \frac{1}{1 - \alpha_k \beta_m \gamma_k}, \quad \text{donde} \quad \alpha_k := \frac{n - k}{n - (k + 1)}, \beta_m := \frac{m - 1}{m} \text{ y } \gamma_k := 1 - \frac{1 + o(1)}{\lambda_k}.$$

Con el desarrollo de primer orden del logaritmo se obtiene entonces

$$A \leq \alpha_k \beta_m \gamma_k - \frac{m}{n - k} \alpha_{k+1} \beta_{m-1} \gamma_{k+1} - \frac{n - k - m}{n - k} \alpha_{k+1} \beta_m \gamma_{k+1}$$

Un cálculo algebraico muestra que, como es de esperar, el peor caso (cuando cambia menos el costo) es $k = cn$. Sustituyendo entonces se llega a la cota deseada. \square

2.9. Caso discreto

Consideramos en esta sección un modelo a tiempo discreto y vemos que la estrategia adaptativa óptima se caracteriza de una forma análoga al caso continuo. Se conjetura que en los modelos discretos se tiene también una brecha de adaptividad acotada, esto pues el proceso es menos aleatorio, sin embargo la formalización de esta intuición no es trivial.

En este modelo hay n nodos y el *deadline* τ se considera entero. En cada ronda $t = 0, 2, \dots, \tau - 1$ todos los nodos inactivos buscan comunicarse con probabilidad $q \in [0, 1]$, independiente de los demás. Los nodos que busquen comunicarse eligen a otro nodo al azar y se vuelven activos si éste estaba activo en la etapa anterior. Formalmente, sea Y_t en número de nodos activos al comienzo de la etapa t y X_t el número de nodos que se activaron en la etapa t . Se tiene la relación $Y_{t+1} = Y_t + X_t$ y evidentemente $X_t|Y_t \sim \text{Bin}(n - Y_t, \frac{Y_t}{n}q)$.

Dotamos a la estrategia adaptativa de la capacidad de informar a un número cualquiera de nodos luego de cada etapa, más precisamente, dada la realización $X_t = x$ y la información $Y_t = y$, Adaptativo empuja un número adicional de $A(y, x, t)$ nodos. Ahora se tiene la relación $Y_{t+1} = Y_t + X_t + A(Y_t, X_t, t)$. De forma análoga, se debe pagar una unidad por cada empuje y por cada nodo no informado al comienzo de la etapa τ , i.e. $n - Y_\tau$.

El siguiente resultado caracteriza el comportamiento de la estrategia adaptativa.

Proposición 2.24 *La estrategia adaptativa óptima está completamente caracterizada por unos parámetros $a_1, a_2, \dots, a_{\tau-1}$, dependiendo en (q, τ, n) , tales que al final de la etapa t Adaptativo empuja hasta que hayan al menos a_t nodos activos. Más aún, $a_{\tau-1} = 1$ y $a_{\tau-2} = 1$, es decir, en las últimas dos etapas no se interfiere.*

Asumiendo que la estrategia se caracteriza con los a_t es fácil ver que la elección de $a_{\tau-1}$ es irrelevante y se fija en 1 por simplicidad. Por otro lado, para ver que $a_{\tau-2} = 1$, se desprende de la proposición 2.24 que se puede asumir $Y_{\tau-2} + X_{\tau-2} = 1$. En este escenario, si se empuja un número p de rumores al final de la etapa $\tau - 2$, se sabe que el costo final está dado según

$$\min_{0 \leq p \leq n-1} \{p + n - \mathbb{E}(\text{Bin}(n - p - 1, \frac{p+1}{n}q))\},$$

haciendo el cálculo se obtiene que el óptimo es $p = 0$, independiente de q . En conclusión, en la etapa $\tau - 2$ no se empujarán rumores, por lo que $a_{\tau-2} = 1$ como estaba enunciado.

La demostración de la caracterización en la proposición 2.24 es una consecuencia directa de los dos siguientes lemas.

Lema 2.25 *Si la estrategia adaptativa óptima es $A(y, x, t)$, entonces A depende sólo en t y en $x + y$, i.e. $A(y, x, t) = \tilde{A}(y + x, t)$.*

DEMOSTRACIÓN. Sigue del hecho que la distribución de $X_{t+1}, X_{t+2}, \dots, X_{\tau-1}$ depende sólo de $X_t + Y_t$, formalmente $X_{t+1}|(Y_t, X_t)$ es igual en distribución a $X_{t+1}|(Y_t + X_t)$. \square

Lema 2.26 *Sea $A(y, x, t)$ la estrategia adaptativa óptima. Si $A(y, 0, t) > 0$, entonces para $y' < y$ se tiene $A(y - y', 0, t) = A(y, 0, t) + y'$.*

DEMOSTRACIÓN. Se define $C_t(y, a)$ como el costo esperado si $Y_t + X_t = y$, se empujan a rumores y en las etapas posteriores se sigue la estrategia óptima. Se puede escribir

$$C_t(y, a) = n - \mathbb{E}(Y_\tau | Y_{t+1} = y + a) + \sum_{i=t+1}^{\tau-1} \mathbb{E}(A(Y_i, X_i, i) | Y_{t+1} = y + a) + a,$$

entonces $A(y, 0, t) = \underset{a}{\operatorname{argmín}}\{C_t(y, a)\}$. Observar que de la caracterización anterior sigue que $C_t(y - y', a) = C_t(y, a - y') + y'$. Finalmente, para obtener $A(y - y', 0, t)$ se toma mínimo sobre a en ambos lados de la igualdad $C_t(y - y', a) = C_t(y, a - y') + y'$, con lo que se concluye $A(y - y', 0, t) = y' + A(y, 0, t)$. \square

Con los dos lemas anteriores sigue que A es de la forma

$$A(y, x, t) = \begin{cases} a_t - (x + y) & y + x \leq a_t \\ 0 & y + x > a_t, \end{cases}$$

con lo que la proposición 2.24 queda demostrada.

Demostremos una convergencia de un modelo discreto, este resultado puede ser útil por ejemplo para simular el caso continuo (que es computacionalmente mucho más inestable).

Proposición 2.27 *En el caso homogéneo, para todo τ es posible describir el proceso a tiempo continuo como límite del caso discreto.*

DEMOSTRACIÓN. Fijemos el modelo continuo con tasa λ , veremos que en un modelo discreto se tiene convergencia hacia las variables T_i . Consideremos el modelo discreto con n nodos y el protocolo de comunicación siguiente: en cada etapa los nodos inactivos buscarán comunicarse con probabilidad q , aquellos que busquen comunicación le preguntan a un nodo al azar por el rumor. Para un $h > 0$ pequeño, el *deadline* en este modelo discreto se define como $\tau' := \frac{\tau}{h}$, es decir, luego de τ' etapas el controlador pagará una unidad por aquellos nodos inactivos.

Definamos Y_i como la primera etapa en que hay i nodos activos, veremos que $h(Y_{k+1} - Y_k)$ converge a una v.a. $\operatorname{Exp}(\lambda_k)$, notar que por la discretización del tiempo hY_k está en la misma “escala” que en el modelo continuo.

Sin pérdida de generalidad asumamos que en tiempo $t = 0$ hay k nodos activos. En cada etapa no se activa un nodo adicional con probabilidad $p := 1 - (1 - \frac{kq}{n})^{n-k}$, luego $Y_{k+1} \sim \operatorname{Geo}(p)$. Es fácil verificar que la v.a. $h\operatorname{Geo}(p)$ converge en ley a una exponencial de tasa $-\frac{\ln(1-p)}{h}$. Fijemos ahora el parámetro del modelo $q = \lambda h$, entonces hY_{k+1} converge a una exponencial de tasa $-\frac{\ln((1 - \frac{kq}{n})^{n-k})}{h} = \lambda_k + o(h)$. \square

Capítulo 3

Conclusión

Se analizó una estrategia que utiliza la evolución del proceso de rumor para controlar el sistema de forma óptima. El beneficio de esta información adicional se probó constante en el caso homogéneo.

Se mostró que hay redes donde la estrategia adaptativa supera a la no-adaptativa con amplio margen. Además, la caracterización mediante *checkpoints* sugiere que la complejidad de implementación no es alta, pues Adaptativo no necesita inspeccionar continuamente la red. Los resultados mostrados permiten calcular de forma numérica los parámetros t_i , estos mismos resultados se podrían generalizar a otros grafos con similar estructura, por ejemplo una red formada por comunidades muy densas.

3.1. Trabajos futuros

Es interesante también estudiar bajo que condiciones, más débiles que asumir todas las tasas iguales, se conserva la brecha de adaptividad constante. En particular, el resultado en el grafo completo debiera capturar bien a redes que tengan propiedades de buena expansividad o conductancia, por lo tanto se conjetura que una brecha constante se debiera observar también en estos casos.

Las simulaciones sugieren que una propiedad más fuerte se tiene, la brecha de adaptividad *aditiva* podría ser constante, i.e. , $\text{costo}_{NA}(\tau) - \text{costo}_A(\tau) \leq O(1)$ para todo τ y n . Se deja esto como una pregunta abierta interesante.

Bibliografía

- Acan, H., Colavecchio, A., Mehrabian, A., & Wormald, N. 2014. On the push&pull protocol for rumour spreading. *arXiv:1411.0948v1*.
- Andersson, H., & Britton, T. 2000. *Stochastic Epidemic Models and Their Statistical Analysis*. Lecture Notes in Statistics. Springer.
- Asadpour, A., Nazerzadeh, H., & Saberi, A. 2008. Stochastic submodular maximization. *In: WINE*.
- Bailey, N. 1950. A Simple Stochastic Epidemic. *Biometrika*, **37**, 193–202.
- Barabasi, Albert-Laszlo, & Albert, Reka. 1999. Emergence of scaling in random networks. *science*, 11.
- Barry, K. 2009 (September). *Ford Bets the Fiesta on Social Networking*. <http://www.wired.com/2009/04/how-the-fiesta/>.
- Bartlett, M. 1978. *An Introduction to Stochastic Processes, with Special Reference to Methods and Applications*. Cambridge University Press.
- Bollobas, B., & Thomason, A. 1997. *On Richardson's Model on the Hypercube*. Cambridge University Press.
- Borgs, C., Brautbar, M., Chayes, J., & Lucier, B. 2014. Maximizing Social Influence in Nearly Optimal Time. *In: SODA*.
- Boyd, S., A., Ghosh, Prabhakar, B., & Shah, D. 2006. Randomized Gossip Algorithms. *IEEE Transactions in Information Theory*, **52**(6), 2508–2530.
- Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., & Wiener, J. 2000. Graph Structure in the Web. *Comput. Netw.*, **33**(1-6), 309–320.
- Chen., N. 2008. On the approximability of influence in social networks. *In: SODA*.
- Chen, W., Wang, C., & Wang, Y. 2010. Scalable influence maximization for prevalent viral marketing in large-scale social networks. *In: KDD*.
- Chen, Y., & Krause, A. 2013. Near-optimal Batch Mode Active Learning and Adaptive Submodular Optimization. *In: ICML*.

