Universidad de Valladolid

Facultad de Ciencias

TRABAJO FIN DE GRADO

Grado en Matemáticas.

Códigos de Red Lineales.

Autora: Esther Fernández Sanz

Tutora: Carolina Ana Núñez Jiménez

Índice general

Resumen

En el problema de transmisión de información a través de una red, en la que hay varios receptores, aparece el problema de los "cuellos de botella", cuando por un conjunto de aristas se quiere transmitir más información de la que admiten, generando un retraso en el sistema. Esto se puede solucionar codificando la información que pasa por dichas aristas, lo que se llama códigos de red (network coding). Cuando esta codificación es lineal en función de los mensajes, hablaremos de códigos en red lineales. En este trabajo introduciremos el problema, daremos una solución a la codificación y estudiaremos la probabilidad de éxito cuando los coeficientes de eligen de forma aleatoria.

Introducción

Un proceso de comunicación, en el más puro sentido de la palabra, siempre se ha compuesto en términos generales de emisor, receptor, un canal, un alfabeto conocido por ambos, y por supuesto, un mensaje que compartir.

En un proceso de este tipo se plantean distintos tipos de problemas, como la confidencialidad (criptografía), la fiabilidad (códigos correctores), y la eficacia de la comunicación.

Respecto a la eficacia, un punto de vista es el de los códigos compresores, que consiguen reducir la cantidad de información a transmitir. Pero hay otra perspectiva muy relacionada con las comunicaciones tal como se realizan en la actualidad, en la que el canal de comunicación es una red y hay varios receptores de la información.

La situación se puede modelar como sigue: se tiene un grafo dirigido (que supondremos simple y sin ciclos), en el cual parte de las redes son los emisores de la información y parte son los receptores. Se quiere que la información se transmita de la forma más eficiente posible, en términos de tiempo y coste computacional.

Si hay un único receptor, la solución óptima se encuentra si en la red se pueden encontrar caminos disjuntos desde los emisores hasta el receptor, pues entonces cada mensaje viaja por el camino correspondiente sin retrasos. Pero si hay más de un receptor, aún existiendo tales caminos disjuntos para cada uno de ellos, se puede producir en algunas aristas un "cuello de botella" por el que debe pasar la información distinta para cada receptor, en cuyo caso se produce un retraso.

Un ejemplo sobre qué tipo de problemas veremos a lo largo del trabajo es el que vemos en la figura 1. En él, el emisor s desea enviar el mensaje a y el mensaje b tanto al receptor r_1 como al r_2.

Por separado, el emisor podrá enviar sin problema ambos mensajes a cada receptor como vemos en la figura 2, a los caminos marcados, los llamamos flujos.

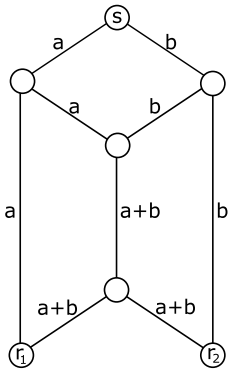


Figura 1: ejemplo de un problema de codificación en red.

No obstante, si queremos enviar esa información simultáneamente, efectivamente, en la arista central nos encontramos con mensajes distintos intentando pasar a la vez.

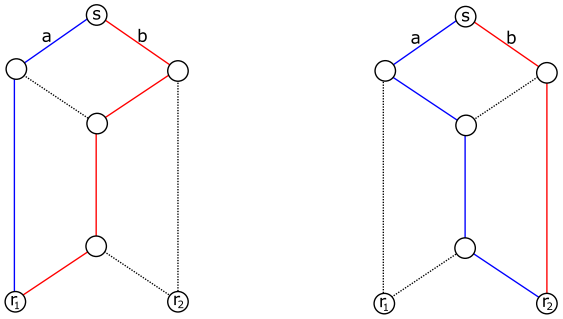


Figura 2: caminos para llevar la información a cada receptor

De esta forma, a r_1 llega a y $a+b$ y puede recuperar el mensaje original, al igual que r_2 recibe b y $a+b$.

La idea consiste por tanto en transmitir por cada arista no uno de los mensajes originales, sino una función de los mismos. Este tipo de codificación se llama **codificación en red** (network coding). En el caso en que las funciones de codificación son lineales se habla de **códigos de red lineales** (linear network coding). En los problemas de códigos de red lineales se plantea el problema de cómo escoger las funciones de codificación. Una posibilidad es escoger los coeficientes de dicha combinación lineal de forma aleatoria, y entonces se habla de **códigos en red lineales aleatorios** (random linear network coding). En este caso es preciso estudiar la probabilidad de éxito de

la transmisión.

Este trabajo de fin de grado ha consistido en el estudio del artículo de Olav Geil y Casper Thomsen *Aspects of random network coding*, [7], en el que se hace una presentación del problema y se estudia cuándo y cómo éste tiene solución, y la probabilidad de éxito en los procesos aleatorios. Para su comprensión ha sido necesario un estudio preciso de los problemas de flujos en redes, así como de la cota de la huella (footprint bound) de un ideal, que está relacionado con el uso de bases de Gröbner del mismo.

Antes de comenzar, establezcamos de forma esquemática el contenido de la memoria:

En el capítulo 1 haremos un pequeño repaso sobre los conceptos necesarios de ideales y órdenes monomiales, para poder hacer la demostración de la cota de la huella y obtener como conclusión poder estudiar el número de ceros en un polinomio.

A partir de la teoría de grafos, en el capítulo 2 veremos la definición de flujo en una red, y cómo optimizarlo. Veremos también qué es un flujo entero y qué relación tiene con los caminos mutuamente disjuntos.

A lo largo del capítulo 3 encontraremos el núcleo central del trabajo, en el qué definiremos el problema de codificación en red, particularizando en la codificación lineal sobre un cuerpo finito. Finalmente veremos un algoritmo por el cual encontrar coeficientes adecuados para que el problema tenga solución.

El capítulo 4, que corresponde a la última parte del trabajo, se dan cotas a la probabilidad de tener éxito asignando los coeficientes en la codificación lineal de forma aleatoria.

Capítulo 1

Huella de un ideal

Dado un cuerpo K, consideramos el anillo de polinomios $K[X_1, ..., X_m]$ en m indeterminadas. Para simplificar la notación, denotaremos dicho anillo como $K[X]$.

Dado $\alpha = (\alpha_1, ..., \alpha_m) \in \mathbb{Z}_{\geq 0}^m$, denotaremos por X^α al **monomio** $X_1^{\alpha_1}...X_m^{\alpha_m}$, y dado $h \in K$ diremos que hX^α es un **término** de $K[X]$.

Dado un orden $\prec$ en $\mathbb{Z}_{\geq 0}^m$, podemos asociarle un orden, que denotaremos igual, en el conjunto de monomios de $K[X]$, por

$$X^\alpha \prec X^\beta \Leftrightarrow \alpha \prec \beta.$$

Definición 1.1. Un **orden monomial** sobre $K[X]$ es cualquier relación binaria $\prec$ sobre $\mathbb{Z}_{\geq 0}^m$, o equivalentemente, cualquier relación sobre el conjunto de los monomios de $K[X]$, que cumple:

1. $\prec$ es un *buen orden* sobre $\mathbb{Z}_{\geq 0}^m$
2. Si $\alpha \prec \beta$, entonces $\alpha + \gamma \prec \beta + \gamma$, $\forall \gamma \in \mathbb{Z}_{\geq 0}^m$.

Fijaremos un orden monomial $\prec$ sobre $K[X]$ de aquí en adelante.

Definición 1.2. Si consideramos un polinomio $f = \sum_{\alpha \in \Gamma} a_\alpha X^\alpha \in K[X]$, con $\Gamma \subset \mathbb{Z}_{\geq 0}^n$ finito y $f \neq 0$, definimos:

- El **multigrado** de f es

$$multideg(f) = \text{máx}_\prec \{\alpha \in \Gamma \ / \ a_\alpha \neq 0\}$$

- El **coeficiente dominante** de f es

$$LC(f) = a_{multideg(f)}$$

- El **monomio dominante** de f es

$$LM(f) = X^{multideg(f)}$$

- El **término dominante** de f es

$$LT(f) = LC(f) \cdot LM(f)$$

Dado un ideal $I \subset K[X]$ denotamos

$$LT(I) = (\{LT(f) \ / \ f \in I - \{0\}\}),$$

ideal generado por los términos dominantes de los elementos del ideal I.

Para cualquier polinomio $f \in K[X]$, puesto que un orden monomial es total, se pueden ordenar los términos de f de acuerdo a sus exponentes; f se escribe de forma única ordenando los monomios de mayor a menor.

Teorema 1.3 (Teorema de división)**.** *Sea $\prec$ un orden monomial y sea $F = \{f_1, ..., f_s\}$ un conjunto ordenado de polinomios de $K[X]$. Entonces para cada $f \in K[X]$ existen $a_1, ..., a_s, r \in K[X]$ que cumplen:*

1. $f = a_1 f_1 + ... + a_s f_s + r$
2. *Si $r \neq 0$, entonces ninguno de los términos de r es divisible por ninguno de los términos $LT(f_1), ..., LT(f_s)$.*

O dicho de otra forma, el polinomio r será combinación lineal de monomios que no pueden ser divisibles por ninguno de los $LT(f_1), ..., LT(f_s)$.

El teorema se demuestra a partir del algoritmo de división, el cual nos proporciona una forma de escribir $f = a_1 f_1 + ... + a_s f_s + r$. A r, queda determinado por el algoritmo y el orden finado en F, lo llamamos resto de la división de f por F, y lo denotamos por $r = \bar{f}^F$.

Tanto el algoritmo, la demostración completa de este teorema, como las demostraciones correspondientes a los resultados de la siguiente sección se encuentran en el libro [3].

1.1. Ideales monomiales

Definición 1.4. Un ideal I de $K[X]$ es **monomial** si tiene un conjunto de generadores formado por monomios, esto es, si existe un subconjunto $A \subset \mathbb{Z}^n_{\geq 0}$ tal que

$$I = (\{X^\alpha \ / \ \alpha \in A\})$$

Definición 1.5. Un conjunto finito de polinomios de un ideal I, $G = \{g_1, .., g_t\}$ es una **base de Gröbner** de I si

$$LT(I) = (LT(g_1), ..., LT(g_t)).$$

Proposición 1.6. Si G es una base de Gröbner de un ideal $I \subset K[X]$, entonces G es un sistema de generadores de I.

Teorema 1.7. *Dado un ideal $I \neq 0$ de $K[X]$, existe una base de Gröbner de I para el orden monomial dado.*

Como consecuencia del teorema de división tenemos el siguiente resultado:

Proposición 1.8. Sea $G = \{g_1, ..., g_t\}$ una base de Gröbner de un ideal $I \subset K[X]$, y sea f un polinomio de $K[X]$. Entonces existe un único $r \in K[X]$ que cumple las siguientes propiedades:

1. Ningún término de r es divisible por ninguno de los $LT(g_1), ..., LT(g_t)$, siempre que $r \neq 0$.
2. $f - r \in I$.

En particular, r se puede calcular, haciendo uso del algoritmo de división, como el $r = \bar{f}^G$, independientemente del orden de G.

1.2. Huella de un ideal

Definición 1.9. Dados un ideal $I \subseteq K[X]$ y un orden monomial $\prec$, se define la **huella** de I respecto del orden $\prec$ como el conjunto de monomios que no son términos dominantes de ningún polinomio en I. Se denota por $\Delta_\prec(I)$.

$$\Delta_\prec(I) = \{X^\alpha \ / \ X^\alpha \neq LT(f) \ \ \forall f \in I\}$$

Proposición 1.10. Sea $I \subset K[X]$ un ideal. Entonces $K[X]/I$ es isomorfo como K-espacio vectorial a

$$S = \mathbb{L}(\{X^\alpha / \ X^\alpha \notin LT(I)\}) = \mathbb{L}(\{X^\alpha / \ X^\alpha \in \Delta_\prec(I)\})$$

Demostración. Se considera $\Phi : \ S \longrightarrow K[X]/I$, donde Φ es la aplicación del paso al cociente de $K[X]$ a $K[X]/I$, que es K-lineal.

Sea $G = \{g_1, ..., g_t\}$ una base de Gröbner de I, es decir, tenemos $LT(I) = (LT(g_1), ..., LT(g_t))$. Hay que tener en cuenta que, por la proposición 1.8,

$\bar{f}^G \in \Delta_\prec(I)$, y $f - \bar{f}^G \in I$ para todo $f \in K[X]$, donde $\bar{f}^G$ es único con estas propiedades.

Dado $f \in K[X]$, se tiene que $\Phi(\bar{f}^G) = \bar{f}^G + I = f + I$, por lo tanto Φ es sobre.