- Chierichetti, F., Lattanzi, S., & Panconesi, A. 2010. Almost tight bounds for rumour spreading with conductance. *In: STOC*.
- Chierichetti, Flavio, Lattanzi, Silvio, & Panconesi, Alessandro. 2011. Rumor spreading in social networks. *Theoretical Computer Science*, **412**(24), 2602–2610.
- Cisco. 2014 (Feb.). *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018*. http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html, [Online; accessed 28-Oct-2014].
- Demers, A, Greene, D, & Hauser, C. 1987. Epidemic algorithms for replicated database maintenance. *Proceedings of the sixth*
- Doerr, B., & Künnemann, M. 2014. Tight Analysis of Randomized Rumor Spreading in Complete Graphs. *Pages 82–91 of: Proc. of the Meeting on Analytic Algorithmics and Combinatorics*. ANALCO’14.
- Doerr, B., Fouz, M., & Friedrich, T. 2011. Social networks spread rumors in sublogarithmic time. *In: STOC*.
- Doerr, B, Fouz, M, & Friedrich, T. 2012. Asynchronous Rumor Spreading in Preferential Attachment Graphs. *Algorithm Theory - SWAT 2012*, 412.
- Domingos, P., & Richardson, M. 2001. Mining the Network Value Customers. *In: KDD*.
- Dubhashi, D., & Panconesi, A. 2009. *Concentration of Measure for the Analysis of Randomized Algorithms*. Cambridge University Press.
- Erdős, Paul, & Rényi, Alfréd. 1960. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, **5**, 17–61.
- Feige, U., Peleg, D., Raghavan, P., & Upfal, E. 1990. Randomized broadcast in networks. *Random Struct. Alg.*, **1**, 447—460.
- Fill, J.A., & Pemantle, R. 1993. Percolation, first-passage percolation, and covering times for Richardson’s model on the n-cube. *Annals of Applied Probability*, **3**, 593–629.
- Fountoulakis, N., Panagiotou, K., & Sauerwald, T. 2012. Ultra-fast rumor spreading in social networks. *In: SODA*.
- Fountoulakis, Nikolaos, & Panagiotou, Konstantinos. 2013. Rumor spreading on random regular graphs and expanders. *Random Structures & Algorithms*, **43**(2), 201–220.
- Frieze, A., & Grimmett, G. 1985. The shortest-path problem for graphs with random arc-lengths. *Discrete Appl. Math.*, **10**, 57–77.
- Gallo, Giorgio, Grigoriadis, Michael D, & Tarjan, Robert E. 1989. A fast parametric maximum flow algorithm and applications. *SIAM Journal on Computing*, **18**(1), 30–55.

- Giakkoupis, G. 2011a. Tight bounds for rumor spreading in graphs of a given conductance. *Pages 57–68 of: Proceedings of the 28th International Symposium on Theoretical Aspects of Computer Science*. STACS’11, vol. 9. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- Giakkoupis, G. 2014. Tight bounds for rumor spreading with vertex expansion. *In: SODA*.
- Giakkoupis, George. 2011b. *Tight bounds for rumor spreading in graphs of a given conductance*.
- Golovin, D., & Krause, A. 2011. Adaptive Submodularity: Theory and Applications in Active Learning and Stochastic Optimization. *J. Artificial Intelligence Research*, **42**, 427–486.
- Grimmett, G. 2010. *Probability on Graphs: Random Processes on Graphs and Lattices*. Institute of Mathematical Statistics Textbooks. Cambridge University Press.
- Han, B., Hui, P., Kumar, A., Marathe, M., Pei, G., & Srinivasan, A. 2010. Cellular Traffic Offloading through Opportunistic Communications: A Case Study. *Pages 31–38 of: Proceedings of the 5th ACM workshop on Challenged networks*.
- Hirschhorn, M.D. 2011. Approximating Euler’s constant. *Fibonacci Quart.*, **49**, 243–248.
- Horel, T., & Singer, Y. 2015. Scalable Methods for Adaptively Seeding a Social Network. *In: WWW*.
- Immorlica, N., Karger, D., Minkoff, M., & Mirrokni, V. 2004. On the costs and benefits of procrastination: approximation algorithms for stochastic combinatorial optimization problems. *In: SODA*.
- Ioannidis, S., Chaintreau, A., & Massoulie, L. 2009. Optimal and scalable distribution of content updates over a mobile social network. *Pages 1422–1430 of: Proc. IEEE INFOCOM*.
- Janson, S. 1999. One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combinatorics, Probability & Computing*, **8**, 347–361.
- Karp, R., Schindelhauer, C., Shenker, S., & Vocking, B. 2000. Randomized Rumor Spreading. *Pages 565–474 of: Proc. of the 41st Annual Symposium on Foundations of Computer Science*. FOCS’00. IEEE Computer Society.
- Keller, L, Le, A, Cici, B, & Seferoglu, H. 2012. MicroCast: cooperative video streaming on smartphones. *Proceedings of the 10th*
- Kempe, D., Kleinberg, J., & Tardos, E. 2003. Maximizing the Spread of Influence through a Social Network. *In: SIGKDD*.
- Kempe, D., Kleinberg, J., & Tardos, E. 2005. Influential nodes in a diffusion model. *In: ICALP*.

- Kleinberg, J. 2013. Cascading Behavior in Social and Economic Networks. *In: ACM EC*.
- Newman, Mark EJ. 2006. Finding community structure in networks using the eigenvectors of matrices. *Physical review E*, **74**(3), 036104.
- Norris, J. R. 1997. *Markov Chains*. Cambridge University Press.
- Norris, J. R. 1998. *Markov Chains (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press.
- Pelusi, Luciana, Passarella, Andrea, & Conti, Marco. 2006. Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *Communications Magazine, IEEE*, **44**(11), 134–141.
- Pittel, B. 1987. On Spreading a Rumor. *SIAM J. Appl. Math.*, **47**(1), 213–223.
- Ravi, R., & A. Sinha, A. 2004. Hedging uncertainty: Approximation algorithms for stochastic optimization problems. *In: IPCO*.
- Ross, Sheldon M. 1995. *Stochastic Processes (Wiley Series in Probability and Statistics)*. 2 edn. Wiley.
- Sauerwald, T., & Stauffer, A. 2011. Rumor spreading and vertex expansion on regular graphs. *In: SODA*.
- Sciancalepore, V., Giustiniano, D., Banchs, A., & Picu, A. 2014. Offloading Cellular Traffic through Opportunistic Communications: Analysis and Optimization. *ArXiv e-prints*, May.
- Seeman, L., & Singer, Y. 2013. Adaptive Seeding in Social Networks. *In: FOCS*.
- Shmoys, D., & Swamy, C. 2004. Stochastic optimization is (almost) as easy as deterministic optimization. *In: FOCS*.
- Whitbeck, J., & Amorim, M. 2011. Relieving the wireless infrastructure: When opportunistic networks meet guaranteed delays. *World of Wireless, . . .*
- Whitbeck, J., Amorim, M., Lopez, Y., Leguay, J., & Conan, V. 2011. Relieving the Wireless Infrastructure: When Opportunistic Networks Meet Guaranteed Delays. *Pages 1–10 of: Proc. of the 2011 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*. WOWMOM'11. IEEE Computer Society.

Apéndice A

Adaptive Rumor Spreading

ADAPTIVE RUMOR SPREADING

JOSÉ CORREA¹, MARCOS KIWI¹, NEIL OLVER², AND ALBERTO VERA¹

¹*Universidad de Chile, Chile* ²*VU Amsterdam, The Netherlands*

ABSTRACT. Rumor spreading is a basic model for information dissemination in a social network. In this setting a central concern for an entity, say the service provider, is to find ways to speed up the dissemination process and to maximize the overall information spread. In the last decade there have been multiple approaches to deal with this loosely defined problem, including the well known influence maximization problem. A central issue absent in the first model is that of adaptivity. How can the service provider use information about the current state of the network to cost effectively speed up the process?

Motivated by the recent emergence of the so-called opportunistic communication networks, we take a novel approach by considering the issue of adaptivity in the most basic continuous time (asynchronous) rumor spreading process. In our setting a rumor has to be spread to a population and the service provider can push it at any time to any node in the network and has unit cost for doing this. On the other hand, as usual in rumor spreading, upon meeting, nodes share the rumor and this imposes no cost on the service provider. Rather than fixing a budget on the number of pushes, we consider the cost version of the problem with a fixed deadline and ask for a minimum cost strategy that spreads the rumor to every node. A non-adaptive strategy can only intervene at the beginning and at the end, while an adaptive strategy has full knowledge and intervention capabilities. Our main result is that in the homogeneous case (where every pair of nodes randomly meet at the same rate) the benefit of adaptivity is bounded by a constant. This requires a subtle analysis of the underlying random process that is of interest in its own right.

Acknowledgments. J. Correa thanks Nicole Immorlica, Brendan Lucier, and Yaron Singer for stimulating discussions. M. Kiwi would like to thank Antonio Fernández for making possible his visit to IMDEA Networks, as well as triggering the discussions with Albert Banchs and Domenico Giustiniano of preliminary versions of their work and the key aspects of the model they were studying. N. Olver thanks J. Komjáthy for helpful pointers to the literature.

Key words and phrases. Adaptivity, Rumor Spreading, Opportunistic Networks, Viral Marketing.

This work is supported by the Millennium Nucleus Information and Coordination in Networks ICM/FIC RC130003, CONICYT via Basal in Applied Mathematics, and an NWO Veni grant.

A.1. Introduction

A basic question in the study of social networks concerns the diffusion of a rumor, which may refer to adopting a new technology, updating content on a cell phone, or buying a new product or service. In this setting we are given a network in which vertices represent agents and edges represent social links. Initially a single agent knows the rumor and we would like to estimate the time by which the full network is informed. The flow of information is governed by a certain stochastic process which may evolve in discrete or continuous time. The most widely studied discrete time models are the push model and the pull model; in the former at each time step a vertex knowing the rumor pushes it to a random neighbor, while in the latter a vertex not knowing pulls the rumor from a random neighbor. On the other hand in the continuous time, or asynchronous, model every pair of connected vertices meet at random times following a Poisson process. The asynchronous model was first formulated by [Boyd *et al.* \[2006\]](#) as a way around the unrealistic time-synchronization issue implicit in discrete time models.

The diffusion of information through a social network has also posed a number of fundamental algorithmic questions particularly in so-called *viral marketing campaigns*. [Domingos & Richardson \[2001\]](#) started the study of how the initial selection of vertices (who adopt a new product or gets it for free) influences further adoption through a cascading effect. These problems were rigorously addressed by [Kempe *et al.* \[2003\]](#) who designed approximation algorithms for the influence maximization problem subject to a budget constraint on the number of initial nodes to which the rumor is pushed. Interestingly, these viral marketing ideas have permeated not only the technological industry but also more traditional markets, including automotive ones (e.g. see [Barry \[2009\]](#)). Unlike the rumor spreading process described above, Kempe *et al.* consider “static” diffusion models, particularly the *independent cascade model*, in which the spread of the rumour is probabilistic, but time plays no role. An alternative approach, which we take in this paper (see also [Han *et al.* \[2010\]](#)), is to keep the standard rumor spreading process but rather than fixing a budget on the initial set of selected vertices, fix a time horizon and only account for the vertices that receive the rumor within this time.

When working with a dynamic diffusion model a new problem pops up, namely that of *adaptivity*. Already Domingos and Richardson identify this issue and state that: *A more sophisticated alternative would be to plan a marketing strategy by explicitly simulating the sequential adoption of a product by customers given different interventions at different times, and adapting the strategy as new data on customer response arrives.* Along these lines [Seeman & Singer \[2013\]](#) consider a two stage extension of the Kempe *et al.* model. Although not so common in the social networks literature it is worth mentioning that the issue of adaptivity is a basic question in stochastic optimization as reflected by [Asadpour *et al.* \[2008\]](#); [Golovin & Krause \[2011\]](#); [Immorlica *et al.* \[2004\]](#); [Ravi & A. Sinha \[2004\]](#); [Shmoys & Swamy \[2004\]](#).

The central concern of this paper is that of adaptivity when speeding up rumor spreading on a social network. More precisely we are given a network and a fixed deadline. As stated above, the diffusion model is the standard for asynchronous rumor spreading. Thus, every pair of connected vertices meet at random times following a Poisson process, and the rumor

is spread whenever, upon meeting, one vertex knows the rumor and the other one does not. Along the way we are able to *push* the rumor to any vertex in order to speed up the diffusion process. We consider both the profit maximization and the cost minimization versions of the problem which are equivalent from an optimization viewpoint. When maximizing profit we get zero profit for every vertex to which the rumor is pushed (say because we are giving the product for free) and get unit profit whenever a vertex gets the rumor through the diffusion process. The objective is thus to maximize the number of vertices that got the rumor through the diffusion process within the time horizon. The cost minimization version is exactly the opposite; every time we push the rumor we make a unit payment and if a vertex gets the rumor through the diffusion process this cost is not realized. The goal is to minimize the total payment made by the time horizon subject to the constraint that everyone should be informed by then.

A non-adaptive policy does not track the evolution of the process and therefore can only push the rumor at the starting time (and also at the deadline in the cost minimization version). In contrast, an adaptive policy may monitor the evolution of the diffusion process and intervene by pushing the rumor to additional vertices. This paper’s main contribution is to show that the advantage of adaptivity is small (in terms of cost or profit).

The seemingly less natural cost minimization version of our problem actually constitutes our main motivation and find its roots in opportunistic communication networks. The widespread adoption of networked mobile devices and the deployment of new technologies (3G, 4G), through which ever increasing data intensive services can be delivered, has generated an explosion of mobile data traffic. This trend is likely to continue, thus exacerbating current cellular network data overload [Cisco \[2014\]](#). Therefore it is critical for operators and service providers to design networks and communication mechanisms that not only can handle the current traffic overload, but also allow the sort of rapid data dissemination that will be required by the multitude of mobile-enabled devices and applications of the near future. A promising, less than a decade old, proposal to address the cellular network data overload consists in offloading traffic through so called opportunistic communications.

The key idea is for service providers to push mobile application content to a small subset of interested users through the cellular network and let them opportunistically spread the content to other interested users upon meeting them. Opportunistic communication can occur when mobile device users are (temporarily) in each others proximity, making it possible for their devices to establish local peer-to-peer connections (e.g., via Wifi or Bluetooth). Opportunistic communication based services have been proposed across several domains. Dissemination of dynamic content such as news is studied by [Ioannidis *et al.* \[2009\]](#) and empirical tests conducted on two real world data sets. [Whitbeck *et al.* \[2011\]](#) study information dissemination of traffic update information, they use a data set derived from large-scale vehicular mobility excerpts of a city-wide dataset of the municipality of Bologna (Italy). In several of the aforementioned application scenarios, the usefulness and/or relevance of the disseminated content crucially depends on it being opportunely delivered. Moreover, quality of service contractual obligations in subscription based contexts might entail strict deadlines for the delivery of data. Another key issue that arises in this context is the presumed feedback capabilities of network nodes and the service providers’ knowledge of how data has propagated through the network up to a given moment of time.

Whitbeck *et al.* [2011] were the first to study a fixed deadline scenario. They propose a *Push-and-Track* framework where a subset of users receive the content from the infrastructure and start disseminating it epidemically. The main feature of *Push-and-Track* is the closed control loop that supervises the injection of copies of the content via the infrastructure whenever it estimates that the ad hoc mode alone will fail to achieve full dissemination within the target horizon. Upon reaching the deadline, the system enters into a “panic zone” and pushes the content to all nodes that have not yet received it. Sciancalepore *et al.* [2014] initiate a mathematically more rigorous analysis of *Push-and-Track* type proposals. In particular, they derive formulas (although not explicit algebraic expressions) for the optimal number of nodes to initially push data in order to minimize the overall number of pushes. Furthermore, they propose a control theory based adaptive heuristic.

Our work is thus motivated by the natural question left open by Sciancalepore *et al.* [2014]: whether or not an adaptive strategy, that harnesses the accrued information of how data has propagated through the network up to any given instant, can actually outperform an optimal non-adaptive strategy, and to what extent.

Model and Main Contributions

Consider a network of n nodes labeled by the elements of $[n] := \{1, \dots, n\}$. Nodes are presumed mobile and such that the encounters of any two nodes i and j , $i \neq j$, are governed by a Poisson process of rate $\lambda_{i,j}$. Thus the time elapsed between two consecutive encounters of i and j is distributed as an exponential random variable of rate $\lambda_{i,j}$, henceforth denoted by $\text{Exp}(\lambda_{i,j})$. All these random variables are independent, including those associated to distinct inter-encounter intervals for the same pair of nodes. As usual, if upon encountering each other one node is informed (i.e., is active) and the other is not, then the information is spread.

We assume that there is a service provider who wishes to cost efficiently disseminate one unit of information to all nodes within a *deadline* of time $\tau > 0$. The set of nodes that posses the unit of information of interest at time t will be denoted $S(t) \subseteq [n]$ and referred to as the set of *active nodes at time* t . Initially (at time $t = 0$), the service provider selects a set of nodes $S(0)$ and *activates* them by pushing to them the unit of data of interest. Subsequently, nodes become active by either one of the following two mechanisms:

- **Opportunistic communication:** If nodes $i \neq j$ encounter each other at time t and either i or j belong to $S(t^-)$,¹ say j for concreteness, then i becomes informed at time t , i.e., i belongs to $S(t')$ for all $t' \geq t$. When a node becomes informed via opportunistic communication it signals the service provider that his state has changed.
- **Pushes:** Because of the network’s feedback capabilities, at any time $0 \leq t \leq \tau$, the service provider has full knowledge of the evolution of the set of active nodes, i.e., of $(S(t'))_{0 \leq t' < t}$, and based on this knowledge she decides whether or not, and which nodes to activate. Formally, at time t a subset of nodes $C(t) \subseteq [n] \setminus S(t^-)$ is chosen and included in $S(t)$, in which case we say the nodes in $C(t)$ are activated and that $|C(t)|$ pushes are performed at time t .

¹Where $S(t^-) := \cup_{0 \leq t' < t} S(t')$.

For most of the paper we deal with the cost minimization version of the problem in which the service provider incurs a unit cost for activating a single node, independent of the nodes label and the time it happens. When the deadline is reached, all nodes not in $S(\tau)$ must be activated, at a total cost of $n - |S(\tau)|$. Note that this is equivalent, from an optimization perspective, to the maximization problem where the service provider gets unit profit for each node informed via opportunistic communication and zero profit for the pushes she makes. We say that the service provider’s strategy is *non-adaptive* if it can only activate nodes at time $t = 0$ and $t = \tau$. Otherwise, we say that its strategy is *adaptive*.

Of course, the cost of an optimal adaptive strategy is at most that of a non-adaptive strategy that initially activates an optimal number of nodes. A natural question is thus to determine the *adaptivity gap*, denoted $\rho(\tau)$, and defined as the ratio between the cost of an optimal non-adaptive strategy and that of an optimal adaptive strategy. Of particular interest is to determine upper bounds on $\rho(\tau)$ that are independent of $\tau > 0$. This is the main endeavor we undertake in this work. This question is certainly of practical significance – if $\rho(\tau)$ turns out to be close to 1 for realistic ranges of the relevant parameters (n , τ , and $\lambda_{i,j}$ ’s), at least from a purely cost effective point of view, there is no justification for incurring the overhead of relying in the network’s feedback capabilities, nor the extra cost required to implement a more computational demanding adaptive on-line strategy.

From the analysis of standard rumor spreading models, a.k.a. stochastic epidemic models, for which a vast literature and long history of research exists, starting with one active node ensures that the whole network will be active w.h.p in time $2 \log n + o(\log n)$ (for an overview of the area see [Grimmett \[2010\]](#) and [Andersson & Britton \[2000\]](#)). Thus, for very large values of τ , say $\tau \geq (2 + o(1)) \log n$, non-adaptive just pays for one node and the expected number of inactive nodes at the deadline is approximately zero, on the other hand adaptive has to pay for the first node and $\rho(\tau)$ is approximately 1.

Our main result is the following.

Theorem A.1.1 In the homogenous case, i.e., $\lambda_{i,j} = \lambda$ for all $i, j \in [n]$, the adaptivity gap is bounded by a small constant, irrespective of the deadline.

We also show that the adaptivity gap with respect to the profit objective, defined in the obvious way, is at most $(1 + o(1))$. Interestingly our analysis turns out to be significantly different for small and large values of τ . For the case of small τ we use a coupling argument to formalize the intuition that the process is “too deterministic” for adaptivity to win much. This already implies a $(1 + o(1))$ bound on the adaptivity gap of the profit maximization version of the problem. On the other hand, the case when τ is large (though not too large) is significantly harder. Intuitively, in our asynchronous setting the evolution of the process is much more unpredictable when only a small number of nodes are active. The latter is in sharp contrast to the discrete time setting that is more concentrated and thus easier to deal with. Consequently for the large τ case we consider a relaxation of an adaptive strategy which pushes for free but may only do so when a certain condition, one which is always satisfied by an optimal adaptive strategy, is met. The analysis then involves understanding an underlying martingale accounting for the expected final cost given the current situation.

Further Related Work

The existing literature on rumor spreading is vast, particularly in the discrete time (synchronous) model. In this context the push protocol works as follows: In each time step, each informed node (initially just one) chooses a random neighbor to which the rumor is sent. If the neighbor did not have the rumor then it gets informed (possibly with some fixed probability). The process continues until every node is informed. In the pull protocol the situation is reversed so that in each time step each uninformed node chooses a random neighbor from which the rumor is pulled. The natural problem here is to estimate time at which every node in the networks has the rumor. This question is quite well understood and extremely precise estimates are known when the network is a complete graph [Frieze & Grimmett \[1985\]](#); [Karp *et al.* \[2000\]](#); [Pittel \[1987\]](#); [Doerr & Künnemann \[2014\]](#). These estimates state that the time is highly concentrated around a logarithmic function of n , depending on the specific protocol. The situation is similar in more complex networks including hypercubes, random graphs, preferential attachment graphs, power-law random graphs, and expander graphs [Feige *et al.* \[1990\]](#); [Doerr *et al.* \[2011\]](#); [Fountoulakis *et al.* \[2012\]](#); [Chierichetti *et al.* \[2010\]](#); [Giakkoupis \[2014\]](#); [Sauerwald & Stauffer \[2011\]](#). The arguably more realistic continuous time (asynchronous) model is not as well understood [Boyd *et al.* \[2006\]](#). This is possibly due to inherent additional randomness of this process, particularly in the beginning, although logarithmic estimates for the expected time to activate the whole network, starting with one node, have been obtained for various classes of graphs [Bollobas & Thomason \[1997\]](#); [Fill & Pemantle \[1993\]](#); [Janson \[1999\]](#); [Acan *et al.* \[2014\]](#).

Viral marketing is also an area of much interest, particularly in the marketing community, and has received a lot of attention under different models for the diffusion of information [Kleinberg \[2013\]](#). Closest to our work is the influence maximization problem in which the goal is to find a subset of at most k nodes maximizing the total final number of informed nodes. The most studied underlying diffusion model is that of *independent cascades* in which when node v becomes informed it has a single chance of informing each currently uninformed neighbor w and succeeds with probability p_{vw} . This problem was studied by [Kempe *et al.* \[2003\]](#) who showed that the underlying optimization problem is a monotone submodular maximization problem and therefore it can be approximated efficiently within a factor of $(1 - 1/e)$. A long list of follow-up papers have studied the problem (see e.g. [Chen. \[2008\]](#); [Chen *et al.* \[2010\]](#); [Kempe *et al.* \[2005\]](#)) until recently [Borgs *et al.* \[2014\]](#) obtained an almost optimal algorithm for it. Variations of the model that consider a dynamic information diffusion model and where the issue of adaptivity plays a central role have been considered in particular for submodular objectives [Asadpour *et al.* \[2008\]](#); [Golovin & Krause \[2011\]](#); [Chen & Krause \[2013\]](#) and for an adaptive seeding problem [Seeman & Singer \[2013\]](#); [Horel & Singer \[2015\]](#).

A.2. Preliminaries

In this section we further specify the model and the notation we will work with. While introducing the model we try to build some intuition and elicit how it behaves. We also

establish some basic facts, which both capture some of the aforementioned intuition and will be needed in subsequent sections for deriving our main contribution; the $O(1)$ bound on the model's adaptivity gap. Due to their length and technical nature, we postpone the proofs of this section's results to the end of the section.

Recall that our study concerns the homogeneous case, i.e., when the rate $\lambda_{i,j}$ is a fixed value independent of the pair of nodes $i \neq j$ considered. Moreover, everything is invariant if the rates and the deadline τ are both scaled by the same amount, so we assume $\lambda_{i,j} = 1/n$ for all $i \neq j$.

Because of symmetry considerations, the specific labels of active nodes is irrelevant and only their number at any given time matters. We henceforth denote by $K(t)$ the number of active nodes at time t for a non-adaptive scheme. Observe that $K(\cdot)$ is right continuous (formally, $K(t^-) \leq K(t) = K(t^+)$). Also, define

$$u_k(t) := \mathbb{E}(n - K(\tau) | K(t) = k),$$

i.e., the expected number of pushes to be made at the end of the process given $K(t) = k$. We will need some information about the optimal non-adaptive choice $k_N := k_N(\tau)$ for the number of initially active nodes, i.e., the value of $k \in [n - 1]$ that minimizes $k + u_k(0)$. For small values of n , one can compute k_N and $u_{k_N}(0)$ explicitly. To do so, it is convenient to consider the elapsed time between the i -th and $(i + 1)$ -th node activation, henceforth denoted X_i . Since X_i is the minimum of $i(n - i)$ random variables distributed according to $\text{Exp}(1/n)$, well known facts imply that X_i is distributed as $\text{Exp}(\lambda_i)$ for $\lambda_i := i(n - i)/n$.