Por la unicidad de $\bar{f}^G$, se tiene que si $f \in \Delta_\prec(I)$ entonces $\bar{f}^G = f$, y que si $f = \sum \lambda_i f_i \in S$ con $f_i \in \Delta_\prec(I)$, entonces

$$\bar{f}^G = \sum \lambda_i \bar{f_i}^G = \sum \lambda_i f_i = f.$$

Además, $S \cap I = \{0\}$. Por tanto si $f \in S$ y $\Phi(f) = 0$ entonces $\bar{f}^G = f \in I$ y tendremos que $\bar{f}^G = f = 0$.

De este modo vemos que efectivamente Φ es un isomorfismo, y que tendremos entonces

$$S \cong K[X]/I.$$

□

En lo que sigue, K será un cuerpo algebraicamente cerrado.

Definición 1.11. Se llama **variedad afín** definida por el ideal $I \subset K[X]$ al conjunto

$$\mathcal{V}_K(I) = \{x \in K^m \ / \ f(x) = 0 \ \ \forall f \in I\}.$$

Teorema 1.12 (Teorema de los ceros de Hilbert)**.** *Sea I un ideal de $K[X]$. Entonces $I(V_K(I)) = \{f \in K[X]/f(x) = 0 \ \forall x \in V_k(I)\} = Rad(I)$.*

Una demostración detallada del teorema se puede encontrar en [5].

Teorema 1.13. *Sea $V = \mathcal{V}_K(I)$ una variedad afín en K^m. Entonces las siguientes afirmaciones son equivalentes:*

i) V es un conjunto finito

ii) El K-espacio vectorial $K[X]/I$ es de dimensión finita.

iii) $\Delta_\prec(I)$ es finita.

Demostración. La equivalencia de *ii*) y *iii*) es consecuencia directa de la proposición 1.10.

Supongamos que $\dim_K K[X]/I < \infty$. Probaremos que V es finito probando que existen conjuntos finitos $A_i \subset K$, para $i = 1, ..., m$, tales que $V \subset A_1 \times ... \times A_m$.

Si fijamos un $i \in \{1, ..., m\}$, entonces al considerar el conjunto de las clases de X_i^j,

$$\{X_i^j + I, j \geq 0\} \subset K[X]/I,$$

tenemos que el primer conjunto tiene infinitos elementos. Como suponemos que $K[X]/I$ es de dimensión finita, los X_i^j deben ser linealmente dependientes en $K[X]/I$, esto es, existirán un $r > 0$ y $\lambda_0, ..., \lambda_r \in K$ tales que

$$F_i(X_i) = \lambda_0 + \lambda_1 X_i + \lambda_2 X_i^2 + ... + \lambda_r X_i^r \in I$$

con no todos los λ_i iguales a 0. En particular, el polinomio $F_i(X_i)$ tiene un número finito de raíces en K. Denotamos por A_i al conjunto de sus raíces. Entonces

$$(a_1, ..., a_m) \in V \Rightarrow F_i(a_1, ..., a_m) = 0 \Rightarrow F_i(a_i) = 0 \Rightarrow a_i \in A_i$$

y por tanto $V \subset A_1 \times ... \times A_m$, como queríamos probar.

Recíprocamente, supongamos que V es un conjunto finito, $V = \{p_1, ..., p_t\}$. Fijemos un $i \in \{1, ..., t\}$ y sea a_j la i-ésima coordenada del punto p_j. Definimos el polinomio

$$f_i(X_i) = \prod_{k=1}^{t} (X_i - a_k) \ \in K[X_i],$$

el cual se anula en cada punto de V, y por lo tanto $f_i \in I(V)$. Por el Teorema de los Ceros de Hilbert, existe algún $n \geq 1$ tal que $f_i^n \in I$, y por consiguiente, $X_i^{tn} = LT(f_i) \in LT(I)$. Este argumento se aplica a cada coordenada y tendremos que existen $r_1, ..., r_m$ tales que $X_1^{r_1} \in LT(I)$, ..., $X_m^{r_m} \in LT(I)$.

De esta forma, para $\alpha = (\alpha_1, ..., \alpha_m)$ se tiene que si $X^\alpha \notin LT(I)$, entonces $\alpha_i \leq r_i - 1$, de donde deducimos que$\Delta_\prec(I)$ es finito.

□

Proposición 1.14. Sea $I \subset K[X]$ un ideal tal que $\mathcal{V}_K(I)$ es un conjunto finito. Entonces, el número de puntos de $\mathcal{V}_K(I)$ es como mucho la dimensión de $K[X]/I$ como K-espacio vectorial.

Demostración. Sea $\mathcal{V}_K(I) = \{p_1, ..., p_t\}$ con $p_i \neq p_j$ si $i \neq j$. Vamos a construir polinomios $f_1, ..., f_t \in K[X]$ tales que $f_i(p_j) = \delta_{ij}$. Haremos la construcción de f_1, siendo análogo para el resto.

Escribamos $p_j = (p_{j,1}, ..., p_{j,m})$. Para cada $j > 1$ tenemos que $p_j \neq p_1$, luego, existe t_j tal que $p_{j,t_j} \neq p_{1,t_j}$. Para

$$h_j = \frac{X_{t_j} - p_{j,t_j}}{p_{1,t_j} - p_{j,t_j}} \in K[X],$$

se tiene que $h_j(p_j) = 0$ y que $h_j(p_1) = 1$.

El polinomio $f_1 = \prod_{j \neq 1} h_j$ cumple que $f_1(p_1) = 1$ y que $f_1(p_j) = 0$ para $j > 1$.

A continuación probaremos que $f_1 + I$,...,$f_t + I$ son linealmente independientes, con lo que se concluye la prueba.

Sean $a_1, ..., a_t \in K$ tales que

$$\sum a_i(f_i + I) = 0,$$

es decir $\sum a_i f_i \in I$. Para cada $j \in \{1, ..., t\}$, tendremos que $(\sum a_i f_i)(p_j) = 0$. Por construcción, $(\sum a_i f_i)(p_j) = a_j$ y por lo tanto cada $a_i = 0$, y f_1+I,...,f_t+I son linealmente independientes. Luego

$$|\mathcal{V}_K(I)| \leq \dim K[X_1, ..., X_m]/I.$$

□

Los siguientes resultados se conocen como la **cota de la huella**.

Teorema 1.15. *Sea $I \subseteq K[X]$ un ideal tal que $K[X]/I$ es de dimensión finita. Entonces la variedad $\mathcal{V}_{\bar{K}}(I)$, tiene cardinal como máximo $|\Delta_\prec(I)|$.*

Demostración. Es consecuencia inmediata de lo anterior:

$$|\mathcal{V}_{\bar{K}}(I)| \leq \dim K[X]/I = |\Delta_\prec(I)|.$$

□

Denotamos por $\mathbb{F}_q$ un cuerpo finito y $\mathbb{F}$ un cuerpo que contiene a $\mathbb{F}_q$. Sea $\bar{\mathbb{F}}$ su clausura algebraica.

Corolario 1.16. Sea $F(X_1, ..., X_m) \in \mathbb{F}[X]$ un polinomio no nulo. Si el monomio dominante del polinomio F es $X_1^{i_1} \cdots X_m^{i_m}$, con $0 \leq i_1, ..., i_m < q$, entonces el número de elementos de $\mathbb{F}_q^m$ que no anulan F es al menos $(q - i_1) \cdots (q - i_m)$.

Demostración. Sea $I = (F, X_1^q - X_1, ..., X_m^q - X_m) \subset \bar{\mathbb{F}}[X]$. Como $\mathbb{F}_q = \{a \in \bar{\mathbb{F}} \ / \ a^q = a\}$, se tiene que

$$\mathcal{V}_{\bar{\mathbb{F}}}(I) = \mathcal{V}_{\mathbb{F}_q}(I) = \mathcal{V}_{\mathbb{F}_q}(F).$$

Como $X_i^q \in LT(I)$ para todo $i \in \{1, ..., m\}$ y $X_1^{i_1} \cdots X_m^{i_m} \in LT(I)$, entonces

$$\Delta_\prec(I) \subset \{X^\alpha \ / \ 0 \leq \alpha_j \leq q \ \ \forall j \text{ y } \exists j \ / \ \alpha_j > i_j\},$$

lo cual se traduce en que $|\Delta_{\prec}(I)| \leq q^m - (q-i_1)\cdots(q-i_m)$. Por lo tanto, haciendo uso del teorema 1.15 tendremos que

$$|\mathcal{V}_{\mathbb{F}_q}(I)| = |\mathcal{V}_{\bar{\mathbb{F}}}(I)| \leq |\Delta_{\prec}(I)| \leq q^m - (q-i_1)\cdots(q-i_m).$$

Aplicando esto a lo que buscamos, tendremos que

$$|\mathbb{F}_q^m - \mathcal{V}_{\mathbb{F}_q}(F)| = |\mathbb{F}_q^m - \mathcal{V}_{\mathbb{F}_q}(I)| = q^m - |\mathcal{V}_{\mathbb{F}_q}(I)| \geq (q-i_1)\cdots(q-i_m)$$

como queríamos demostrar.

□

Capítulo 2

Flujos en redes

En este capítulo se presupone que el lector tiene nociones básicas sobre teoría de grafos. Trabajaremos con grafos dirigidos, simples y sin ciclos.

2.1. Flujos en una red

Sea $G = (V, E)$ un grafo dirigido, simple y sin ciclos, siendo V el conjunto de vértices y E el conjunto de aristas. Dada una arista $e \in E$ denotaremos $e = (x, y)$ si x es su extremo inicial e y su final.

Notación 2.1. Para cada vértice $v \in V$, definimos los siguientes conjuntos de aristas:

- $out(v)$, formado por las aristas cuyo origen es el vértice v.
- $in(v)$, con las aristas que tienen como final el vértice v.

Para una arista $j = (u, v)$, deifinimos:

- $out(j) = out(v)$
- $in(j) = in(u)$.

Definición 2.2. En un grafo dirigido, una **fuente** es un vértice del que únicamente salen aristas y un **sumidero** es un vértice al que solo llegan aristas.

Definición 2.3. Una **red** es un par (G, c), donde G es un grafo dirigido y c es una aplicación que asocia a cada arista del grafo un número real positivo, $c(e)$, al cual llamamos **capacidad** de la arista $e \in E$.

Por lo general, nos referiremos a una red simplemente por G, sobrentendiendo que se tiene la función capacidad.

Notación 2.4. Dado un grafo dirigido $G = (V, E)$ y una función $h : E \to \mathbb{R}_{\geq 0}$ denotaremos para cada vértice v:

$$h_+(v) = \sum_{\substack{e \in E \\ e \in out(v)}} h(e).$$

$$h_-(v) = \sum_{\substack{e \in E \\ e \in in(v)}} h(e).$$

Lema 2.5 (Handshaking)**.** Dado un grafo dirigido G, y una función h tal que $h(e) \in \mathbb{R}_{\geq 0}$ para $e \in E$, se tiene

$$\sum_{v \in V} h_+(v) = \sum_{v \in V} h_-(v) = \sum_{e \in E} h(e)$$

Demostración. No hay más que tener en cuenta que

$$E = \cup_{v \in V} out(v) = \cup_{v \in V} in(v),$$

siendo ambas uniones disjuntas.

Por lo tanto tenemos que

$$\sum_{e \in E} h(e) = \sum_{v \in V} \left(\sum_{e \in out(v)} h(e) \right) = \sum_{v \in V} h_+(v)$$

y a su vez

$$\sum_{e \in E} h(e) = \sum_{v \in V} \left(\sum_{e \in in(v)} h(e) \right) = \sum_{v \in V} h_-(v).$$

□

Definición 2.6. Dada una red un **flujo** f en ella es una aplicación que asigna un valor $f(e) \in \mathbb{R}_{\geq 0}$ a cada arista e, tal que satisface las condiciones siguientes:

- restricción de capacidad: $0 \leq f(e) \leq c(e)$, para cada arista $e \in E$.
- restricción de conservación: $f^+(v) = f^-(v)$, es decir, flujos salientes y entrantes en v coinciden para todo nodo v que no sea una fuente o un sumidero de la red.

Por definición de fuente, $f^-(s) = 0$ para todos los s que sean fuentes en la red, por tener únicamente aristas que salen de dichos vértices. De la misma manera, $f^+(t) = 0$ para aquellos vértices que sean sumideros, pues sólo llegan aristas.

El lema de handshaking puede aplicarse en una red para las capacidades:

$$\sum_{e\in E} c(e) = \sum_{v\in V} c^+(v) = \sum_{v\in V} c^-(v).$$

Así mismo, podemos usarlo para los flujos sobre una red:

$$\sum_{e\in E} f(e) = \sum_{v\in V} f^+(v) = \sum_{v\in V} f^-(v).$$

Notación 2.7. Dada una red, denotamos por $V_s \subset V$ a todas las fuentes, y por $V_t \subset V$ a los sumideros de la red.