Analogously, let $K^*(t)$ be the number of active nodes at time t for an **optimal** adaptive scheme, still assuming an explicit deadline τ . Since exponentially distributed random variables are memoryless, the optimal adaptive scheme is completely determined by a sequence $0 \leq t_1 \leq \dots \leq t_n \leq \tau$ (depending on τ), so if at time $t \in \{t_1, \dots, t_n\}$ it holds that $K^*(t^-) < k$, where k is the largest i for which $t_i = t$, then the optimal scheme $K^*(\cdot)$ makes $k - K^*(t^-)$ pushes at time t . Let $P(t)$ denote the number of pushes performed by the optimal scheme up to (and including) time t , but excluding pushes made at the deadline τ (so $P(\tau) = P(\tau^-)$). Clearly, the cost of an optimal adaptive scheme must equal $\mathbb{E}(n - K^*(\tau^-) + P(\tau^-))$. Hence, on average $\mathbb{E}(K^*(\tau^-) - P(\tau^-))$ nodes are activated via opportunistic communication.

In what follows, it will be convenient to denote by $c_k(t)$ the expected cost incurred by an optimal adaptive scheme $K^*(\cdot)$ in the remaining time, given that at time t^- there are exactly k active nodes, i.e.,

$$c_k(t) := \mathbb{E}(n - K^*(\tau^-) + (P(\tau) - P(t^-)) | K^*(t^-) = k).$$

An optimal scheme should push when the expected saving due to having an extra active node at time t compensates the additional cost of making a push – this completely characterizes the values t_k . In fact, whatever τ is, it clearly makes no sense to make n pushes unless the deadline has been reached, so $t_n = \tau$. Also, it makes no sense to wait in order to make a first push, so always $t_1 = 0$. Moreover, given $t_{k+1} \leq t_{k+2} \leq \dots \leq t_n = \tau$, the value of t_k is $\inf\{t \leq t_{k+1} : c_{k-1}(t) - c_k(t) \geq 1\}$ or τ in case the infimum is taken over an empty set.

For $n = 2$, an optimal adaptive scheme is completely characterized by $t_1 = 0$ and $t_2 = \tau$, and thus mimics the optimal non-adaptive scheme. It is not hard to justify that when $n = 3$

the optimal adaptive and non-adaptive strategy also coincide and make a single push at time $t = 0$, whatever the value of τ . When $n = 4$, a (long and tedious) calculation based on our characterization of the t_k 's yields that an optimal adaptive scheme fixes $t_1 = 0$, and $t_3 = t_4 = \tau$. For all τ larger than some absolute constant, the optimal adaptive strategy chooses $t_2 \notin \{0, \tau\}$ and beats an optimal non-adaptive scheme.

We can now start formally stating results that will be useful later on. Our first claim is that an optimal adaptive scheme will not perform pushes once roughly half the network's nodes become active. The intuition is that making a push when $i \geq n/2$ nodes are active costs 1 but henceforth *reduces*, to something less than λ_i , the rate at which nodes become activated (implying higher expected time between successive node activations). Thus, in expectation, there will be less than 1 more active node at time τ^- , implying that one ends saving less than the investment of 1 in the push.

Proposition A.2.1 Optimal non-adaptive never starts with more than $\lceil (n-1)/2 \rceil$ active nodes. Furthermore, an optimal adaptive strategy never pushes at some time t if $K^*(t) \geq \lceil (n-1)/2 \rceil$, i.e., $t_k = \tau$ for all $k \geq \lceil (n-1)/2 \rceil$.

We can think of the optimal adaptive scheme as having a target minimum number of active nodes that depends only on the current time t . The following proposition essentially says that this target is not larger than $k_N(\tau - t)$, the number of initial pushes for the optimal non-adaptive strategy with deadline $\tau - t$.

Proposition A.2.2 Let $k \in [n-1]$ and $0 \leq t < \tau$. If $u_k(t) - u_{k+1}(t) < 1$, then $t_k > t$, i.e., adaptive will not push at time t if $K^*(t^-) \geq k$.

A further insight for the last propositions rely on our next assertion, that we believe is of independent interest. In equal conditions, namely $K(t) = K^*(t)$, one additional active node benefits adaptive more than non-adaptive, in the sense that for both schemes the cost decreases, but for non-adaptive the reduction is more significant. Intuitively, this is explained by the enhanced control an adaptive scheme has over the underlying process, which causes any increment in the number of activated nodes to have a smaller impact.

Lemma A.2.3 For every $k \in [n-1]$, if $0 \leq t \leq t_{k+1}$, then $c_k(t) - c_{k+1}(t) \leq u_k(t) - u_{k+1}(t)$.

Remaining proofs

Throughout this section, we denote the r.v. $X_k \sim \text{Exp}(\lambda_k)$ characterizing the times between activations, for $a, b \in \mathbb{N}$ if $a \leq b$, then $X_{a..b} := X_a + \dots + X_b$, which is easy to see is an absolutely continuous random variable over the non-negative real numbers. Moreover, given a random variable Y we denote its cumulative distribution by $F_Y(y) := \mathbb{P}(Y \leq y)$ and its density by $f_Y(y)$. For the sake of convenience, for any $a \in \mathbb{N}$, we let $X_{a..a-1}$ be the random variable equal to zero everywhere, so in particular, $F_{X_{a..a-1}}(x)$ is the indicator function of $x \geq 0$ and $f_{X_{a..a-1}}(x)$ will denote an everywhere 0 function. Also, it will be useful to define $u_n(\cdot)$ and $c_n(\cdot)$ as zero over $[0, \tau)$.

We start by proving Lemma A.2.3, we will see that the other results are a consequence. The next two lemmas establish some basic continuity and differentiability properties of $u_k(\cdot)$ and $c_k(\cdot)$.

Lemma A.2.4 For all $k \in [n-1]$, the mapping $u_k : [0, \tau] \rightarrow \mathbb{R}$ is continuous, non-decreasing, and differentiable. In fact, for $0 \leq t < \tilde{\tau} < \tau$,

$$u'_k(t) = \lambda_k(u_k(t) - u_{k+1}(t)) = \sum_{a=0}^{n-1-k} f_{X_{k..k+a}}(\tilde{\tau} - t)(u_{k+a}(\tilde{\tau}) - u_{k+a+1}(\tilde{\tau})).$$

PROOF. By conditioning on the events $X_{k..k+a} \leq \tilde{\tau} < X_{k..k+a+1}$ for $a \in \mathbb{N}$, we get

$$u_k(t) = \sum_{a=0}^{n-1-k} (F_{X_{k..k+a-1}}(\tilde{\tau} - t) - F_{X_{k..k+a}}(\tilde{\tau} - t))u_{k+a}(\tilde{\tau}).$$

Since the $X_{a..b}$'s are all non-negative and continuous, $u_k(\cdot)$ is continuous and differentiable, with derivative

$$u'_k(t) = - \sum_{a=0}^{n-1-k} (f_{X_{k..k+a-1}}(\tilde{\tau} - t) - f_{X_{k..k+a}}(\tilde{\tau} - t))u_{k+a}(\tilde{\tau}).$$

Recalling that $u_n(\cdot)$ is everywhere 0, regrouping the terms of the summation yields the second of the stated identities. (Essentially by definition of the model, it follows that $u_k(\cdot)$ is non-decreasing.)

To prove the first stated identity, note that the probability of more nodes being activated via opportunistic communication between time t and $t + \Delta$, for very small Δ , is roughly $\lambda_k \Delta$, so

$$u_k(t) = u_k(t + \Delta)(1 - \lambda_k \Delta) + u_{k+1}(t + \Delta)\lambda_k \Delta + o(\Delta).$$

Dividing by Δ , regrouping, and making $\Delta \rightarrow 0$ yields the desired identity. \square

Lemma A.2.5 For all $k \in [n-1]$, the mapping $c_k : [0, t_{k+1}) \rightarrow \mathbb{R}$ is continuous, non-decreasing, and differentiable with derivative

$$c'_k(t) = \lambda_k(c_k(t) - c_{k+1}(t)) = \sum_{a=0}^{n-1-k} f_{X_{k..k+a}}(t_{k+1} - t)(c_{k+a}(t_{k+1}) - c_{k+a+1}(t_{k+1})).$$

PROOF. A straightforward adaptation of the proof argument of Lemma A.2.4, but now fixing $\tilde{\tau} := t_{k+1}$ and considering $c_k(t)$ instead of $u_k(t)$ yields

$$c_k(t) = 1 - F_{X_k}(t_{k+1} - t) + \sum_{a=0}^{n-1-k} (F_{X_{k..k+a-1}}(t_{k+1} - t) - F_{X_{k..k+a}}(t_{k+1} - t))c_{k+a}(t_{k+1}).$$

What remains of the proof of the second stated identity uses the fact that $c_k(t_{k+1}) - c_{k+1}(t_{k+1}) = 1$ and otherwise follows the proof argument of Lemma A.2.4. The first claimed equality is proved as the analogue result of Lemma A.2.4 concerning $u_k(\cdot)$. \square

Remark A.2.6 The attentive reader would have noticed that the preceding result's proof implicitly relies on the fact that the t_k 's, as functions of the deadline, satisfy $t_k(\tau - t) = t_k(\tau) - t$ for $0 \leq t < t_k(\tau)$. Intuitively, this is due to the memoryless feature of our model, which makes the adaptive scheme's decision of whether or not to push to depend only on the time remaining until the deadline.

We now have all the ingredients needed to achieve this section's main objective.

PROOF OF LEMMA A.2.3. We claim that the statement holds when $t = t_{k+1}$ for some k . Indeed, whatever the deadline τ , an optimal non-adaptive scheme will initially perform as many pushes as an optimal adaptive scheme. Since for a deadline of $\tau - t_k$ the adaptive scheme initially makes at least k pushes, the same must hold for the non-adaptive scheme. In order for k pushes to be cost effective, it must hold that $1 \leq u_k(t_{k+1}) - u_{k+1}(t_{k+1})$. Since $c_k(t_{k+1}) - c_{k+1}(t_{k+1}) = 1$, the asserted claim follows.

Because of Lemmas A.2.4, A.2.5, and the above established claim, it suffices to show that $c'_k(t) \leq u'_k(t)$ for $0 \leq t < t_{k+1}$. We proceed by backwards induction in k . We take as base case $k = n - 1$, which holds since $t_n = \tau$ and $c_{n-1}(t) = u_{n-1}(t) = e^{-(\tau-t)}$ for $0 \leq t \leq \tau$. Assume $k < n - 1$. Let j be the largest integer in $[n - 1 - k]$ such that $t_{k+j} = t_{k+1}$. Hence, by the way in which the t_k 's are defined, $c_{k+a}(t_{k+1}) - c_{k+1+a}(t_{k+1}) = c_{k+a}(t_{k+1+a}) - c_{k+1+a}(t_{k+1+a}) = 1$ for $a \in \{0, \dots, j - 1\}$. Moreover, by inductive hypothesis and Lemmas A.2.4 and A.2.5, for $a \in [n - 1 - k] \setminus [j]$, we have that $t_{k+1} < t_{k+a}$ and $c_{k+a}(t_{k+1}) - c_{k+a+1}(t_{k+1}) = c'_{k+a}(t_{k+1})/\lambda_{k+a} \leq u'_{k+a}(t_{k+1})/\lambda_{k+a} = u_{k+a}(t_{k+1}) - u_{k+a+1}(t_{k+1})$. Now, an application of Lemma A.2.5 and Lemma A.2.4 with $\tilde{\tau} = t_{k+1}$ yields,

$$\begin{aligned} c'_k(t) &= \sum_{a=0}^{j-1} f_{X_{k..k+a}}(t_{k+1} - t) + \sum_{a=j}^{n-1-k} f_{X_{k..k+a}}(t_{k+1} - t)(c_{k+a}(t_{k+1}) - c_{k+a+1}(t_{k+1})) \\ &\leq u'_k(t) + \sum_{a=0}^{j-1} f_{X_{k..k+a}}(t_{k+1+a} - t)(1 - u_{k+a}(t_{k+1+a}) + u_{k+1+a}(t_{k+1+a})). \end{aligned}$$

Applying the claim established in this proof's first paragraph, concludes the inductive proof. \square

Now we can easily prove the remaining results using $c_k(t) - c_{k+1}(t) \leq u_k(t) - u_{k+1}(t)$.