Corolario 2.8. Si f es un flujo de una red G, entonces

$$\sum_{s\in V_s} f^+(s) = \sum_{t\in V_t} f^-(t).$$

Demostración. Sabemos por el lema de handshaking que $\sum f^+(v) = \sum f^-(v)$. Además, por ser f un flujo, cumplirá la restricción de conservación, esto es, para todo v que no sea fuente o sumidero, $f^+(v) = f^-(v)$ y por lo tanto $\sum_{v\notin V_s\cup V_t} f^+(v) = \sum_{v\notin V_s\cup V_t} f^-(v)$. Separando los sumandos dependiendo del carácter de $v \in V$ tendremos:

$$\sum f^+(v) = \sum_{s\in V_s} f^+(s) + \sum_{v\notin V_s\cup V_t} f^+(v) = \sum_{t\in V_t} f^-(t) + \sum_{v\notin V_s\cup V_t} f^-(v) \sum f^-(v),$$

luego

$$\sum_{s\in V_s} f^+(s) = \sum_{t\in V_t} f^-(t)$$

como queríamos demostrar. □

Definición 2.9. Se define el **valor** de un flujo en red, $val(f)$, como el número real

$$\sum_{s\in V_s} f^+(s) = \sum_{t\in V_t} f^-(t).$$

2.2. Cortes en redes

Definición 2.10. Dada una red, un **corte** $[S, T]$ es el conjunto de las aristas de G que salen de un vértice de S y llegan a uno de T, donde S y T forman una partición de V en la cual todas las fuentes pertenecen a S y todos los sumideros pertenecen a T.

Definición 2.11. Se define la **capacidad** de un corte $[S, T]$, $cap(S, T)$, como la suma de las capacidades de sus aristas.

Nótese que las capacidades de las aristas que vayan desde T hasta S no afectan a la capacidad del corte.

Todos los caminos que van desde una fuente a un sumidero pasarán por alguna de las aristas del corte $[S, T]$, por lo tanto parece claro que el flujo no podrá tener valor mayor que la capacidad de dicho corte. Veremos más adelante que esto es cierto.

Antes de continuar, extenderemos la definición de flujo saliente y entrante a los conjuntos de vértices, esto es, para algún $U \subset V$, de la siguiente forma

- $f^+(U)$ suma de flujos de las aristas que salen de U.
- $f^-(U)$ suma de flujos de las aristas que entran en U.

De esta forma, diremos que el flujo saliente del conjunto U es $f^+(U) - f^-(U)$, y el flujo entrante es $f^-(U) - f^+(U)$.

Lema 2.12. Si U es un conjunto de nodos en una red, entonces el flujo de red saliente de U es la suma de los flujos salientes de los nodos en U.

En particular, si f es un flujo y $[S, T]$ es un corte, entonces el flujo saliente de S y el entrante de T son equivalentes a $val(f)$.

Demostración. La primera parte del lema establece que

$$f^+(U) - f^-(U) = \sum_{v \in U} [f^+(v) - f^-(v).]$$

Para comprobar que esta igualdad se cumple, vamos a considerar un flujo $f(e)$, siendo e la arista de vértices x e y, y veamos de qué forma contribuye esta arista en ambos lados de la igualdad. Distinguimos 4 casos:

- $\underline{x, y \in U}$: en el lado izquierdo la arista e no aporta nada, ya que ni sale de U ni entra en U, mientras que en el lado derecho se suma $f(e)$ en $f^+(x)$ y se resta en $f^-(y)$, por lo que la aportación también es nula.

- $\underline{x, y \notin U}$: $f(e)$ no contribuye en ningún lado de la igualdad.

- $\underline{x \in U, y \notin U}$: entonces el lado izquierdo se ve incrementado en $f(e)$ por ser e una arista que sale de U. Por otra parte, el lado derecho también se incrementa $f(e)$ en el flujo saliente del nodo x, $f^+(x)$. Ambas partes han aumentado lo mismo.

- $\underline{x \notin U, y \in U}$: al contrario que en el caso anterior, esta vez el flujo entra en el conjunto U, haciendo una contribución negativa de $f(e)$ en el lado izquierdo de la igualdad. En el lado derecho $f(e)$ contribuye nuevamente de forma negativa en $f^-(y)$.

Por lo tanto se cumple la igualdad.

Por último, como f es un flujo, cumple la restricción de conservación. Entonces:

$$f^+(S) - f^-(S) = \sum_{v \in S}(f^+(v) - f^-(v)) = \sum_{s \in V_s}(f^+(s) - f^-(s)) = \sum_{s \in V_s} f^+(s)$$

de la misma forma, para el flujo entrante de T tenemos:

$$f^-(T) - f^-(T) = \sum_{v \in T}(f^-(v) - f^+(v)) = \sum_{t \in V_t}(f^-(t) - f^+(t)) = \sum_{t \in V_t} f^-(t)$$

entonces $f^+(S) - f^-(S) = f^-(T) - f^+(T)$ por el corolario 2.8, teniendo además que es igual a $val(f)$. □

Teorema 2.13. *Si f es un flujo y $[S, T]$ un corte, entonces*

$$val(f) \leq cap(S, T).$$

Demostración. Teniendo en cuenta el lema anterior, el valor de f es igual al flujo saliente del conjunto de nodos S, $val(f) = f^+(S) - f^-(S) \leq f^+(S)$.

Por otra parte, al tratarse de un flujo, la restricción de capacidad requiere que $f^+(S) \leq c^+(S) \leq cap(S, T)$ y uniendo ambas obtenemos

$$val(f) \leq cap(S, T).$$

□

2.3. Flujo máximo y corte mínimo

El problema que ahora nos planteamos es si dada una red podremos encontrar un flujo con el máximo valor posible y cómo. Teniendo el cuenta el teorema 2.13, nos interesará encontrar la mejor cota inferior para un corte en una red dada.

Definición 2.14. Un **flujo máximo** es un flujo con $val(f)$ máximo de entre todos los flujos posibles en una red.

Definición 2.15. Un **corte mínimo** es un corte $[S, T]$ con $cap(S, T)$ mínima de entre todos los cortes posibles en una red.

El siguiente teorema responde al problema planteado:

Teorema 2.16 (Flujo máximo-corte mínimo)**.** *Dada una red, el valor máximo de un flujo es igual a la capacidad mínima de un corte.*

El teorema se puede demostrar con el algoritmo de **Ford Fulkerson**, que se puede ver detalladamente en la referencia [11]. En cada etapa de este algoritmo se aumenta el valor del flujo de la misma, buscando unos caminos concretos, según explicaremos en breve. Pero antes, vamos a reducir el problema.

Sea $G = (V, E)$ una red con $\{s_1, ..., s_n\}$ fuentes y $\{t_1, ..., t_m\}$ sumideros. Consideramos el conjunto de vértices V', formado por todos los vértices de V y dos nuevos vértices s' y t'. Consideramos a su vez el conjunto de aristas E' que contiene todas las aristas de E, y $m+n$ aristas nuevas: $(s', s_1), ..., (s', s_n)$ y $(t', t_1), ..., (t', t_m)$. De esta forma, las antiguas fuentes dejan de serlo por tener aristas entrantes, de la misma forma que los antiguos sumideros dejan de serlo por tener aristas salientes, quedando en el conjunto V' solamente una fuente y un sumidero.

A estas nuevas aristas les asignamos capacidades suficientemente altas, por ejemplo, la suma $\sum_{e \in E} c(e)$. Es claro entonces que un flujo en la red G se extiende de forma única a un flujo en la red extendida, con el mismo valor. Recíprocamente, un flujo en la red extendida se restringe a un flujo del mismo valor en G.

Así, los resultados que veremos a continuación están enfocado a redes con una única fuente y un único sumidero, ya que el problema con varios se reduce a éste.

Definición 2.17. Si f es un flujo en una red, llamamos **camino aumentador** de f en G a un camino P, no necesariamente dirigido, desde la fuente

hasta el sumidero el cual no pasa dos veces por la misma arista o el mismo vértice, donde para cada arista $e \in E(P)$ se tiene que:

- Si una arista e se recorre en el mismo sentido en el camino aumentador P y en G, entonces $f(e) < c(e)$. En este caso, se define $\varepsilon(e) = c(e) - f(e)$.
- Si una arista e se recorre en sentido contrario en P y en el grafo, entonces $f(e) > 0$. En este caso, definimos $\varepsilon(e) = f(e)$.

Definición 2.18. Se define la **tolerancia** del camino aumentador P como $\min_{e \in E(P)} \varepsilon(e)$, el cual siempre toma valores estrictamente positivos.

En el siguiente resultado veremos cómo a partir de un flujo dado podemos encontrar otro flujo cuyo valor sea mayor del que partíamos.

Lema 2.19. Si P es un camino aumentador de f con tolerancia z, se define la aplicación f' en las aristas de la siguiente manera:

- Si la arista tiene el mismo sentido en P y en el grafo dirigido, entonces $f'(e) = f(e) + z$.
- Si la arista va en sentido contrario en P en el grafo, entonces $f'(e) = f(e) - z$.

Entonces f' es un flujo en la red y $val(f') = val(f) + z$.

Demostración. Veamos que efectivamente f' es un flujo cuando aplicamos esos cambios. Para ello hay que comprobar que cumple las restricciones de capacidad y de conservación. Tengamos en cuenta que un camino aumentador no afecta a las capacidades de la red.

1. *Restricción de capacidad*:

   En las aristas por las que no haya pasado el camino no se habrá producido ningún cambio en su flujo, por lo tanto $f'(e) = f(e)$, y seguirá cumpliendo la restricción de capacidad.

   En el caso de las aristas que tienen el mismo sentido en el camino aumentador P y en el grafo dirigido, tendremos que el flujo en cada arista será $f'(e) = f(e) + z \leq f(e) + \varepsilon(e) = f(e) + c(e) - f(e) = c(e)$ y por lo tanto:

   $$0 \leq f'(e) \leq c(e).$$

Por último tendríamos las aristas del camino aumentador que llevan sentido contrario al grafo. En este caso $f'(e) = f(e) - z \leq f(e) \leq c(e)$, y también $f'(e) = f(e) - z \geq f(e) - \varepsilon(e) = f(e) - f(e) = 0$ por lo tanto tendremos:

$$0 \leq f'(e) \leq c(e)$$

Luego f' cumple la restricción de capacidad para todas sus aristas.

2. *Restricción de conservación*:

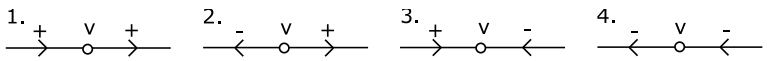
   En aquellos vértices v por los que no pasa el camino aumentador no se produce ningún cambio en el flujo y se tendrá que $f'^{+}(v) = f'^{-}(v)$ en estos vértices.

   Veamos ahora qué sucede en los vértices por los que pasa P, teniendo en cuenta que $f^{+}(v) = f^{-}(v)$ para dichos vértices. Consideraremos las aristas $e_1 = in(v) \cap P$ y $e_2 = out(v) \cap P$.

   Dependiendo del sentido que lleve cada una respecto de P y del grafo dirigido, el flujo se modificará de una forma u otra:

   - e_i lleva el mismo sentido: $f'(e_i) = f(e_i) + z$ para $i = 1, 2$.
   - e_i lleva sentido opuesto: $f'(e_i) = f(e_i) - z$ para $i = 1, 2$.

   Diferenciamos 4 situaciones en P. En la figura está representado el sentido que llevan en el grafo.



   Veamos dos de ellos con más detalle:

   1. e_1 mismo sentido, e_2 mismo sentido:

      $$f'^{-}(v) = \sum_{\substack{e \in in(v) \\ e \neq e_1}} f'(e) + f'(e_1) = \sum_{\substack{e \in in(v) \\ e \neq e_1}} f(e) + f(e_1) + z = f^{-}(v) + z$$

      $$f'^{+}(v) = \sum_{\substack{e \in out(v) \\ e \neq e_2}} f'(e) + f'(e_2) = \sum_{\substack{e \in out(v) \\ e \neq e_2}} f(e) + f(e_2) + z = f^{-}(v) + z$$

      y por lo tanto $f'^{-}(v) = f'^{+}(v)$.

2. e_1 sentido contrario, e_2 mismo sentido, ambos están en $out(v)$:

$$f'^{-}(v) = \sum_{e \in in(v)} f'(e) = \sum_{e \in in(v)} f(e) = f^{-}(v)$$

$$f'^{+}(v) = \sum_{\substack{e \in out(v) \\ e \neq e_1, e_2}} f'(e) + f'(e_2) = \sum_{\substack{e \in out(v) \\ e \neq e_1, e_2}} f(e) + f(e_2) - z + z = f^{-}(v)$$

y por lo tanto $f'^{-}(v) = f'^{+}(v)$.

El resto de casos se hacen de forma similar, teniendo en cuenta si e_1 y e_2 están en $in(v)$ ó $out(v)$.

Por lo tanto, cumple también la restricción de conservación.

Luego, hemos demostrado que efectivamente f' es un flujo.