PROOF OF PROPOSITION A.2.2. Adaptive pushes when the benefit of activating an additional node is 1, i.e. $c_k(t) - c_{k+1}(t) = 1$. If $u_k(t) - u_{k+1}(t) < 1$ as a direct consequence of Lemma A.2.3 adaptive will not push at such time. \square

Let us show the first part of Proposition A.2.1, that non-adaptive will not start with more than $n/2$ active nodes.

Proposition A.2.7 No optimal non-adaptive scheme makes more than $m := \lceil (n - 1)/2 \rceil$ pushes at time $t = 0$.

PROOF. Let $j \geq m$ and observe that for a non-adaptive scheme that starts by making m pushes, it holds that $X_{m..j} = X_m + \dots + X_j > \tau$ if and only if at least $n - j$ nodes remain inactive at time τ^- . Since the number of inactive nodes at time τ^- is precisely $n - K(\tau^-)$,

$$u_m(0) = \sum_{s=1}^{n-m} \mathbb{P}(n - K(\tau^-) \geq s | K(0) = m) = \sum_{j=m}^{n-1} \mathbb{P}(X_{m..j} > \tau). \quad (\text{A.1})$$

On the other hand, for all $m \leq i < n - 1$, since $\lambda_i \geq \lambda_{i+1}$ it follows that X_i is stochastically dominated by X_{i+1} , implying that $X_{m..j}$ is stochastically dominated by $X_{m+a..j+a}$, where $0 \leq a < n - j$. So from (A.1) it follows that $u_m(0) \leq a + u_{m+a}(0)$, which implies the desired conclusion. \square

PROOF OF PROPOSITION A.2.1. Thanks to Proposition A.2.7, non-adaptive start with less than $n/2$ active nodes for any deadline, so for any $k \geq \lceil (n - 1)/2 \rceil$, $u_k(0) - u_{k+1}(0) < 1$. Now fix $t < \tau$, with deadline $\tau - t$ non-adaptive still does not start with k , by memoryless property $u_k(t) - u_{k+1}(t) < 1$ and invoking Proposition A.2.2 the result follows. \square

A.3. Estimates on the evolution of the non-adaptive process

In this section we give a number of useful estimates on the evolution of the non-adaptive process. We also characterize the optimal non-adaptive strategy and its cost.

Proposition A.3.1 If $t \in [0, \tau]$ and $k \in [n - 1]$, then $u_k(t) = (1 + o(1)) \frac{n}{1 + \frac{k}{n-k} \cdot e^{(\tau-t)}} + o(1)$.

This result is essentially well-known (see e.g., Bailey [1950]; Bartlett [1978]), so we only briefly sketch its proof. The evolution of the process starting from (say) $n/\log n$ active nodes, and all the way until all but $n/\log n$ nodes are active, is highly concentrated. With very high probability, it closely follows the solution of the deterministic differential equation $\frac{dx}{dt} = x(1 - x)$, where $x(t)$ denotes the proportion of active nodes at time t . This yields the logistic curve of the above proposition. What remains is to consider the evolution when there are either very few active nodes, or very few inactive ones. Consider the former case. For $i = o(n)$, $\lambda_i \approx i$, and this allows the process to be well-approximated by a linear birth process. For this, exact analytic results are available. Namely, given k initially active nodes, the number of active nodes after time t has a negative binomial distribution, so long as t is small enough that the estimate applies. A similar linear death process approximation holds for the end of the process; stitching together these estimates yields Proposition A.3.1.

We need some more refined estimates on how $u_k(t)$ varies with k and t . These do not follow from Proposition A.3.1, but notice that they would follow immediately if $u_k(t)$ was *exactly* described by the logistic curve.

Lemma A.3.2 For all $k < n/2$, if $\tau - t = \omega(1)$, then

$$u_{k+1}(t) = u_k(t) \left(1 - \frac{1 + o(1)}{\lambda_k} \right) \quad \text{and} \quad u_k(t+h) \leq u_k(t)e^h \quad \forall h \leq \tau - t.$$

PROOF. The proof for the first fact is presented at the end of the section, as requires some additional results. We prove the second fact for all k by backwards induction. It is convenient to artificially extend the functions $u_k(t)$ as $n - k + 1$ for all $t > \tau$. With this convention the recurrence $u_k(t) = \mathbb{E}(u_{k+1}(t + X_k))$ for all $t \leq \tau$ holds, where $X_k \sim \text{Exp}(\lambda_k)$. Note that $u_{n-1}(t) = e^{-\lambda_{n-1}(\tau-t)}$ for all $t \leq \tau$, so $u_{n-1}(t+h) \leq u_{n-1}(t)e^h$ for all $t+h \leq \tau$ and we have the base case. Assuming the property true for $k+1$, we get that for all $t+h \leq \tau$,

$$u_k(t+h) = \mathbb{E}(u_{k+1}(t+h+X_k)) \leq \mathbb{E}(u_{k+1}(t+X_k)e^h) = u_k(t)e^h.$$

□

The expected cost of a non-adaptive strategy starting with k pushes is $k + u_k(0)$. This cost is in fact a convex function of k . Again, this would follow immediately if $u_k(t)$ was precisely described by the logistic formula.

Lemma A.3.3 For every $t \in [0, \tau]$ the sequence $\{k + u_k(t)\}_{k \in [n]}$ is convex. As a consequence k_N can be taken to be the smallest k such that $u_k(0) - u_{k+1}(0) < 1$.

PROOF. The function $\mathbb{E}[K(\tau) | S(t)]$ is a submodular function of $S(t)$, even in the generality of inhomogeneous networks (see [Han et al. \[2010\]](#)). Since in the homogenous case the mapping depends only on $|S(t)|$, we deduce that $\mathbb{E}[K(\tau) | K(t)]$ is a concave function of $K(t)$, and so $u_k(t) = n - \mathbb{E}[K(\tau) | K(t) = k]$ is a convex function of k . The result follows. □

Now we obtain an estimate of the optimal non-adaptive strategy, i.e., the number k_N of nodes activated at the start. The rates λ_k are unimodal (increasing until $n/2$, and then decreasing). Intuitively then, the optimal non-adaptive strategy aims to have $n/2$ nodes activate at time $\tau/2$, so that the rates are on average as large as possible during the evolution. In conclusion the expected amount paid at the end is roughly the same as the cost paid at the start; cf. [Sciancalepore et al. \[2014\]](#).

Proposition A.3.4 Given a deadline τ , the optimal non-adaptive pick is such that

$$k_N = (1 + o(1)) \frac{n}{1 + e^{\tau/2}} \quad \text{and} \quad u_{k_N}(0) = k_N(1 + o(1)).$$

Thus, the total expected cost of the optimal non-adaptive strategy is $2(1 + o(1))k_N$.

We omit the proof, which follows immediately by optimizing using the estimate of Proposition [A.3.1](#).

Now we proceed to prove the remaining part of lemma [A.3.2](#). We will use the usual notation for Markov chains, $\mathbb{E}_k(\cdot) = \mathbb{E}(\cdot | K(0) = k)$ and similarly $\mathbb{P}_k(\cdot) = \mathbb{P}(\cdot | K(0) = k)$.

Lemma A.3.5 For all $t \geq 0$, $k \in [n - 1]$,

$$\frac{d}{dt} \mathbb{E}_k(K(t)) = \mathbb{E}_k(\lambda_{K(t)}). \quad (\text{A.2})$$

PROOF. Take a small h and compare $K(t + h)$ with $K(t)$, if at most one activation occurs in $[t, t + h]$ then the difference is either zero or one:

$$\begin{aligned} \mathbb{E}_k(K(t + h) - K(t)) &= \mathbb{P}_k(\text{one activation in } [t, t + h]) + o(h) \\ &= \mathbb{E}_k(\mathbb{P}(\text{one activation in } [t, t + h] | K(t))) + o(h) \\ &= \mathbb{E}_k(\lambda_{K(t)} h) + o(h). \end{aligned}$$

Taking limit in the last expression we conclude the claim. \square

Lemma A.3.6 For all $t \geq 0$, $k \in [n - 1]$,

$$u_k(t) - u_{k+1}(t) = \frac{\mathbb{E}(\lambda_{K(\tau)} | K(t) = k)}{\lambda_k}. \quad (\text{A.3})$$

PROOF. Notice that $u_k(t) = \mathbb{E}_k(n - K(\tau - t))$ and taking derivatives we conclude $u'_k(t) = -\frac{d}{dt} \mathbb{E}_k(K(\tau - t)) = \mathbb{E}_k(\lambda_{K(\tau-t)})$ thanks to Lemma A.3.5. Finally using Lemma A.2.4 we also have $u'_k(t) = \lambda_k(u_k(t) - u_{k+1}(t))$ and the result follows. \square

PROOF OF LEMMA A.3.2. First note that if $\tau - t = \omega(1)$ then $n - K(\tau - t) = o(n)$ with probability $1 - n^{-2}$. Using Lemma A.3.6, it suffices to show $\mathbb{E}_k(\lambda_{K(\tau)}) = u_k(t)(1 + o(1))$. Write $\lambda_k = (n - k) - \frac{(n-k)^2}{n}$, using $n - K(\tau - t) = o(n)$ w.h.p we conclude $\mathbb{E}_k((n - K(\tau - t))^2) = o(n)\mathbb{E}_k(n - K(\tau - t)) + o(1)$ and thus

$$\mathbb{E}_k(\lambda_{K(\tau-t)}) = \mathbb{E}_k(n - K(\tau - t)) - \frac{1}{n} \mathbb{E}_k((n - K(\tau - t))^2) = u_k(t) - \frac{o(n)u_k(t) + o(1)}{n}.$$

Finally, $\mathbb{E}_k(\lambda_{K(\tau-t)}) = (1 - o(1))u_k(t)$ as needed. \square

A.4. Additive Gap for Small τ

In this section we consider the case in which τ is small. Specifically, we prove that for $\tau \leq 2 \log \log n$ the optimal adaptive and non-adaptive strategies are within an additive factor of $O(\sqrt{n} \log^2 n)$. Note that, thanks to Proposition A.3.4, in this situation the optimal non-adaptive strategy activates $k_N = k_N(\tau) = (1 + o(1)) \frac{n}{1 + e^{\tau/2}} = \Omega(\frac{n}{\log n})$ nodes initially. We see that the non-adaptive evolution is highly concentrated, intuitively this should be enough to conclude that adaptive cannot obtain a significant advantage; we use a coupling argument to make this precise.

Let S be any countable collection of points on \mathbb{R}_+ , with an infinite number of points, and with at least one point at 0, and denote by S_i the position of the i -th point. Associate to S a

counting process $(K^S(t))_{0 \leq t \leq \tau}$ as follows. Let $X_i^S := (S_{i+1} - S_i)/\lambda_i$, and let $T_i^S := \sum_{j=1}^{i-1} X_j^S$, so $T_1^S = 0$. Then for $i \leq n-1$, we set $K^S(t) = i$ for all $t \in [T_i^S, T_{i+1}^S) \cap [0, \tau]$, and we set $K^S(t) = n$ for all $t \in [T_n^S, \tau]$.

Now let N be a Poisson point process of unit intensity, and let N' be obtained by adding k_N additional points at the origin to N . Since the inter-activation times $X_i^{N'}$ are 0 for $i < k_N$, and distributed exponentially of rate λ_i for $k_N \leq i < n$, we have that the law of $(K^{N'}(t) : t \in [0, \tau])$ is precisely that of the evolution of the non-adaptive process, with k_N pushes at time 0.

We can interpret an adaptive strategy directly in this perspective. For each $s \in \mathbb{R}_+$, it can decide whether to add a new point at position s , based only on $N \cap [0, s]$. In other words, it is a function φ that takes a collection of points S and returns $\varphi(S) \supseteq S$, with $0 \in \varphi(S)$, and where $\varphi(S) \cap [0, t]$ depends only on $S \cap [0, t]$, for any t . The resulting evolution is simply $K^{\varphi(S)}$; the points in $\varphi(S) \setminus S$ correspond to pushes. One easily checks that this always has the correct law of an evolution of an adaptive process, and that any adaptive strategy can be described in this way.