En particular, puesto que P llega al sumidero, para la última arista del camino P, e_t, siempre tendremos el mismo sentido que el grafo dirigido, y por tanto

$$f'(e_t) = f(e_t) + z,$$

luego $f'^{-}(t) = f^{-}(t) + z$, y $val(f') = val(f) + z$, como queríamos ver. □

El algoritmo de Ford Fulkerson parte de un flujo (por ejemplo, el flujo 0, esto es, $f(e) = 0$). En cada etapa busca un camino aumentador para así poder aumentar el valor del flujo. Así mismo, en cada paso se encuentra un corte $[S, T]$ de la red. Se prueba que cuando no sea posible encontrar ningún camino aumentador, entonces habremos llegado al máximo, quedándonos con el flujo y el corte obtenidos en la última etapa, los cuales son óptimos.

Ejemplo 2.20. Dada la red de la figura 2.1 buscremos posibles caminos aumentadores para así ampliar el valor del flujo. En la figura 2.1 el flujo de cada arista está entre paréntesis, mientras que la capacidad de cada arista es el número a la derecha del flujo.

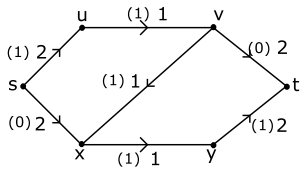


Figura 2.1: El valor del flujo es $val(f) = 1$.

Observamos que el camino $P = (s, x, v, t)$ cumple las condiciones para que P sea un camino aumentador. Busquemos su tolerancia.

Denotemos a las aristas de P por $e_1 = (s, x)$, $e_2 = (x, v)$ y $e_3 = (v, t)$. Entonces

$$\varepsilon(e_1) = c(e_1) - f(e_1) = 2$$
$$\varepsilon(e_2) = f(e_2) = 1$$
$$\varepsilon(e_3) = c(e_3) - f(e_3) = 2 - 1 = 1.$$

Por lo tanto la tolerancia de P es $z = 1$. Definimos para estas aristas:

$$f'(e_1) = f(e_1) + z = 1$$
$$f(e_2) = f(e_2) - z = 0$$
$$f(e_3) = f(e_3) + z = 1$$

y $f'(e) = f(e)$ para el resto de aristas. Ahora, $val(f') = val(f) + z = 2$, el cual podemos ver en la figura 2.2.

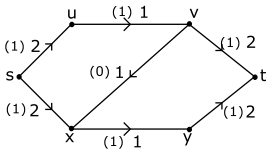


Figura 2.2: Ahora valor del flujo ha cambiado siendo $val(f') = 2$.

Ningún camino más sobre la red cumple las condiciones para que sea un camino aumentador, luego habremos llegado al flujo máximo, de valor 2.

2.4. Flujos enteros y caminos mutuamente disjuntos

Definición 2.21. Dada una red, un flujo es entero cuando sus valores $f(e)$ para todas las aristas de la red son números enteros.

El algoritmo de Ford Fulkerson permite encontrar flujos enteros máximos si se parte de capacidades enteras.

Esto es debido a que en cada etapa la tolerancia del camino aumentador será el mínimo $\varepsilon(e)$, definido como $\varepsilon(e) = c(e) - f(e)$ si la arista e lleva el mismo sentido en el grafo que en P, o $\varepsilon(e) = f(e)$ si lleva sentido opuesto en P, y en ambos casos ε es un entero positivo.

Por lo tanto, en cualquier etapa el flujo aumentará su valor en un número entero, dándonos como resultado al finalizar el algoritmo un flujo entero y máximo. Es decir:

Corolario 2.22 (Teorema integridad)**.** Si todas las capacidades de una red son enteras, entonces hay un flujo máximo que asigna un flujo entero a cada arista.

Definición 2.23. Un conjunto de caminos mutuamente disjuntos es un conjunto de caminos dirigidos en G, los cuales no tienen aristas comunes.

Proposición 2.24. Dado el grafo G, encontrar un conjunto de t caminos mutuamente disjuntos equivale a encontrar un flujo entero de valor t en G con función capacidad $c \equiv 1$.

Demostración. En primer lugar, dado un flujo entero de valor t, vamos a ver si se pueden construir t caminos mutuamente disjuntos tales que el conjunto de sus aristas coincide con el conjunto de aristas del flujo 1.

Ordenemos los vértices de G de forma que si en algún camino v está antes que v', entonces $v < v'$. Vamos a definir los caminos inductivamente.

Para cada fuente s se construyen los caminos de longitud 1 correspondientes a las aristas de flujo 1 que salen de s. En total, tendremos t caminos. Supongamos que v no es un sumidero y que se han recorrido los vértices anteriores a v, habiendo construido t caminos mutuamente disjuntos tales que sus aristas son todas las aristas de flujo 1 "anteriores" a v, es decir, tales que su extremo final w es $w < v$. Sean $C_{i_1}, ..., C_{i_r}$ caminos que terminan en v. Eso quiere decir que $f^-(v) = r$, y por ser un flujo, $f^+(v) = r$. Por lo tanto, de v salen exactamente r aristas de flujos 1, $a_1, ..., a_r$. Sustituimos entonces los caminos $C_{i_1}, ..., C_{i_r}$ por $(C_{i_1}, a_1), ..., (C_{i_r}, a_r)$.

Cuando se alcanzan todos los sumideros, entonces se tienen exactamente t caminos mutuamente disjuntos.

Recíprocamente, si al conjunto de aristas de t caminos mutuamente disjuntos se le asigna flujo 1 y al resto flujo 0, se obtiene un flujo de valor t. Veamos que efectivamente cumple las restricciones:

- *Capacidad.* Como el flujo en las aristas es 1 ó 0, entonces $f(e) \leq c(e) = 1$ para todas las aristas.
- *Conservación.* En un vértice v entran r aristas de flujo 1, donde cada una de ellas forma parte de un camino, y por lo tanto, como cada camino saldrá por aristas disjuntas, salen también $f^+(v) = r = f^-(v)$.

□

Capítulo 3

Códigos en red

3.1. Problema de los códigos en red lineales.

Vamos a empezar por explicar qué es un problema de códigos en red. La idea es que partimos de un grafo dirigido y acíclico $G = (V, E)$, en el que se distinguen dos conjuntos de nodos $S, R \subset V$. Los nodos de S son los **emisores**, esto es, los nodos en los que se genera la información. Esta información ha de llegar a los nodos R, llamados **receptores**. En principio, cada receptor **demanda** parte de la información. Cuando todos los receptores demandan toda la información se trata de **multidifusión**.

La información generada en S es un vector $\vec{X} = (X_1, ..., X_h)$, donde cada X_i se genera en único nodo de S. Dicha información se transmite a través de las aristas de la red, de forma que en cada vértice se **codifica** la información recibida, que se enviará a través de las aristas salientes, con distintas codificaciones para cada arista.

El propósito de este capítulo es ver de qué forma la información recibida por los receptores se puede **descodificar** para recuperar la parte del mensaje demandada. Veremos de qué forma se pueden definir funciones adecuadas de codificación y descodificación para que la transmisión tenga éxito.

El término "red" se utiliza en el sentido en que entendemos ya dada la función capacidad $c \equiv 1$, de forma que se puede aplicar la proposición 2.24.

Consideraremos de ahora en adelante una red, dada por un grafo dirigido y libre de ciclos, $G = (V, E)$ cuyas capacidades serán 1 para todas las aristas. En el conjunto de los nodos V diferenciamos dos subconjuntos:

- $S = \{s_1, ..., s_{|S|}\}$, los **emisores**.

- $R = \{r_1, ..., r_{|R|}\}$, los **receptores**.

Los caminos que se formarán irán desde los emisores hasta los receptores, llevando la información a través de las aristas, conjunto E. En las aristas se define un **orden ancestral** $\prec$, esto es, un orden tal que si en algún camino la arista e_i pasa antes que la arista e_j, entonces $e_i \prec e_j$.

Definición 3.1. Un problema de códigos en red consiste en :

- Una red con emisores y receptores marcados, $G = (V, E, S, R)$.
- Un vector-mensaje $\vec{X} = (X_1, ..., X_h)$, que toma valores en A^h, siendo A el alfabeto, al cual le impondremos que sea un grupo abeliano finito.
- Una **función fuente**, que es una función sobreyectiva $K : \{X_1, ..., X_h\} \longrightarrow S$ (que nos da información sobre cuál de los elementos de S ha generado el mensaje X_i).
- Una **función demanda** D, que asigna a cada $r \in R$ un subconjunto ordenado de $\{X_1, ..., X_h\}$,$(X_{i_1}, ..., X_{i_t})$ (como su nombre indica, nos dice qué parte del mensaje hay que hacer llegar a cada receptor.)

Cuando $D(r) = (X_1, ..., X_h)$ para todo $r \in R$, siendo $(X_1, ..., X_h)$ el mensaje, se dice que es un problema de **multidifusión**.

Dado un problema como el anterior, a cada arista $j = (u, v) \in E$, se le asigna una variable $Y(j)$ que toma valores en A. Estos valores se asignan a través de una función de recurrencia llamada **función de codificación**, siguiendo el orden ancestral dado en el conjunto de las aristas. Es decir, de la forma:

$$Y(j) = f_j\left(\{Y(i) \ / \ i \in in(j)\}, \{X_k \ / \ K(X_k) = u\}\right).$$

Cuando el argumento de f_j es el conjunto vacío, se le asigna el valor 0 a $Y(j)$.

A su cada vez, a cada receptor r se le asignan $|D(r)|$ variables, $Z_1^{(r)}, ..., Z_{|D(r)|}^{(r)}$, que toman valores a través de una **función de descodificación**, $b_j^{(r)}$, del tipo:

$$Z_j^{(r)} = b_j^{(r)}\left(\{Y(i) \ / \ i \in in(r)\}, \{X_k \ / \ K(X_k) = r\}\right).$$

Definición 3.2. Se dice que un problema de codificación en red **tiene solución** si existe un alfabeto no trivial A y un conjunto de funciones de codificación f_{ij} y de descodificación b_{ij}^r tales que para todo $r \in R$,

$$(Z_1^{(r)}, ..., Z_{|D(r)|}^r) = D(r).$$

Definición 3.3. Decimos que el problema es lineal si el alfabeto A es un cuerpo finito, $\mathbb{F}_q$, y si todas las funciones f_j y $b_j^{(r)}$ son lineales, es decir,

$$Y(j) = \sum_{i \in in(j)} f_{ij} Y(i) + \sum_{\substack{X_i \in S \\ K(X_i)=u}} a_{ij} X_i \qquad con\ f_{ij}, a_{ij} \in \mathbb{F}_q \tag{3.1}$$

y además

$$Z_j^{(r)} = \sum_{i \in in(r)} b_{ij}^{(r)} Y(i) + \sum_{\substack{X_i \in S \\ K(X_i)=r}} \tilde{b}_{ij}^{(r)} X_i \qquad con\ b_{ij}^{(r)} \in \mathbb{F}_q. \tag{3.2}$$

De ahora en adelante sólo trataremos redes en los que el conjunto de emisores y de receptores son disjuntos.

Esto no supone una limitación, ya que si se tiene $v \in S \cap R$, se transforma el problema en uno equivalente en el que $|S \cap R|$ disminuye en 1, como sigue: Se añade a G un nuevo vértice v', que se considera emisor y en el que se generan los mensajes que antes se generaban en v, el cual deja de ser emisor, junto con una arista que va de v' a v por cada mensaje que se generaba en v, y por la cual se transmitirá dicho mensaje. Procediendo de esta forma con todos los vértices de $S \cap R$ se consigue un problema equivalente con $S \cap R$.

En un problema lineal, utilizando de forma recursiva la expresión 3.1 se tiene que para toda arista j existen coeficientes $c_1(j)$,...,$c_h(j) \in \mathbb{F}_q$ tales que $Y(j) = c_1(j)X_1 + ... + c_h(j)X_h$.

Definición 3.4. Para cada arista j de una red, se define el vector de codificación global de j como

$$c_g(j) = (c_1, ..., c_h)$$

si $Y(j) = c_1X_1 + ... + c_hX_h$.

Identificando X_i con el vector $e_i = (0, ..., 1, ..., 0)$ de la base canónica de $\mathbb{F}_q^h$, es claro que para que $D(r) = (X_1, .., X_h)$ es necesario que los vectores $\{c_g(j)\ /\ j \in in(r)\}$ generen $\mathbb{F}_q^{|D(r)|}$.

Veamos que esta condición nos proporciona una explicación sobre posibles "cuellos de botella" en una red.

3.1.1. Red extendida

Dado un problema de red como en la definición 3.1, vamos a construir lo que llamaremos **red extendida asociada al problema**, de la siguiente forma. Añadimos al grafo

- Los vértices $\chi_1, ..., \chi_h$ tales que en cada vértice χ_i se genera el mensaje X_i, y h aristas $(\chi_i, K(X_i))$, para $1 \leq i \leq h$.
- Para cada receptor $r \in R$ se añaden $r^{(1)}, ..., r^{(|D(r)|)}$ vértices, y $|D(r)|$ aristas $(r, r^{(i)})$, $1 \leq i \leq |D(r)|$.

Considerando, como ya hemos dicho, capacidad 1 en todas las aristas, la red extendida es una red cuyas fuentes son los vértices $\chi_1, ..., \chi_h$ y cuyos sumideros son todos los vértices añadidos a cada receptor.

Definición 3.5. Dado un problema de red y $r \in R$, llamaremos **flujo extendido** en r a todo conjunto de $|D(r)|$ caminos mutuamente disjuntos desde $\{\chi_1, ..., \chi_h\}$ hasta el conjunto $\{r^{(1)}, ..., r^{(|D(r)|)}\}$.

Proposición 3.6. Si existe una solución para el problema de codificación en red, entonces existe un flujo de tamaño $t = |D(r)|$ para cada receptor.

Demostración. Dado un receptor $r \in R$, podemos tomar la red de forma que eliminemos el resto de receptores y las aristas que van a ellos, quedándonos ahora una red con un solo receptor. Vimos por la proposición 2.24 la existencia de t caminos mutuamente disjuntos de las fuentes a r equivale a la existencia de un flujo entero con valor t en la red extendida, y que ello equivale a que todo corte en G tenga capacidad $\geq t$, por el teorema 2.13.

Por lo tanto, la no existencia de los caminos implica la existencia de un corte $[S, T]$ de capacidad estrictamente menor que t, pero eso quiere decir que el subespacio vectorial de $\mathbb{F}_q^t$ formado por los vectores de codificación globales de las aristas de $[S, T]$ tienen dim $< t = |D(r)|$, así que no se puede recuperar todo el mensaje. □

Ejemplo 3.7. En la figura 3.1 vemos una red con un sólo emisor s y dos receptores r_1 y r_2. Dependiendo de la función demanda, podrá haber solución al problema. En el caso en que tengamos multidifusión no será posible encontrar un flujo para el receptor r_1, ya que nos encontramos con un *cuello de botella* en la arista vertical, esto es, la dimensión del mensaje es como mmucho 1. Por lo tanto el problema de multidifusión no tendrá solución a pesar de que para el receptor r_2 si que exista un flujo de tamaño 2.

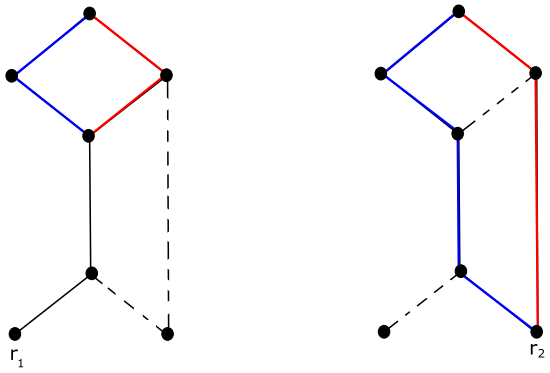


Figura 3.1: En esta red el receptor r_1 sólo puede tener $|D(r_1)| = 1$.

3.2. Códigos en red lineales para multidifusión

En esta sección trataremos problemas de códigos de red que sean lineales y de multidifusión. Esto es, trabajaremos en cuerpos finitos del tipo $\mathbb{F}_q$, donde las variables $Y(j)$ asociadas a cada arista del conjunto E del grafo son de la forma

$$Y(j) = \sum_{i \in in(j)} f_{ij} Y(i) + \sum_{\substack{X_i \in S \\ K(X_i)=u}} a_{ij} X_i,$$

y cada $D(r) = (X_1, ..., X_h)$.

Vamos a codificar de forma matricial la relación entre las variables $Y(j)$, $Z_j^{(r)}$ y las X_i. Para ello, definimos las siguientes matrices, donde se entiende que la$i$-ésima fila o columna con $1 \leq i \leq h$ se corresponde con el mensaje X_i, y que j-ésima fila o columna con $1 \leq j \leq |E|$ se corresponde con la j-ésima arista, siguiendo el orden ancestral:

- La matriz $A \in \mathcal{M}_{h \times |E|}(\mathbb{F}_q^h)$, que tiene en cada entrada (i, j) el coeficiente de codificación a_{ij} en aquellos casos en que $j \in out(K(X_i))$ y 0 en caso contrario.

- La matriz de coeficientes de codificación $F \in \mathcal{M}_{|E| \times |E|}(\mathbb{F}_q^h)$, con f_{ij} en la entrada (i, j) si $i \in in(j)$ y 0 en el resto. La matriz F tiene siempre forma triangular superior, pues al tener fijado un orden ancestral sobre el conjunto de las aristas, no podremos encontrar ningún elemento f_{ij} con $j \prec i$.

- La matriz de coeficientes de descodificación de cada $r \in R$, $B^{(r)} \in \mathcal{M}_{|E|\times h}(\mathbb{F}_q^h)$, con $b_{i,j}^{(r)}$ en la entrada (i,j) si $i \in in(r)$.

Dado un entero no negativo n, tenemos que la entrada (i,j) de la matriz F^n guarda relación con todos los caminos posibles desde la arista i a la arista j y que además son de tamaño $n+1$. Más precisamente, dicha entrada es

$$\sum_{\substack{i=i_0,i_1,\ldots,i_n=j \\ \text{camino de } G}} f_{i,i_1} f_{i_1,i_2} \cdots f_{i_{n-1},j}.$$

También, para $n=0$, al tener $F^0 = I$, podemos entender la matriz identidad como una matriz en cuyas entradas están todos los caminos de tamaño 1. Esto es, cada 1 en (i,i) representa el camino cuya arista única es i.

Así mismo, al ser un grafo simple y libre de ciclos, existirá $N \in \mathbb{N}$ tal que $F^N = 0$, siendo N el número máximo de aristas que puede tener un camino en G.

Teniendo en cuenta estas consideraciones, la matriz

$$\tilde{F} = (I + F + F^2 + ... + F^{N-1})$$

se puede entender como una matriz en la que cada entrada (i,j) nos aporta información sobre todos los caminos posibles cuya primera arista es i y cuya última arista es j, del tamaño que sean.

Además $(I + F + F^2 + ... + F^{N-1}) = (I - F)^{-1}$, pues

$$(I - F)(I + F + F^2 + ... + F^{N-1}) = I - F^N = I.$$

Una vez hemos definido estas matrices, podemos escribir la expresión 3.1 de la siguiente manera:

$$(Y(1), ..., Y(|E|)) = (Y(1), ..., Y(|E|))F + (X_1, ..., X_h)A.$$

Iterando el procedimiento, obtenemos ahora:

$$(Y(1), ..., Y(|E|)) = (Y(1), ..., Y(|E|))F^2 + (X_1, ..., X_h)A(I + F)$$

Aplicando inducción, se prueba entonces que

$$Y(1), ..., Y(|E|)) = (Y(1), ..., Y(|E|))F^N + (X_1, ..., X_h)A(I + F + F^2 + ... + F^{N-1})$$

$$= (X_1, ..., X_h) \cdot A \cdot \tilde{F}$$

y por lo tanto, tendremos la siguiente relación

$$(Y(1), ..., Y(|E|)) = (X_1, ..., X_h) \cdot A \cdot \tilde{F}.$$

De esta forma, la relación 3.2 se escribe como:

$$(Z_1^{(r)}, ..., Z_h^{(r)}) = (Y(1), ..., Y(|E|)) \cdot B^{(r)} = (X_1, ..., X_h) A \tilde{F} B^{(r)}.$$

Definición 3.8. Definimos la **matriz de transferencia** del receptor r como

$$M^{(r)} = A \cdot \tilde{F} \cdot B^{(r)} \in \mathcal{M}_{h \times h}(\mathbb{F}_q^h).$$

Ya que los problemas que vamos a tratar son de multidifusión, todos nuestros vectores demanda son $(Z_1^{(r)}, ..., Z_h^{(r)}) = (X_1, ..., X_h)$, y si conociésemos $M^{(1)} = ... = M^{(|R|)} = I$ entonces el problema de codificación lineal estaría resuelto, siendo las matrices A, F y cada una de las $B^{(r)}$ la solución del problema.

No obstante, si los coeficientes de la matriz $B^{(r)}$ se pueden modificar de forma que $M^{(r)} = I$, el problema de recuperar la información estaría también resuelto. Para esto, es necesario que la matriz $M^{(r)}$ sea inversible, esto es, que su determinante sea distinto de cero para todos los receptores. De esta forma, definiríamos $\tilde{B}^{(r)} = B^{(r)} \cdot (M^{(r)})^{-1}$, teniendo ahora con la nueva matriz

$$(Y(1), ..., Y(|E|))\tilde{B}^{(r)} = (X_1, ..., X_h) A \tilde{F} \tilde{B}^{(r)} = (X_1, ..., X_h) \cdot I.$$

Las matrices que ahora serían la solución del problema de multidifusión son A, $\tilde{F}$ y las $\tilde{B}^{(1)}, ..., \tilde{B}^{(|R|)}$. Decimos entonces que $A, \tilde{F}, B^{(1)}, ..., B^{(|R|)}$ son parte de la solución, pues ésta se obtiene con un simple cálculo de álgebra lineal.

A continuación veremos un ejemplo con el cual ilustrar todo lo que hemos expuesto en este apartado.

Ejemplo 3.9. En la red de la figura 3.2 diferenciamos por una parte el conjunto de los emisores, $S = \{s_1, s_2\}$, y por otra el conjunto de los receptores: $R = \{r_1, r_2\}$.

El mensaje que se va a enviar a través de la red va a ser (X_1, X_2), luego $h = 2$. Suponemos que $K(X_1) = s_1$ y $K(X_2) = s_2$.

Veamos cómo son las matrices únicamente para el receptor r_1, por lo que aunque el número total de aristas de la red sea 11, tomaremos solo las 9 aristas etiquetadas en la figura 3.2, correspondientes al emisor r_1.

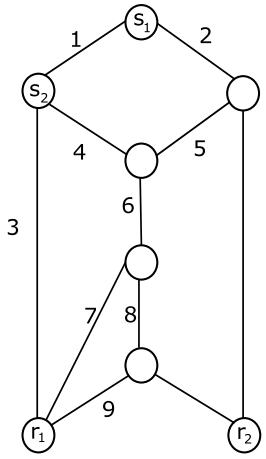


Figura 3.2: ejemplo de multidifusión para el receptor r_1.

En este caso la matriz $A \in \mathcal{M}_{2\times 9}(\mathbb{F}_q^h)$ será

$$A = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{23} & a_{24} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

y la matriz $F \in \mathcal{M}_{9\times 9}(\mathbb{F}_q^h)$

$$F = \begin{pmatrix} 0 & 0 & f_{13} & f_{14} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & f_{25} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & f_{46} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & f_{56} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & f_{67} & f_{68} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & f_{89} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Podemos hacer unos sencillos cálculos con dicha matriz, llegando a que $F^5 = 0$, coincidiendo con el número máximo de aristas que podemos encontrar en un camino, en este caso los caminos $(1, 4, 6, 8, 9)$ ó $(2, 5, 6, 8, 9)$.

Haciendo los cálculos de las potencias de F tenemos la matriz $\tilde{F} = 1 + F + F^2 + F^3 + F^4$ de la siguiente forma:

$$\tilde{F} = \begin{pmatrix} 1 & 0 & f_{13} & f_{14} & 0 & f_{14}f_{46} & f_{14}f_{46}f_{67} & f_{14}f_{46}f_{68} & f_{14}f_{46}f_{68}f_{89} \\ 0 & 1 & 0 & 0 & f_{25} & f_{25}f_{56} & f_{25}f_{56}f_{67} & f_{25}f_{56}f_{68} & f_{25}f_{56}f_{68}f_{89} \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & f_{46} & f_{46}f_{67} & f_{46}f_{68} & f_{46}f_{68}f_{89} \\ 0 & 0 & 0 & 0 & 1 & f_{56} & f_{56}f_{57} & f_{56}f_{68} & f_{56}f_{68}f_{89} \\ 0 & 0 & 0 & 0 & 0 & 1 & f_{67} & f_{68} & f_{68}f_{89} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & f_{89} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Por último, definimos la matriz $B^{(r_1)}$, a la que nos referiremos simplemente como B, con los coeficientes de descodificación b_{ij}.

$$B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ b_{31} & b_{32} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ b_{71} & b_{72} \\ 0 & 0 \\ b_{91} & b_{92} \end{pmatrix}.$$

Con el propósito de tener una notación más cómoda, vamos a definir a partir de los coeficientes de codificación unas variables d_i, donde cada d_i está asociada a un camino desde un emisor a un receptor, dependiendo de los subíndices de los coeficientes de descodificación f_{ij}:

- Desde s_1:

  $$\begin{array}{lll} d_1 = f_{13} & d_2 = f_{14}f_{46}f_{67} & d_3 = f_{25}f_{56}f_{67} \\ d_4 = f_{14}f_{46}f_{68}f_{89} & d_5 = f_{25}f_{56}f_{68}f_{89} & \end{array}$$

  Por ejemplo, d_1 es el camino que recorre las aristas 1 y 3 o d_2 el que pasa por las aristas 1, 4, 6 y 7.