This provides a (somewhat non-obvious, though quite natural) coupling between the evolution of non-adaptive and adaptive. To exploit this, we relax the requirement that an adaptive strategy may only look at the past when making its decisions. We define a *clairvoyant* strategy as *any* function ξ where $\xi(S) \supseteq S$ and $0 \in \xi(S)$. Clearly the optimal clairvoyant strategy has lower cost than the optimal adaptive one.

Lemma A.4.1 There is an optimal clairvoyant strategy which adds points only at the origin.

PROOF. Given a collection of points S , the clairvoyant strategy adds points to obtain a collection $R = \xi(S)$. The cost is then $|R \setminus S| + |\{j \in [n] : T_j^R > \tau\}|$ (the number of pushes plus the number of activation times that are larger than the deadline). Consider any point R_j in $R \setminus S$, and suppose that it is not at the origin. If $j \leq n/2$, moving this point to the origin reduces (or keeps the same) X_i^R for every $i \leq j$, since the λ_i 's are increasing for $i \leq n/2$, and has no effect of X_i^R for $i > j$. Hence the cost can only decrease. If $j > n/2$, simply removing this point can only decrease the cost, this time using that the rates λ_i are decreasing for $i \geq n/2$. \square

So we are comparing the optimal non-adaptive strategy, which picks some number k_N of initial pushes without any knowledge of N , to the optimal clairvoyant strategy, which picks some optimal number of initial pushes based on N . A concentration argument shows that the extra information is very unlikely to be useful. More precisely, we argue that for any number k of initial pushes, the probability that the total cost paid is less than $2k_N - O(\sqrt{n} \text{polylog}(n))$ is polynomially small, and then apply a union bound to conclude that no number k of initial pushes can significantly improve the cost.

Lemma A.4.2 Let $\frac{\log^2 n}{\sqrt{n}} < \tau \leq 2 \log \log n$. Then given a unit intensity Poisson point process N , the expected cost of the optimal clairvoyant strategy is $2k_N - O(\sqrt{n} \text{polylog}(n))$.

The proof can be found at the end of the section.

Consequences for the profit maximization version

Note that the previous result already implies that for the profit maximization version of the problem the adaptivity gap is $1 + o(1)$. Indeed, if $\tau \geq 2 \log \log n$ then $k_N = o(n)$ and thus the profit of non-adaptive is $n - o(n)$ while adaptive certainly gets at most n . On the other hand, if $\tau \leq 2 \log \log n$ but is at least a constant, Lemma A.4.2 implies that the activations that adaptive and non-adaptive make differ by a sub linear term, and since both get a profit which is linear in n the ratio is $1 + o(1)$. Finally, if $\tau = o(1)$ then $k_N = n/2 - o(n)$ so that throughout the rumor spreading process the rates λ_i equals $4/(n + o(n))$ and this cannot be changed by an adaptive strategy.

Remaining proofs

We use the notation $\mathbb{E}_k(\cdot) = \mathbb{E}(\cdot | K(0) = k)$ and similarly $\mathbb{P}_k(\cdot) = \mathbb{P}(\cdot | K(0) = k)$. We will need an estimate of $\mathbb{E}(T_m - T_k)$ throughout this section.

Lemma A.4.3 If $1 \leq k \leq m < n$, then $\mathbb{E}(T_m - T_k) = \log \frac{m(n-k)}{k(n-m)} + O\left(\frac{1}{k} + \frac{1}{n-m}\right)$.

PROOF. Just observe that $\lambda_i^{-1} = \frac{1}{i} + \frac{1}{n-i}$ so

$$\sum_{i=k}^{m-1} \frac{1}{\lambda_i} = H_{m-1} - H_{k-1} + H_{n-k} - H_{n-m} = \log \frac{m(n-k)}{k(n-m)} + O\left(\frac{1}{k} + \frac{1}{n-m}\right).$$

□

We begin with a concentration result for $T_{n/2} - T_k$, the time to activate $n/2$ nodes starting with k .

Lemma A.4.4 Let $\mu := \mathbb{E}(T_{n/2} - T_k)$, where $k \geq \log(n)$. Then for any $0 < \delta \leq \frac{1}{\log(n/k)}$,

$$\begin{aligned} \mathbb{P}_k(T_{n/2} \geq (1 + \delta)\mu) &\leq e^{-(\delta\mu)^2 k/20} \\ \mathbb{P}_k(T_{n/2} \leq (1 - \delta)\mu) &\leq e^{-(\delta\mu)^2 k/20}. \end{aligned}$$

PROOF. This is an application of the Chernoff-Cramér method (see for example [Dubhashi & Panconesi \[2009\]](#)). We show the first bound. For any $0 < \theta < \lambda_k$,

$$\mathbb{P}_k(T_{n/2} \geq (1 + \delta)\mu) \leq e^{-\theta(1+\delta)\mu} \mathbb{E}\left(\exp\left(\theta \sum_j X_j\right)\right) = e^{-\theta(1+\delta)\mu} \prod_j (1 - \theta/\lambda_j)^{-1}.$$

Now using the easily verified inequality $1 - x \geq e^{-x-x^2}$, $\forall x \in [0, 1/2]$, we obtain

$$\begin{aligned} \mathbb{P}_k(T_{n/2} \geq (1 + \delta)\mu) &\leq \exp(-\theta(1 + \delta)\mu + \theta \sum_j \lambda_j^{-1} + \theta^2 \sum_j \lambda_j^{-2}) \\ &= \exp(-\delta\theta\mu + \theta^2 \sum_j \lambda_j^{-2}). \end{aligned}$$

Now we use the fact $\sum_{j=k}^{n/2} \frac{1}{j^2} \leq \frac{2}{k}$ for k sufficiently large to compute

$$\sum_{j=k}^{n/2} \frac{1}{\lambda_j^2} = \sum_{j=k}^{n/2} \left(\frac{1}{j^2} + \frac{1}{(n-j)^2} + \frac{2}{n} \left(\frac{1}{j} + \frac{1}{n-j} \right) \right) \leq \frac{5}{k}.$$

Applying the last inequality, $\mathbb{P}_k(T_{n/2} \geq (1 + \delta)\mu) \leq \exp(-\delta\theta\mu + 5\theta^2/k)$.

Now set $\theta = \frac{\delta\mu k}{10}$ and recall that $\mu = \log(n/k - 1) + O(k^{-1})$. Since $\delta < \frac{1}{\log(n/k)}$ and $k \geq \log n$ it follows that θ respects the conditions, and we have the required bound. The second concentration bound is obtained analogously, but with the inequality $1 + x \geq e^{x-x^2}$, for all $x \geq 0$, instead. \square

Note that by symmetry this result easily yields bounds on $T_i - T_k$, for $i \geq n/2$ and $k \leq n/2$, applying an union bound, with $\mu := \mathbb{E}(T_i - T_k)$ one obtains

$$\begin{aligned} \mathbb{P}(T_i - T_k \leq (1 - \delta)\mu) &\leq 2e^{-(\mu\delta)^2 \min\{k, n-i\}/20} \\ \mathbb{P}(T_i - T_k \geq (1 + \delta)\mu) &\leq 2e^{-(\mu\delta)^2 \min\{k, n-i\}/20} \end{aligned}$$

PROOF OF LEMMA A.4.2. We will apply a union bound over k , the number of initial pushes made by clairvoyant. As such, the dependence between the activation times for different choices of k will not concern us (recall that these are determined by the Poisson point process N as well as k). So with k fixed, we will write just T_i to denote the time that the i 'th node is activated.

Define $r := \frac{n}{1+e^{\tau/2}}$, so by Proposition A.3.4 non-adaptive has cost $2r(1 + o(1))$. Let $\ell := \sqrt{n} \log^2 n$. We first prove that starting with $r - \ell$ nodes implies that at the deadline the final cost is more than $r - \ell$ with probability $1 - n^{-2}$. Consequently, starting with $r - \ell$ nodes guarantees a total cost of at least $c_N - 2\ell$.

Define $T := T_{n-r+\ell} - T_{r-\ell}$, note that in the event $T > \tau$ the total cost is at least $2r - 2\ell$, as we want. Let $\mu := \mathbb{E}(T)$, we now invoke Lemma A.4.4, to obtain

$$\mathbb{P}(T < (1 - \delta)\mu) \leq 2 \exp\left(-\frac{(\mu\delta)^2(r - \ell)}{20}\right) \quad \forall \delta \leq \frac{1}{\log(\frac{n}{r-\ell})}.$$

Note that $\mathbb{E}_r(T_{n-r}) = \tau + O(r^{-1})$, using this fact and a direct computation with the usual bounds for the logarithm, $\frac{x}{x+1} \leq \log(1+x) \leq x$, one shows

$$2\ell \left(\frac{1}{r-\ell} + \frac{1}{n-r} \right) > \mu - \tau > 2\ell \left(\frac{1}{r} + \frac{1}{n-r+\ell} \right)$$

Choose $\delta = \frac{1}{2 \log \log n}$, as $\tau > \frac{2 \log^2 n}{\sqrt{n}}$, with the above bounds one checks $(1 - \delta)\mu > \tau$. An easy algebraic computation shows that

$$(\mu\delta)^2(r - \ell) > \left(\left(\tau + \frac{2\ell}{r} \right) \frac{1}{2 \log \log n} \right)^2 (r - \ell) > \left(\frac{\ell}{r \log \log n} \right)^2 (r - \ell) > \log^2(n).$$

Finally, we conclude $\mathbb{P}(T < \tau) \leq n^{-3}$, so the amount paid at the end is at least $r - \ell$ with probability $1 - n^{-3}$.

Now we wish to prove a similar bound on the probability of paying less than $c_N - 2\ell$ for any other initial number of starting activations. Intuitively, starting with a value different from $r - \ell$ should only make this less likely; clairvoyant needs $n - 2r - 2\ell$ nodes to be randomly activated, and having these be the “middle ones” with the highest rates should be best. More precisely, it can easily be argued that $T_{n-2\ell+k} - T_k$ is stochastically dominated by T for any appropriate choice of k . So the probability that the cost is smaller than $c_N - 2\ell$ when starting with k active nodes is less than $1/n^3$ for any k , and hence the overall probability is at most $1/n^2$.

□

A.5. Bounding the adaptivity gap

We now consider the case where $\tau \geq 2 \log \log n$. Here we need to do more than exploit the concentration of the evolution of the process. If the optimal non-adaptive scheme starts with relatively few pushes at the starting time, there will be a substantial amount of randomness at the start of the process before the epidemic phase transition. Our goal is to show that an adaptive scheme cannot substantially exploit this.

The optimal adaptive strategy is difficult to handle; for example, the optimal choices of t_k are determined via a complicated recurrence. As in the last section, it will be very useful to use a more tractable lower bound; however the lower bound we use here is quite different from the clairvoyant lower bound of the previous section.

As seen in Proposition A.2.2, conditioned on $K^*(t^-) = k$ adaptive does not push at t if $u_k(t) - u_{k+1}(t) < 1$. We consider a modified set of rules for adaptive. Suppose it may push **for free**; however, if there are k active nodes at some time t , it may only push if $k < n/2$ and $u_k(t) - u_{k+1}(t) \geq 1$. Obviously the optimal adaptive strategy satisfies these restrictions, and just pays less under these different rules. So the optimal strategy under these modified rules pays even less. The optimal “modified adaptive” strategy is very simple to describe however: since pushes are free, it will push whenever it is allowed to. We will show that the cost of the optimal modified adaptive strategy, which is simply the expected number of inactive nodes at time τ , is within a constant factor of the cost of non-adaptive.

Let $\tilde{K}(t)$ be the number of active nodes at time t using this optimal modified adaptive strategy. Let $\tilde{T}_i := \min\{t : \tilde{K}(t) \geq i\}$, and let $\tilde{P}(t)$ denote the number of pushes up to and including time t . Observe that $\tilde{K}(0) = k_N$, since k_N is the first k such that $u_k(t) - u_{k+1}(t) < 1$ by Lemma A.3.3.

If one considers the non-adaptive evolution $K(t)$, $u_{K(t)}(t)$ is precisely the Doob martingale for the number of inactive nodes at time τ . It will be very useful to consider a variant of this for the modified adaptive process. So define $\tilde{U}(t) := u_{\tilde{K}(t)}(t)$. This describes the expected number of inactive nodes at time τ , *given that no pushes are made from time*

t until the deadline. It is not a martingale, but rather a supermartingale, since any pushes made by the modified adaptive strategy will decrease the amount paid at the end. Since $\tilde{U}(0) = u_{k_N}(0) = (1 + o(1))k_N$, we will be interested in how much smaller (in expectation) $\tilde{U}(\tau)$ is compared to $\tilde{U}(0)$.

Define $\tilde{t}_k := \inf\{t \in [0, \tau] : u_k(t) - u_{k+1}(t) \geq 1\}$, or $\tilde{t}_k = \tau$ in the case the infimum is taken over the empty set. Then, if $\tilde{K}(t) = k$ and $t < \tilde{t}_k$, the modified adaptive strategy clearly cannot push, so \tilde{t}_k is interpreted as the first time when it is convenient to activate more than k nodes. Since $u_k(t)$ is a strictly increasing function of t , we can equivalently state this as: no push will occur at time t if $\tilde{U}(t^-) < \phi_{\tilde{K}(t^-)}$, where $\phi_k := u_k(\tilde{t}_k)$ for all k . Conversely, if $\tilde{U}(t^-) = \phi_{\tilde{K}(t^-)}$, then the optimal modified adaptive strategy will certainly push. This causes $\tilde{U}(t)$ to jump down by precisely 1 unit.

So we will refer to the values ϕ_k as **thresholds**; the process $\tilde{U}(t)$ is always below the current threshold $\phi_{\tilde{K}(t)}$, and if the threshold is reached, a push will be performed. Also note that by convexity of the sequence $\{u_k(\cdot)\}_k$ the times \tilde{t}_k are increasing, so only a single push occurs at any moment in time.

The following proposition connects the number of pushes made, i.e., the number of times \tilde{U} reaches the current threshold, with the cost saved by the modified adaptive strategy.

Proposition A.5.1 The optimal modified adaptive strategy saves one unit of cost with respect to non-adaptive for each push after $t = 0$, i.e.,

$$\mathbb{E}(n - \tilde{K}(\tau)) = u_{k_N}(0) - \mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0)).$$

PROOF. The crucial point is that $\tilde{U}(t) + \tilde{P}(t)$ is a martingale. This is because between pushes, the evolution is exactly the same as the evolution of the Doob martingale for the non-adaptive process; and if a push occurs at time t , then

$$\tilde{U}(t) + \tilde{P}(t) = \tilde{U}(t^-) - 1 + \tilde{P}(t^-) + 1.$$

Using the optional stopping theorem for martingales, we conclude that

$$\mathbb{E}(\tilde{U}(\tau) + \tilde{P}(\tau)) = \mathbb{E}(\tilde{U}(0) + \tilde{P}(0)) = u_{k_N}(0) + \tilde{P}(0).$$

Since $\tilde{U}(\tau) = n - \tilde{K}(\tau)$, the result follows. \square

We will use this as follows. Suppose $\mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0)) \leq C$. Call the non-adaptive cost $c_N := k_N + u_{k_N}(0)$; from Proposition A.3.4 we know that $c_N = 2(1 + o(1))u_{k_N}(0)$. The adaptivity gap $\rho(\tau)$ is clearly bounded by $c_N/\mathbb{E}(n - \tilde{K}(\tau))$, so applying Proposition A.5.1 we have

$$\rho(\tau) \leq \frac{c_N}{u_{k_N}(0) - \mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0))} = 2(1 + o(1)) \left(1 + \frac{1}{u_{k_N}(0)/C - 1}\right).$$

We also have the trivial upper bound $\rho(\tau) \leq c_N = 2(1 + o(1))k_N$, just because an adaptive strategy will certainly need to push at least once. The required constant bound on $\rho(\tau)$ for the case $\tau \geq 2 \log \log n$ thus follows.

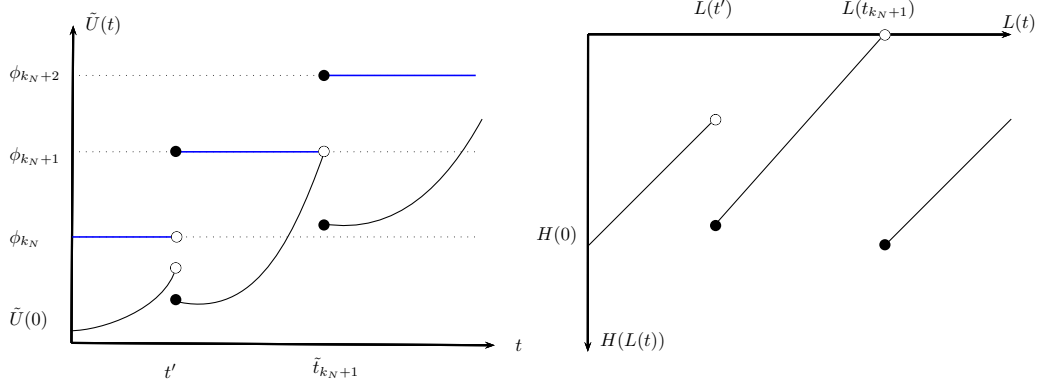


Figura A.1: A sample evolution of \tilde{U} , and the corresponding evolution of the transformed process H . A random activation occurred at time t' , and a push occurred at time \tilde{t}_{k_N+1} .

So the aim for the rest of the section is to bound $\mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0))$ by a constant. To exploit the characterization of modified adaptive we will use some of the estimates on $u_k(t)$ that we derived in Section A.3. Recall Lemma A.3.2, which states that $u_{k+1}(t) = u_k(t)(1 - \frac{1+o(1)}{\lambda_k})$ and $u_k(t+h) \leq u_k(t)e^h$ for $t+h \leq \tau$. This has a very straightforward interpretation in terms of $\tilde{U}(t)$: between activations $\tilde{U}(t)$ grows sub-exponentially, but if at time t there was an activation, then roughly $\tilde{U}(t)$ is multiplied by the factor $1 - \frac{1}{\lambda_{\tilde{K}(t^-)}}$.

Now we have all the ingredients to bound how many times the process $\tilde{U}(t)$ hits the thresholds $\phi_{\tilde{K}(t)}$, which is exactly the number of pushes. We proceed by transforming the process in a number of ways.

- First, given the (sub)exponential growth, it makes sense to take logarithms, yielding a process that between jumps grows no faster than a linear function with unit slope.
- Secondly, we locally shift the resulting process so that the threshold at any moment of time is moved to zero. So the process will always be non-positive, and we will be interested in the number of times that it hits zero.
- Finally, we locally rescale time, in precisely the same way as in the previous section, so that the distribution of the times of random activations are described by a Poisson point process of unit intensity. We locally rescale the value at the same time, so that still the process increases linearly at unit rate in between jumps.

Formally, we define the following transformed process $H(s)$:

$$H(L(t)) := \lambda_{\tilde{K}(t)} (\log \tilde{U}(t) - \log \phi_{\tilde{K}(t)}), \quad \text{where } L(t) := \int_0^t \lambda_{\tilde{K}(x)} dx.$$

An illustration of the evolution of \tilde{U} and the corresponding transformed process H is shown in Figure A.1. Notice how upon each random activation or push the threshold increases, while \tilde{U} jumps down.

We are interested in the number of times that H reaches 0, as this corresponds to the

number of pushes after time 0. It is convenient to consider instead

$$H'(L(t)) := H(L(t)) + \tilde{P}(t) - \tilde{P}(0).$$

The process H jumps down immediately whenever it reaches 0, and in fact the size of this jump is larger than 1. Very roughly speaking, H' cancels out these jumps (actually it may still jump down, but by a smaller amount), while the jumps corresponding to random activations are unaffected. A bound on the number of pushes is very easily read off of H' , as shown in the following lemma.

Lemma A.5.2 For every $t \in [0, \tau]$, the number of pushes $\tilde{P}(t) - \tilde{P}(0)$ is at most $\max\{0, 1 + \sup_{0 \leq r \leq t} H'(L(r))\}$.

PROOF. Let θ be the last time a push was performed in $[0, t]$; thus $\tilde{P}(t) = \tilde{P}(\theta)$. If $\theta = 0$ the Lemma clearly holds, otherwise there was at least one push and $\tilde{P}(\theta^-) = \tilde{P}(t) - 1$. Notice that by definition $H(L(\theta^-)) = 0$, because H always reaches zero before pushing. Now compute,

$$\lim_{r \nearrow \theta} H'(L(r)) = \lim_{r \nearrow \theta} (H(L(r)) + \tilde{P}(r) - \tilde{P}(0)) = \tilde{P}(t) - 1 - \tilde{P}(0).$$

Obviously $\sup_{0 \leq r \leq t} H'(L(r)) \geq \lim_{r \nearrow \theta} H'(L(r))$, so the proof is complete. \square

So all that remains is to bound the expected supremum of H' . The reason that this is possible is very simple; through most of its evolution, the process has a negative drift. More precisely, we have the following. Here and for the remainder of this section, N will denote a Poisson process of unit intensity.

Proposition A.5.3 For any constant $0 < c \leq 1/2$, there is a $c' = (1 + o(1))c$ so that for any $\sigma \leq L(\tilde{T}_{cn})$, $H'(s + \sigma) - H'(\sigma)$ is stochastically dominated by $s - 2(1 - c')N(s)$ on the interval $[0, L(\tilde{T}_{cn}) - \sigma]$.

The proof is presented at the end of the section. Note in particular that $s - 2(1 - c)N(s)$ has negative drift for any $c < 1/2$. Primarily this entails converting our estimates on the behaviour of \tilde{U} into estimates on H .

Unfortunately, the bound we obtain on the drift gets worse as the number of active nodes increases. So a small amount of care is required. We give a sketch of the argument, and postpone the details for the end of the section. First, on the interval $[0, L(\tilde{T}_{n/4})]$, $H'(s)$ is dominated by $s - \frac{3}{2}N(s)$. This process starts from 0 and has negative drift (since $\mathbb{E}[s - \frac{3}{2}N(s)] = -s/2$). Proving that its expected maximum value is constant reduces to a straightforward concentration bound. Since $\tau \geq 2 \log \log n$, so that $k_N = O(n/\log n)$, this negative drift also implies that the process will be very negative (below say $-n/64$) at time $L(\tilde{T}_{n/4})$, with very high probability. On the interval $[L(\tilde{T}_{n/4}), L(\tilde{T}_{n/2})]$, $H'(s)$ can be dominated by $s - N(s)$ conditioned on $N(L(\tilde{T}_{n/4})) = H'(L(\tilde{T}_{n/4})) \leq -n/64$. Again a concentration argument shows that with very high probability the process remains negative in this interval, and so the expected number of further pushes is again at most a constant.

Remaining proofs

Before proving Proposition A.5.3 we need the following intermediate result.

Lemma A.5.4 For all $k < n/2$,

$$\frac{\phi_k}{\phi_{k+1}} \leq 1 - \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}}(1 + o(1)).$$

PROOF. Recall that \tilde{t}_k is such that $u_k(\tilde{t}_k) = u_{k+1}(\tilde{t}_k) + 1$ and by construction $\phi_k = u_k(\tilde{t}_k)$. By Lemma A.3.2 $u_{k+1}(t) = (1 - \frac{1+o(1)}{\lambda_k})u_k(t)$, replacing the identity $u_k(\tilde{t}_k) = u_{k+1}(\tilde{t}_k) + 1$ one gets $\phi_k = \lambda_k(1 + o(1))$, finally

$$\frac{\phi_k}{\phi_{k+1}} = \frac{\lambda_k}{\lambda_{k+1}}(1 + o(1)) = 1 - \frac{\lambda_{k+1} - \lambda_k(1 + o(1))}{\lambda_{k+1}}.$$

□

PROOF OF PROPOSITION A.5.3. Note that the map $L(t)$ is a bijection and the processes involved are Markovian. Thus a comparison of the infinitesimal increments of the two processes is sufficient.

Recall from the previous section that the times of new opportunistic activations can be coupled to the points of N ; if the i 'th random activation (not push) occurs at time t , then $L(t) = N_i$.