- Desde s_2:

  $$d_6 = 1 \qquad d_7 = f_{46}f_{67} \qquad d_8 = f_{46}f_{68}f_{89}$$

  Por ejemplo, d_6 es el camino que recorre la arista 3.

De esta forma podemos calcular ahora la matriz de transferencia, utilizando la nueva notación d_i para representar los productos de f_{ij}, la cual queda:

$$M^{(r)} = \begin{pmatrix} a_{11}d_1b_{31} + a_{11}d_2b_{71} + a_{12}d_3b_{71} + & a_{11}d_1b_{32} + a_{11}d_2b_{72} + a_{12}d_3b_{72} + \\ a_{11}d_4b_{91} + a_{12}d_5b_{91} & a_{11}d_4b_{92} + a_{12}d_5b_{92} \\ \\ a_{23}d_6b_{31} + a_{24}d7b_{71} + a_{24}d_8b_{91} & a_{23}d_6b_{32} + a_{24}d_7b_{72} + a_{24}d_8b_{92} \end{pmatrix}.$$

Al igual que los monomio que aparecen en las entradas de la matriz $\tilde{F} = I + F + ... + F^{N-1}$ se corresponden con los caminos en el grafo G, es claro que los monomios que aparecen en las entradas de la matriz de transferencia $M^{(r)} = A \cdot \tilde{F} \cdot B^{(r)}$ se pueden entender como caminos en el grafo extendido.

Por ejemplo, en la matriz $M^{(r)}$ que hemos calculado en el ejemplo:
El monomio $a_{11}a_{23}d_2d_6b_{32}b_{71}$ corresponde al flujo

$$\{(1,4,6,7),(3)\}.$$

El monomio $a_{11}a_{24}d_1c_7b_{31}b_{72}$ corresponde al flujo

$$\{(1,3),(4,6,7)\}.$$

Por lo tanto, cada monomio de $\det M^{(r)}$ representa una unión de caminos. En este sentido se tiene el siguiente resultado:

Proposición 3.10. Cada monomio de $\det M^{(r)}$ es, en el sentido anterior, un flujo.

La demostración es puramente combinatoria pero elemental, aunque de notación muy pesada, por lo que la vamos a omitir. No obstante, intentaremos entender a través del ejemplo que acabamos de ver por qué esto es verdad.

El determinante de $M^{(r)}$ de nuestro ejemplo es la siguiente forma:

$$\begin{aligned} \det M^{(r)} = \; & a_{11}a_{23}d_2d_6b_{71}b_{32} - a_{11}a_{23}d_2d_6b_{31}b_{72} + a_{11}a_{24}d_1c_7b_{31}b_{72} \\ & -a_{11}a_{24}d_1d_7b_{31}b_{72} + a_{12}a_{23}d_3d_6b_{71}b_{32} - a_{12}a_{23}d_3d_6b_{31}b_{72} \\ & +a_{11}a_{23}d_4d_6b_{91}b_{32} - a_{11}a_{23}d_4d_6b_{31}b_{92} + a_{11}a_{24}d_1d_8b_{31}b_{92} \\ & -a_{11}a_{23}d_1d_8b_{91}b_{32} + a_{12}a_{23}d_5d_6b_{91}b_{32} - a_{12}a_{23}d_5d_6b_{31}b_{92} \end{aligned}$$

Lo sorprendente (y lo que afirma la proposición) es que cada sumando del determinante se corresponde justamente con un "producto" de caminos disjuntos. Esto se debe a que aquellos sumandos correspondientes a caminos no disjuntos se anulan unos con otros, como veremos en algunos casos a continuación:

1. Caminos que tienen común una arista final o inicial: estos monomios se anulan al estar repetidos pero con distinto signo. En el ejemplo estos caminos son:
   - d_1d_6, que tienen en común la arista 3 y aparece una vez con signo $+$ y otra con signo $-$.
   - d_2d_7 y d_3d_7 con la arista 7 en común, los cuales aparecen una vez con signo $+$ y otra con signo $-$.
   - d_4d_8 y d_5d_8 cuya arista común es la 9 y nuevamente aparecen una vez con signo $+$ y otra con signo $-$.
2. Caminos que tienen aristas intermedias en común: en este caso lo que ocurre es que monomios distintos dan lugar al mismo camino, y la combinatoria de determinantes hace que también se anulen. En el ejemplo tenemos:
   - $d_2d_8 = \mathbf{f_{14}f_{46}}f_{67} \cdot f_{46}\mathbf{f_{68}f_{89}} = \mathbf{d_4}d_7$, los cuales tienen en común las aristas 4 y 6. Así, sustituyendo los monomios, éstos se anulan 2 a 2:

     $$\begin{aligned} &a_{11}a_{24}d_4d_7b_{91}b_{72} + a_{11}a_{24}d_2d_8b_{71}b_{92} - a_{11}a_{24}d_4d_7b_{71}b_{92} - a_{11}a_{24}d_2d_8b_{91}b_{72} \\ &= a_{11}a_{24}d_2d_8b_{91}b_{72} + a_{11}a_{24}d_2d_8b_{71}b_{92} - a_{11}a_{24}d_2d_8b_{71}b_{92} - a_{11}a_{24}d_2d_8b_{91}b_{72} \\ &= 0 \end{aligned}$$

   - $d_3d_8 = \mathbf{f_{25}f_{56}}f_{67} \cdot f_{46}\mathbf{f_{68}f_{89}} = \mathbf{d_5}d_7$, con la arista 6 en común. Escribiendo aquellos monomios con d_5d_7 con los caminos d_3d_8,los monomios se anulan dos a dos, como en el caso anterior.

3.2.1. Matriz de Edmonds

Recordamos que la condición para que el problema de codificación en red tenga solución es que el determinante de cada matriz $M^{(r)}$ sea distinto de cero, o, equivalentemente, que $\prod_{r\in R} \det M^r \neq 0$. No obstante estos cálculos pueden resultar tediosos, empezando por tener que calcular el N tal que $F^N = 0$, y por tanto todas las potencias de F. A continuación, vamos a introducir una matriz cuyo cálculo resulta más sencillo, pues no es necesario hacer dichas potencias:

Definición 3.11. Definimos la matriz de Edmonds del problema de red como la matriz $E^{(r)} \in \mathcal{M}_{h+|E|\times h+|E|}(\mathbb{F}_q^h)$ dada por

$$E^{(r)} = \begin{pmatrix} A & 0 \\ I - F & B^{(r)} \end{pmatrix}.$$

En el siguiente lema veremos que su determinante tiene el mismo valor que el de la matriz $M^{(r)}$, y que por tanto podremos sustituir la condición de $\prod_{r\in R}\det M^{(r)} \neq 0$ por la de $\prod_{r\in R}\det E^{(r)} \neq 0$, ahorrándonos así encontrar la N.

Lema 3.12.

$$\det M^{(r)} = (-1)^{h\cdot(|E|+1)}\det E^{(r)}.$$

Demostración. Consideramos la matriz

$$\begin{pmatrix} I & -A(I-F)^{-1} \\ 0 & I \end{pmatrix},$$

cuyo determinante es 1. Multiplicando por la matriz de Edmonds obtenemos:

$$\begin{pmatrix} I & -A(I-F)^{-1} \\ 0 & I \end{pmatrix}\begin{pmatrix} A & 0 \\ I-F & B^{(r)} \end{pmatrix} = \begin{pmatrix} 0 & -A(I-F)^{-1}B^{(r)} \\ I-F & B^{(r)} \end{pmatrix}.$$

Haciendo los determinantes de ambas partes tenemos

$$\det E^{(r)} = \det\begin{pmatrix} A & 0 \\ I-F & B^{(r)} \end{pmatrix} = \det\begin{pmatrix} 0 & -A(I-F)^{-1}B^{(r)} \\ I-F & B^{(r)} \end{pmatrix}$$

$$= (-1)^{h\cdot|E|}\det\begin{pmatrix} -A(I-F)^{-1}B^{(r)} & 0 \\ B^{(r)} & I-F \end{pmatrix}$$

$$= (-1)^{h\cdot|E|}(-1)^{h}\det\begin{pmatrix} A(I-F)^{-1}B^{(r)} & 0 \\ B^{(r)} & I-F \end{pmatrix}$$

$$= (-1)^{h\cdot(|E|+1)}\det\left(A(I-F)^{-1}B^{(r)}\right)\det(I-F).$$

Como la matriz F es triangular superior con ceros en la diagonal, la matriz $I-F$ es triangular superior con unos en su diagonal, y por tanto, $\det(I-F) = 1$. Concluyendo así que

$$\det E^{(r)} = (-1)^{h\cdot(|E|+1)}\det\left(A(I-F)^{-1}B^{(r)}\right) = (-1)^{h\cdot(|E|+1)}\det M^{(r)}$$

como queríamos demostrar. □

3.2.2. Polinomio de transferencia

Definición 3.13. Se define el **polinomio de transferencia** de una red de codificación como el polinomio de coeficientes en $\mathbb{F}_q$ en las m indeterminadas a_{ij}, f_{ij} y $b_{ij}^{(r)}$

$$\prod_{r \in R} \det M^{(r)}.$$

Dado que la condición para que el problema tenga solución es que $\det M^{(r)} \neq 0$ para cada uno de los receptores, entonces tendremos que el polinomio de transferencia ha de ser distinto de cero, y encontrar una solución es equivalente a encontrar elementos de $\mathbb{F}_q^m$ que sean no ceros del polinomio.

Además, cada monomio del polinomio de transferencia es producto de monomios de $\det M^{(r)}$, siendo cada uno de estos un flujo de tamaño h del receptor r, y por tanto cada sumando del polinomio de transferencia será el producto de $|R|$ flujos de tamaño h, uno por cada receptor. Es decir, los a_{ij} y f_{ij} del polinomio de transferencia aparecerán como mucho con exponente $|R|$, mientras que los coeficientes de descodificación $b_{ij}^{(r)}$ aparecen como mucho con potencia 1.

Definición 3.14. Llamamos **sistema de flujos** de tamaño h a un conjunto $\mathcal{F} = \{F_1, ..., F_{|R|}\}$, donde cada F_t es un flujo de tamaño h del receptor $r_t \in R$.

Podemos entender entonces cada sumando del polinomio de transferencia como un sistema de flujos de tamaño h, por ser el producto de un flujo por cada receptor.

Proposición 3.15. El polinomio de transferencia es distinto de cero si y sólo si para todo receptor existe un flujo de tamaño h.

Demostración. Es consecuencia de la proposición 3.10, pues existe un flujo de tamaño h para r si y sólo si $\det M^{(r)} \neq 0$. □

Es decir, dada una red de codificación somos capaces de encontrar un sistema de flujos de tamaño h empleando el algoritmo de Ford Fulkerson, y de esta forma poder resolver el problema, pues garantizamos que su polinomio de transferencia será distinto de cero.

Teorema 3.16. *Un problema de códigos en red de multidifusión tiene solución si y sólo si el polinomio de transferencia correspondiente es distinto de cero.*

Además, si el problema tiene solución, entonces tiene solución en $\mathbb{F}_q$ siempre que $q > |R|$.

Demostración. Si el problema tiene solución, debe existir por la proposición 3.6 un flujo de tamaño h para cada receptor r, y por la proposición 3.15 esto implica que el determinante de la matriz de transferencia será distinto de cero.

Recíprocamente, si $\det M^{(r)} \neq 0$, puesto que no aparecen potencias mayores que $|R|$, haciendo uso del corolario 1.16 hay al menos $(q-|R|)^m$ elementos en $\mathbb{F}_q^m$ que no anulan al polinomio, siempre que $q > |R|$. Cada uno de estos elementos proporcionará una solución del problema.

□

3.3. Algoritmo

Como ya hemos comentado, para cada receptor $r \in R$ de una red, si los vectores de codificación global de las aristas $j \in in(r)$ generan $\mathbb{F}_q^h$, entonces el problema tiene solución.

Dado un problema de codificación en red de multidifusión, vamos a tratar de dar una solución al mismo, a partir de un algoritmo en tiempo polinómico, que explicamos a continuación.

Vamos a empezar por explicar esquemáticamente en qué consiste.

Consideraremos la red extendida. A las h aristas que se añaden, las denotaremos por 01, 02, ..., $0h$. La arista $0i$ conecta el vértice χ_i con el emisor $K(X_i)$, y de esta forma, para cada arista $j \in out(K(X_i))$ se puede entender $a_{ij} = f_{0i,j}$.
Además, si la arista $0i$ corresponde a $(\chi_i, K(X_i))$, asignamos a $0i$ el vector de codificación global $(0, ..., 1, ..., 0)$ con 1 en el lugar i-ésimo, esto es, $Y(0i) = X_i$.