Condition on the event $\tilde{K}(t) = k$, by construction $t \leq \tilde{t}_k$, let us assume $t < \tilde{t}_k$ and take a small Δt so there will be no pushes in the interval $[t, t + \Delta]$ and $H'(s)$ evolves just as $H(s)$ does. Set $s = L(t)$, notice that $L(t + \Delta t) = L(t) + \lambda_k \Delta t$, so when rescaling time Δt yields $\Delta s = \lambda_k \Delta t$. The time until the next activation is $\text{Exp}(\lambda_k)$, then with probability $1 - \lambda_k \Delta t + o(\Delta t)$ there was not an activation, thus thanks to Lemma A.3.2 $\tilde{U}(t + \Delta t) \leq \tilde{U}(t)e^{\Delta t}$, $H(s)$ changes according to

$$H(L(t + \Delta t)) = \lambda_k \log \left(\frac{\tilde{U}(t + \Delta t)}{\phi_k} \right) \leq \lambda_k \log \left(\frac{\tilde{U}(t)e^{\Delta s/\lambda_k}}{\phi_k} \right) = H(s) + \Delta s.$$

With probability $\lambda_k \Delta t + o(\Delta t)$ there was an activation, and thanks to Lemma A.3.2 $\tilde{U}(t + \Delta t) \leq \tilde{U}(t)e^{\Delta t}(1 - \frac{1+o(1)}{\lambda_k})$, hence the process changes according to

$$\begin{aligned} H(L(t + \Delta t)) &= \lambda_{k+1} \log \left(\frac{\tilde{U}(t + \Delta t)}{\phi_{k+1}} \right) \\ &\leq \lambda_{k+1} \log \left(\frac{\tilde{U}(t)e^{\Delta s/\lambda_k}(1 - \frac{1+o(1)}{\lambda_k})}{\phi_{k+1}} \right) \\ &= \lambda_{k+1} \log \left(\frac{\tilde{U}(t)}{\phi_k} \right) + \lambda_{k+1} \left(\frac{\Delta s}{\lambda_k} + \log \left(1 - \frac{1 + o(1)}{\lambda_k} \right) + \log \left(\frac{\phi_k}{\phi_{k+1}} \right) \right). \end{aligned}$$

By construction $\log\left(\frac{\tilde{U}(t)}{\phi_k}\right) \leq 0$, using $\lambda_{k+1} > \lambda_k$ it follows $\lambda_{k+1} \log\left(\frac{\tilde{U}(t)}{\phi_k}\right) \leq H(s)$. To bound the other terms, by Lemma A.5.4 $\frac{\phi_k}{\phi_{k+1}} \leq 1 - \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}}(1 + o(1))$, recalling the well known inequality $\forall x > -1, \log(1 - x) \leq -x$,

$$\log\left(1 - \frac{1 + o(1)}{\lambda_k}\right) + \log\left(\frac{\phi_k}{\phi_{k+1}}\right) \leq -\frac{1 + o(1)}{\lambda_k} - \frac{\lambda_{k+1} - \lambda_k}{\lambda_{k+1}}(1 + o(1)).$$

Finally, when there is one activation, as $k \leq cn$

$$H(s) \leq \Delta s - \left(\frac{\lambda_{k+1}}{\lambda_k} + \lambda_{k+1} - \lambda_k\right)(1 + o(1)) \leq \Delta s - 2(1 - c).$$

By inspection we find the process to be dominated by $s - 2(1 - c)N(s)$. We are done with the case $t < \tilde{t}_k$, now the case when $t = \tilde{t}_k$ is easy because a push was made in such event, with the same calculation as before $H(s) \leq H(s^-) - 2(1 - c)$, thus $H'(s) \leq H'(s^-) - 2(1 - c) + 1 < H'(s^-)$. The process $s - 2(1 - c)N(s)$ clearly dominates, because the latter changed zero just in time t . \square

We prove a number of lemmas, all based on simple concentration arguments.

Lemma A.5.5 Let Z_1, \dots, Z_m be i.i.d. exponential random variables of unit rate. Then

$$\mathbb{P}\left(\sum_{i=1}^m Z_i > (1 + \delta)m\right) \leq \left(\frac{1 + \delta}{e^\delta}\right)^m.$$

PROOF. A standard application of the Chernoff-Cramèr method yields that for any $\delta > 0$,

$$\begin{aligned} \mathbb{P}\left(\sum_{j=1}^m Z_j > (1 + \delta)m\right) &\leq \inf_{0 \leq \theta < 1} e^{-\theta(1+\delta)m}(1 - \theta)^{-m} \\ &\leq \left(\frac{1 + \delta}{e^\delta}\right)^m. \end{aligned}$$

\square

Lemma A.5.6 For any $c > 1$, there exists a D s.t.

$$\mathbb{P}(\sup_{s \geq 0} \{s - cN(s)\} \geq \gamma) \leq De^{-(1-1/c)\gamma}$$

for all $\gamma > 0$, and hence

$$\mathbb{E}(\sup_{s \geq 0} \{s - cN(s)\}) = O(1).$$

PROOF. Let $\Psi = \sup_{s \geq 0} \{s - cN(s)\}$. Write $N_i = \sum_{j=1}^i Z_j$, where $Z_i \sim \text{Exp}(1)$. Then

$$\mathbb{P}(\Psi \geq \gamma) = \mathbb{P}\left(\sum_{j=1}^m Z_j \geq cm + \gamma \text{ for some } m \geq 1\right).$$

Applying Lemma A.5.5 with $\delta = c - 1 + \gamma/m$,

$$\begin{aligned} \mathbb{P}\left(\sum_{j=1}^m Z_j \geq cm + \gamma\right) &\leq e^{-\gamma} \left(\frac{c}{e^{c-1}}\right)^m \cdot \left(1 + \frac{\gamma}{cm}\right)^m \\ &\leq e^{-(1-1/c)\gamma} \kappa^m, \end{aligned}$$

where $\kappa = \frac{c}{e^{c-1}} < 1$. A union bound over m now yields

$$\mathbb{P}\left(\sum_{j=1}^m Z_j \geq cm + \gamma \text{ for some } m \geq 1\right) \leq e^{-(1-1/c)\gamma} \sum_{m=1}^{\infty} \kappa^m = e^{-(1-1/c)\gamma} \frac{\kappa}{1-\kappa},$$

proving the first part of the lemma. We also obtain

$$\begin{aligned} \mathbb{E}\Psi &= \int_0^{\infty} \mathbb{P}(\Psi > \gamma) d\gamma \\ &\leq \frac{\kappa}{1-\kappa} \int_0^{\infty} e^{-(1-1/c)y} dy \\ &= O(1). \end{aligned}$$

□

Lemma A.5.7 For any $c > 1$, there exists a constant $D > 0$ s.t.

$$\mathbb{P}(s - cN(s) \geq -\frac{1}{2}(c-1)s) \leq e^{-Ds} \quad \text{for all } s \geq 0.$$

PROOF. Write $N_i = \sum_{j=1}^i Z_j$, where $Z_i \sim \text{Exp}(1)$. Simply rewrite the stated probability as

$$\mathbb{P}(N(s) \leq \frac{s}{2}(1+1/c)) = \mathbb{P}\left(\sum_{i=1}^{\lfloor \frac{s}{2}(1+1/c) \rfloor} Z_i > s\right),$$

and then apply Lemma A.5.5 with $\delta = \frac{c-1}{c+1}$. □

Lemma A.5.8 The event that $L(\tilde{T}_{n/4})$ is smaller than $n/16$ has exponentially small probability.

PROOF. There are only $k_N = o(1)$ pushes at time 0, and the number of pushes after time 0 by the time $n/4$ nodes are active is also $o(n)$ with exponentially high probability by Lemma A.5.6 and Proposition A.5.3. So $L(\tilde{T}_{n/4}) \geq N_{n/8}$ with exponentially high probability. But $N_{n/8} \geq n/16$ with exponentially high probability, directly from Lemma A.5.5; combining all of these, we get the lemma. □

We are now ready to deduce our required bound on the maximum of $H'(s)$.

Lemma A.5.9 $\mathbb{E}(\sup_{s \leq L(\tilde{T}_n)} H'(s)) = O(1)$.

PROOF. Combining Proposition A.5.3, Lemma A.5.7 and Lemma A.5.8, we deduce that $H'(L(\tilde{T}_{n/4})) \leq -n/64$ with exponentially high probability. Thus we can dominate $H'(s + L(\tilde{T}_{n/4}))$ by the process $s - cN(s) - n/64$, where $c = 1 + o(1)$, with exponentially high probability. Moreover $L(\tilde{T}_{n/2}) - L(\tilde{T}_{n/4}) \leq n/2$ with exponentially high probability. Now observe that

$$\mathbb{P}\left(\sup_{s \in [0, n/2]} s - cN(s) - n/64 \geq 0\right) \leq \mathbb{P}\left(\sup_{s \in [0, n/2]} s - cN(s) - s/64 \geq n/128\right).$$

Thus a final application of Lemma A.5.6 with $c = 64/63$ tells us that with exponentially high probability, $\sup_{s \in [L(\tilde{T}_{n/4}), L(\tilde{T}_{n/2})]} H'(s) < 0$. This completes the proof. \square

A.6. Final comments

More complex variants of our base homogeneous model will not exhibit the $O(1)$ adaptivity gap this work's main contribution establishes. As our next result shows, this happens even under mild conditions on the distinct feasible rates at which network nodes encounter each other.

Proposition A.6.1 There is an infinite family of inhomogeneous networks whose adaptivity gap grows arbitrarily large with network size, even when the rates take only two distinct values, one zero and the other one positive.

PROOF. Consider a network consisting of a *root* v , *intermediate* nodes v_1, \dots, v_k , and *leaf* nodes $v_{i,1}, \dots, v_{i,k}$ for each $i \in [k]$. The time between consecutive encounters between nodes v and v_i for $i \in [k]$ and between v_i and $v_{i,j}$ for $i, j \in [k]$, will all be independent identically distributed $\text{Exp}(1/k)$ random variables. All other node pairs never encounter each other (equivalently, they encounter each other at rate 0). Let $\tau := k \log(k \log k)$ and consider the adaptive scheme that activates the root at time $t = 0$, at time $\tilde{\tau} := k \log \log k$ activates all the intermediate nodes that are not active at time $\tilde{\tau}^-$, and otherwise the process evolves on its own. If v_i becomes active at time $t \leq \tilde{\tau}$, then each of its children will be inactive at τ^- with probability $e^{-(\tau-t)/k}$. Some basic calculations show that the expected number of v_i 's inactive children at time τ^- is $(k + \tilde{\tau})e^{-\tau/k} = (1 + \log \log k)/\log k$. The expected number of intermediate nodes that will remain inactive at time $\tilde{\tau}^-$ is easily seen to be $ke^{-\tilde{\tau}/k} = k/\log k$. Thus, the expected cost of the adaptive scheme is $(1+o(1))k \log \log k / \log k$. Now, let's consider a non-adaptive scheme that at time $t = 0$ activates $m \in [k]$ intermediate nodes. To lower bound the optimal non-adaptive cost, we can assume that the root becomes active via opportunistic communication (at no cost) at time $t = 0$. Similar calculations as the ones required in the analysis of the adaptive scheme yield that each of the $k - m$ initially inactive intermediate nodes will have $(\tau + k)e^{-\tau/k} \geq \tau/(k \log k) \geq 1$ expected number of inactive children at time τ^- . Hence, the expected cost of the non-adaptive scheme will be at least $(k - m) + m = k$. An adaptivity gap of $(1 - o(1)) \log k / \log \log k$ follows. \square

A slight variation of this example shows that if we allow rates to be different on different edges, the adaptivity gap can be as large as $\Omega(\sqrt{n})$.

A natural generalization of our problem is a setting where only a specified subset of nodes is required to be activated by the deadline; but where all nodes spread the rumor. Our main result holds in this setting as well.

Proposition A.6.2 The target set problem has a constant adaptivity gap.

It is possible that the expected *difference* between the costs of the optimal adaptive and non-adaptive strategies is always bounded by a constant. We leave this as an interesting open question.

We believe that it is worthwhile to elicit under which hypotheses (weaker than all rates being equal), the adaptivity gap is still $O(1)$. More generally, it would be of interest to understand if and how model enhancements such as network heterogeneity, fairness among user constraints, users' selfish behavior, and adversarial scenarios impact the adaptivity gap of opportunistic communication networks. In particular we expect that our results capture well situations where the underlying graph has good expansion properties (high conductance) and therefore we conjecture that our constant adaptivity gap result holds for these graphs as well.