Redefinimos nuestro orden ancestral, anteponiendo al que ya teníamos en las aristas de G las nuevas aristas, de tal forma que

$$01 \prec 02 \prec ... \prec 0h.$$

El siguiente paso será buscar un sistema de flujos de tamaño h, el cual encontraremos por el algoritmo de Ford Fulkerson, en la red aumentada.

Una vez tenemos un sistema de flujos $\mathcal{F}$ para nuestra red, falta encontrar valores para las matrices A, F y $B^{(r)}$ tales que la matriz $M^{(r)}$ sea inversible y así cada receptor podrá recibir correctamente los mensajes.

Para ello, el algoritmo irá definiendo de forma recurrente, siguiendo el orden ancestral fijado sobre las aristas, los vectores de codificación global

$c_g(j)$ de todas las aristas j del flujo.

Además, para cada camino de longitud 2 (i, j) que no forme parte de $\mathcal{F}$, se define $f_{ij} = 0$.

En cada paso del algoritmo correspondiente a arista j-ésima, se definirán conjuntos $C_1^j, ..., C_{|R|}^j$, de cardinal h, de forma que C_i^j sea un corte del flujo F_i, y tal que el conjunto $B_i^j = \{c_g(t) \ / \ t \in C_i^j\}$ genere $\mathbb{F}_q^h$. Estos conjuntos se irán modificando al paso por cada arista del sistema de flujos, siguiendo el orden ancestral, definiendo nuevos coeficientes $f_{l,j}$ para que B_i^j siga generando $\mathbb{F}_q^h$ en cada iteración. Esto se hará de forma aleatoria, comprobando posteriormente que cada B_i^j realmente es una base. De no ser así, se volverían a definir de forma aleatoria dichos coeficientes, hasta encontrar los adecuados.

Veamos posteriormente que los coeficientes aleatorios proporcionan bases de $\mathbb{F}_q^h$ con probabilidad alta, de forma que el algoritmo se realiza en tiempo polinomial.

Vamos a describir el algoritmo en detalle:

Primero fijamos un sistema de flujos.

Después, procedemos a la inicialización:

Definimos los conjuntos $C_1^0 = ... = C_{|R|}^0$ como el conjunto ordenado de las aristas que hemos añadido a la red,

$$C_i^0 = \{01, ..., 0h\} \quad 1 \leq i \leq |R|.$$

Así mismo, definimos el conjunto de vectores $B_1^0 = ... = B_{|R|}^0$, donde cada B_i^0 contiene los vectores de codificación global del conjunto C_i^0, en este caso, :

$$B_i^0 = \{(1, 0, ..., 0), (0, 1, ..., 0), ..., (0, ..., 0, 1)\}, \quad 1 \leq i \leq |R|.$$

Nótese que:

- $C_1^0, ..., C_{|R|}^0$ son cortes en cada flujo F_i para $1 \leq i \leq |R|$.
- B_i^0 genera $\mathbb{F}_q^h$ para $1 \leq i \leq |R|$.

Paso recursivo:

Tomamos la arista $j \in \mathcal{F}$ (arista j-ésima que denotamos por j), y supongamos definidos los conjuntos $C_1^{j-1}, ..., C_{|R|}^{j-1}$ para la arista anterior. Sean $F_{l_1}, ..., F_{l_k}$ los flujos tales que $j \in F_{l_1}, ..., F_{l_k}$. Del paso anterior tendremos los conjuntos correspondientes a estos flujos para la arista j; $C_{l_1}^{j-1}, ..., C_{l_k}^{j-1}$, los cuales son cortes de los flujos $F_{l_1}, ..., F_{l_k}$ respectivamente. Sea i_t la única arista de C_t^{j-1} que precede a j en F_t.

Definimos

$$C_t^j := C_t^{j-1} - \{i_t\} \cup \{j\},$$

de forma que el nuevo conjunto C_t^j sigue siendo un corte del flujo F_t.

A su vez, queremos que el conjunto B_t^j definido como:

$$B_t^j := B_t^{j-1} - \{c_g(i_t)\} \cup \{c_g(j)\}$$

genere $\mathbb{F}_q^h$.

Para ello hay que definir $c_g(j)$, lo cual se hace a partir de la elección de los coeficientes $f_{i,j}$, ya que

$$Y(j) = \sum_{l \in in(j)} f_{lj} Y(l).$$

Para aquellos índices t tales que F_t no contenga a la arista j;

$$C_t^j := C_t^{j-1} \quad \text{y por tanto} \quad B_t^j := B_t^{j-1}.$$

Una vez hemos pasado por todas las aristas de la red, tendremos para cada receptor r_t, correspondiente al flujo F_t, el corte $C_t^{r_t} = \{in(r_t)\}$ y el conjunto $B_t^{r_t}$ que es base de $\mathbb{F}_q^h$, luego el problema estará resuelto.

Ahora bien, si elegimos los coeficientes de codificación de cada iteración de forma aleatoria, ¿cuáles son nuestras opciones de que finalmente las aristas que llegan a cada receptor generen $\mathbb{F}_q^h$? Para verlo, tendremos que entender en qué situaciones esas elecciones no suponen un éxito a la hora de encontrar una solución a nuestro problema.

Volvamos al paso j, donde asignaremos valores de forma aleatoria a los coeficientes f_{ij}.

Sean $i_1, ..., i_k$ las aristas que en los flujos $F_1, ..., F_k$ preceden a j. Sea $k' = |\{i_1, ..., i_k\}|$, esto es, el número de aristas que preceden a j en $\mathcal{F}$. Supongamos que $\{i_1, .., i_{k'}\}$ son tales aristas sin repetirse.

Para tener una notación más sencilla, vamos a referirnos al vector de codificación global de la arista j como $Y(j)$, teniendo en cuenta que $X_1, ..., X_h$ corresponden a $(1, 0, ..., 0), ..., (0, ..., 0, 1)$ respectivamente, y que $Y(j)$ por tanto es combinación lineal de ellos.

Tomemos un $i_t \in \{i_1, ..., i_k\}$ concreto. El conjunto $B_t^{j-1} = \{b_{t_1}^{j-1}, ..., b_{t_h}^{j-1}\}$ de vectores de codificación globales genera $\mathbb{F}_q^h$, y por lo tanto, como $Y(j) \in \mathbb{F}_q^h$, tenemos que

$$Y(j) \in \mathbb{L}(\{b_{t_1}^{j-1}, ..., b_{t_h}^{j-1}\}).$$

Reordenamos de forma que $b_{t_h}^{j-1} = c_g(Y(i_t))$. Entonces la actualización que hacemos será

$$B_t^j = \{b_{t_1}^j, ..., b_{t_{h-1}}^j, Y(j)\}$$

con cada $b_{t_l}^j := b_{t_l}^{j-1}$ para $l \in \{1, ..., h-1\}$.

El conjunto B_t^j será una base de $\mathbb{F}_q^h$ si y sólo si $Y(j) \notin \mathbb{L}(\{b_{t_1}^j, ..., b_{t_{h-1}}^j\})$ para todo t. Para ver en qué casos B_t^j lo es o no, tenemos el siguiente resultado:

Lema 3.17. Dada una base $\{b_1, ..., b_h\}$ de $\mathbb{F}_q^h$ y dado un $c \in \mathbb{F}_q^h$, entonces existe un único $a \in \mathbb{F}_q$ tal que

$$c + a \cdot b_h \in \mathbb{L}(\{b_1, ..., b_{h-1}\}).$$

Demostración. Si el vector $c \in \mathbb{F}_q^h$, entonces puede escribirse como combinación lineal de los elementos de su base:

$$c = c_1 b_1 + ... + c_h b_h,$$

de tal forma que sólo puede ser $a = -c_h$ tal que

$$c + a b_h = c_1 b_1 + ... + c_{h-1} b_{h-1} \in \mathbb{L}(\{b_1, ..., b_{h-1}\}).$$

□

Proposición 3.18. La probabilidad de éxito en cada etapa del algoritmo es al menos

$$\frac{q^{k'} - kq^{k'-1} + (k-1)}{q^{k'}}.$$

Demostración. Si tenemos

$$Y(j) = \sum_{l \in \{i_1, ..., i_{k'}\} - \{i_t\}} f_{lj} Y(l) + f_{i_t j} Y(i_t)$$

y tomamos

$$c = \sum_{l \in \{i_1, ..., i_{k'}\} - \{i_t\}} f_{lj} Y(l) \in \mathbb{F}_q^h,$$

con f_{lj} fijos, entonces por el lema 3.17 existirá para cada t tal que $1 \leq t \leq k$, un único $f_{i_t j}$ que haga que $Y(j) \in \mathbb{L}(\{b_{t_1}^j, ..., b_{t_{h-1}}^j\})$, y que por tanto B_t^j no sea base de $\mathbb{F}_q^h$.

Por lo tanto, para cada $t \in \{1, ..., k\}$, hay $q^{k'-1}$ elecciones fallidas (pues el resto de coeficientes se pueden elegir de cualquier manera), y en total habrá como mucho $kq^{k'-1}$ de tales elecciones.

No obstante, la opción de que todos los f_{ij} para $i = i_1, ..., i_{k'}$ "malos" aparezcan a la vez se ha descontado k veces, por lo que sumamos $(k-1)$. Por lo tanto, la probabilidad de escoger los coeficientes de codificación de la arista j con éxito será de al menos

$$\frac{q^{k'} - kq^{k'-1} + (k-1)}{q^{k'}}.$$

□

Algoritmo 3.1. Algoritmo en tiempo polinómico:

1. Dada una red, tomamos la red extendida, y llamamos a las aristas nuevas $01, ..., 0h$, con $f_{0i,j} := a_{ij}$ si $0i = (\chi_i, K(X_i))$. Cada $0i$ tendrá vector de codificación global el vector e_i de la base canónica, por lo que $Y(0i) = X_i$.

2. Redefinimos el orden ancestral de la red anteponiendo

$$01 \prec ... \prec 0h \prec 1.$$

3. Encontramos un sistema de flujos, usando el algoritmo de Ford Fulkerson, $\mathcal{F} = \{F_1, ..., F_{|R|}\}$.

   Asignamos $f_{ij} = 0$ a aquellas aristas i, j tales que el camino de tamaño 2 i, j no pertenece a $\mathcal{F}$.

4. Definimos los conjuntos

$$C_1^0 = ... = C_{|R|}^0 = \{01, ..., 0h\},$$

   donde C_i^0 es un corte de F_i. y los conjuntos

$$B_1^0 = ... = B_{|R|}^0 = \{e_1, ..., e_h\}$$

   correspondientes a los vectores de codificación global de cada C_i^0.

5. Para cada arista j-ésima del sistema de flujos, realizamos las siguientes actualizaciones:

   a) Localizamos los flujos $F_{l_1}, ..., F_{l_k}$ tales que j forma parte de ellos.

   b) Para cada F_{l_t}, sea i_t la arista de F_{l_t} tal que $i_t \in in(j)$. Entonces:

$$C_{l_t}^j := C_{l_t}^{j-1} - \{i_t\} \cup j$$

$$B_{l_t} := c_g(C_{l_t}^j)$$

   esto es, los vectores de codificación de sus aristas.

c) Sea $k' = |\{i_1, ..., i_k\}|$, con $i_1, ..., i_{k'}$ distintas.

d) Elegimos los coeficientes $f_{i_1,j}, ..., f_{i'_k,j}$ para $Y(j)$ tales que B_{l_t} genere $\mathbb{F}_q^h$.

3.3.1. Ejemplo

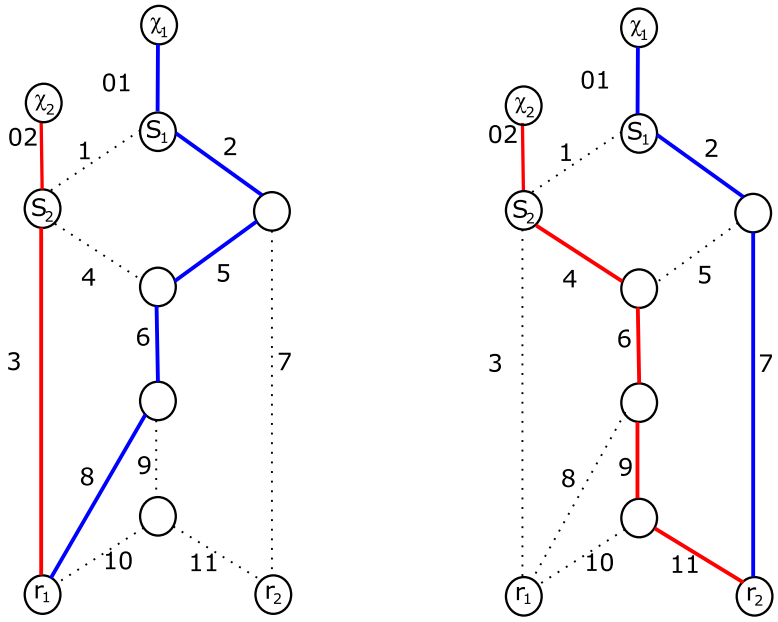


Figura 3.3:

A partir del ejemplo 3.9 que ya desarrollamos con anterioridad. El orden ancestral fijado en el conjunto de las aristas corresponde a las etiquetas en la figura 3.3.

Para implementar el algoritmo que hemos visto, extendemos la red, añadiendo las correspondientes aristas 01 y 02. En la figura 3.3 podemos ver los flujos F_1 y F_2 del sistema de flujos.

Así mismo, damos como valores iniciales igual a 0 a aquellos f_{ij} tales que $(i, j) \notin \mathcal{F}$, esto es:

$$f_{01,1} = f_{13} = f_{14} = f_{9,10} = 0.$$

Tendremos el conjunto $C_1 = \{01, 02\}$, siendo un corte del flujo F_1, y su correspondiente conjunto $B_1 = \{X_1, X_2\}$, entendiendo $X_1 = (1, 0)$ y $X_2 = (0, 1)$. De la misma forma, $C_2 = \{01, 02\}$ con $B_2 = \{X_1, X_2\}$.

Para este ejemplo hemos escogido ya los coeficientes de codificación, de forma que funcionen correctamente, con el propósito de continuar con los valores asignados y ver cómo poder asignar los coeficientes de descodificación. Sigamos no obstante el algoritmo en la siguiente tabla:

$j \in \mathcal{F}$	$Y(j)$	C_1	C_2	B_1	B_2
		$\{01, 02\}$	$\{01, 02\}$	$\{X_1, X_2\}$	$\{X_1, X_2\}$
2	$Y(2) = X_1$	$\{02, 2\}$	$\{02, 2\}$	$\{X_1, X_2\}$	$\{X_1, X_2\}$
3	$Y(3) = X_2$	$\{2, 3\}$	$\{02, 2\}$	$\{X_1, X_2\}$	$\{X_1, X_2\}$
4	$Y(4) = X_2$	$\{2, 3\}$	$\{2, 4\}$	$\{X_1, X_2\}$	$\{X_1, X_2\}$
5	$Y(5) = X_1$	$\{3, 5\}$	$\{2, 4\}$	$\{X_1, X_2\}$	$\{X_1, X_2\}$
6	$Y(6) = f_{46}Y(4) + f_{56}Y(5)$ $= X_1 + X_2$	$\{3, 6\}$	$\{2, 6\}$	$\{X_1 + X_2, X_2\}$	$\{X_1, X_1 + X_2\}$
7	$Y(7) = X_1$	$\{3, 6\}$	$\{6, 7\}$	$\{X_1 + X_2, X_2\}$	$\{X_1, X_1 + X_2\}$
8	$Y(8) = f_{68}Y(6) = X_1 + X_2$	$\{3, 8\}$	$\{6, 7\}$	$\{X_1 + X_2, X_2\}$	$\{X_1, X_1 + X_2\}$
9	$Y(9) = f_{69}Y(6) = X_1 + X_2$	$\{3, 8\}$	$\{7, 9\}$	$\{X_1 + X_2, X_2\}$	$\{X_1, X_1 + X_2\}$
11	$Y(11) = f_{9,11}Y(9) = X_1 + X_2$	$\{3, 8\}$	$\{7, 11\}$	$\{X_1 + X_2, X_2\}$	$\{X_1, X_1 + X_2\}$

Como el algoritmo nos devuelve los coeficientes de codificación, tendremos las matrices A y F:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

F:

$$F = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Conociendo las matrices A y F, podemos tras los cálculos pertinentes hallar las matrices de transferencia para así calcular matrices $B^{(r)}$ adecuadas, y que así el problema tenga solución.

Las matrices de descodificación son

$$B^{(r_1)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ b_{31}^{(1)} & b_{32}^{(1)} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ b_{81}^{(1)} & b_{82}^{(1)} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad B^{(r_2)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ b_{71}^{(2)} & b_{72}^{(2)} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ b_{11,1}^{(2)} & b_{11,2}^{(2)} \end{pmatrix}.$$

Y por lo tanto obtenemos las siguientes matrices de transferencia:

$$M^{(r_1)} = \begin{pmatrix} b_{81}^{(1)} & b_{82}^{(1)} \\ b_{31}^{(1)} + b_{81}^{(1)} & b_{32}^{(1)} + b_{82}^{(2)} \end{pmatrix} \quad \text{y} \quad M^{(r_2)} = \begin{pmatrix} b_{71}^{(2)} + b_{11,1}^{(2)} & b_{72}^{(2)} + b_{11,2}^{(2)} \\ b_{11,1}^{(2)} & b_{11,2}^{(2)} \end{pmatrix}.$$

Igualando cada una de ellas a la matriz identidad, obtenemos las matrices de descodificación :

$$B^{(r_1)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad B^{(r_2)} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Una solución al problema de codificación en red de multidifusión será por lo tanto las matrices A, F, $B^{(r_1)}$, $B^{(r_2)}$ que acabamos de dar.

Capítulo 4

Codificación en red aleatoria

En el capítulo anterior hemos visto que la existencia de una solución a un problema de códigos en red equivale a que el polinomio de transferencia sea no nulo. También hemos dado un algoritmo en que los coeficientes de codificación global se eligen de forma aleatoria, dando lugar a una solución.

Una vez elegidos los coeficientes de codificación de forma aleatoria, los receptores deben aprender cómo descodificar. Esto se hace inyectando en el sistema los vectores-mensaje

$$\vec{X} = (1, 0, ..., 0),\ \vec{X} = (0, 1, 0, ..., 0),\ ...,\ \vec{X} = (0, ..., 0, 1).$$

El receptor ve qué le llega en cada una de estas transmisiones, en forma de h mensajes ordenados, siguiente el orden ancestral en las aristas que llegan a él, las cuales generan $\mathbb{F}_q^h$. De esta forma, puede interpretar, usando álgebra lineal básica, la información recibida cuando se le envía un mensaje cualquiera.

Lo que ahora nos interesa es evaluar la probabilidad de éxito si se eligen de forma aleatoria los coeficientes de codificación. Podremos considerar también el caso en que no solo se escogen todos los coeficientes aleatoriamente, sino que algunos de ellos se fijan *a priori*.

4.1. Aproximación algebraica

A continuación, daremos una serie de cotas para la probabilidad de éxito a la hora de encontrar una solución al problema de codificación en red aleatoria, donde los coeficientes de codificación fijados a priori formen parte de la posible solución.

Proposición 4.1. 1. Sea η el número de coeficientes de codificación que elegimos de forma aleatoria. Entonces

$$P_{éxito} \geq P_{Ho} := \left(\frac{q-|R|}{q}\right)^{\eta} \tag{4.1}$$

si $q \geq |R|$.

2. Fijado un orden monomial, si $X_1^{i_1} \cdots X_m^{i_m}$ es el monomio dominante del polinomio de transferencia, entonces

$$P_{éxito} \geq P_{FP2} := \prod_{j=i}^{m} \frac{q-i_j}{q}. \tag{4.2}$$

3. Dado el polinomio de transferencia F, entonces

$$P_{éxito} \geq P_{FP1} := \text{mín}\left\{ \prod_{j=1}^{m} \frac{q-i_j}{q} \ / \ X_1^{i_1} \cdots X_m^{i_m} \text{ es monomio de } F \right\}. \tag{4.3}$$

Demostración. Tenemos que el polinomio $F = \prod_{r\in R} \det M^{(r)}$ depende de las variables a_{ij}, f_{ij} y $b_{ij}^{(r)}$ de cada receptor. Sustituimos en él las variables a_{ij} f_{ij} elegidas a priori y consideramos las b_{ij} como constantes.

Así, F es un polinomio en η variables $X_1, ..., X_\eta$ sobre el cuerpo $\mathbb{F}_q^h(\{b_{ij}^{(r)}\})$, al cual llamamos "polinomio de transferencia a priori".

Por la construcción del polinomio de transferencia, el máximo exponente que puede tener un coeficiente de codificación es $|R|$, correspondiendo al caso en que dicho coeficiente aparezca en todos los flujos de cada receptor. Por el corolario 1.16 se tiene entonces que será al menos

$$\left(\frac{q-|R|}{q}\right)^{\eta}.$$

Al fijar un orden monomial en el cual se conoce el monomio dominante del polinomio de transferencia a priori F, y haciendo uso de nuevo del corolario 1.16, obtenemos la cota P_{FP2}.

La última cota se obtiene de la misma forma, pero sin fijar un orden monomial. A pesar de ser peor cota que P_{FP2}, es más sencilla de calcular al no ser necesario precisar un orden monomial.

□

Aunque las cotas P_{FP2} y P_{FP1} son más ajustadas que P_{Ho}, la primera tiene más facilidad para ser calculada, pues no es necesario buscar el polinomio de transferencia. Es por eso que intentaremos mejorar dicha cota, conservando la sencillez de su obtención. Para ello, demos unos resultados previos:

Lema 4.2. Sean d,m y N números naturales tales que $N \leq m$. Consideramos todas las m-uplas de número naturales $i_1, ..., i_m$ tales que

1. $i_1, ..., i_m \leq d$.
2. $i_1 + ... + i_m \leq dN$.

Sea $(i_1, ..., i_m)$ tal que $(q - i_1) \cdot \cdot \cdot (q - i_m)$ es mínimo. Entonces necesariamente existe $J \subset \{1, ..., m\}$ tal que $|J| = N$, $i_j = d$ si $j \in J$ e $i_j = 0$ si $j \notin J$.

Demostración. a) Veamos que $i_1 + ... + i_m = dN$. Supongamos que no es así, es decir, $i_1 + ... + i_m < dN \leq dm$, pues $N \leq m$. Entonces $i_1 + ... + i_m < dm$ y por lo tanto existirá un k tal que $i_k < d$. Redefinamos

- $i'_k := i_k + 1 \leq d$.
- $i'_j := i_j$ para todo $j \neq k$.

Ahora los nuevos $(i'_1, ..., i'_m)$ siguen cumpliendo las condiciones:

1. $i'_1, ..., i'_m \leq d$
2. $i'_1 + ... + i'_m = i_1 + ... + i_m + 1 \leq dN$

Pero por otra parte tenemos que $(q - i'_k) < (q - i_k)$ y por tanto

$$\prod_{j=1}^{m}(q - i'_j) < \prod_{j=1}^{m}(q - i_j),$$

en contra de la hipótesis de que el producto $\prod(q - i_k)$ era mínimo.

Por lo tanto efectivamente $i_1 + ...i_m = dN$.

b) Ordenamos los i_j de forma decreciente, $i_1 \geq ... \geq i_N \geq i_{N+1} \geq ... \geq i_m$. Veamos que entonces $i_{N+1} = ... = i_m = 0$, y por tanto, teniendo en cuenta a),

$$i_1 = ... = i_N = d.$$

Supongamos que $i_m \geq 1$, con $m > N$. Entonces si se tuviese

$$i_1 = ... = i_{m-1} = d$$

se tendría $dN = d(m-1)+im$, y puesto que $0 < i_m \leq d$, debe ser $i_m = d$, y $m = N$, en contra de que $m > N$. Luego existe $k \in \{1, ..., m-1\}$ tal que $i_k < d$. Redefinimos

- $i'_k = i_k + 1$
- $i'_l = i'_l$ para $l \neq k, m$
- $i'_m = i_m - 1$

Los $(i'_1, ..., i'_m)$ siguen cumpliendo las propiedades (1) y (2) del lema.

Pero

$$\prod_{j=1}^{m}(q - i_j) > \prod_{j=1}^{m}(q - i'_j)$$

pues

$$(q-i'_k)(q-i'_m) = (q-(i_k+1))(q-(i_m-1)) = q^2-(i_k+i_m)q+i_ki_m+i_m-i_k-1 =$$

$$= (q - i_k)(q - i_m) + i_m - i_k - 1 < (q - i_k)(q - i_m)$$

por ser $i_m \leq i_k$. Esto contradice que $(i_1, ..., i_m)$ hace mínimo dicho producto, y por lo tanto $i_m = 0$.

De la misma forma se probaría de forma inductiva que que $i_h = 0$ para todo $h > N$.

Como tenemos que $i_1 + ... + i_N = dN$ y que todos los $i_j \leq d$, entonces sólo nos queda la opción de que $i_1 = ... = i_N = d$. □

Proposición 4.3. Consideramos un cuerpo $\mathbb{F}$ tal que $\mathbb{F}_q \subset \mathbb{F}$. Sea $F(X_1, ..., X_m) \in \mathbb{F}[X_1, ..., X_m]$, un polinomio distinto de cero tal existen $d, N \in \mathbb{N}$ tal que $d \leq q$ r$N \leq m$, y que todos los monomios $X_1^{i_1} \cdots X_m^{i_m}$ de F satisfacen:

1. $i_1, ..., i_m \leq d$,
2. $i_1 + ... + i_m \leq dN$.

Si $(x_1, ..., x_m) \in \mathbb{F}_q^m$ son elegidos de forma aleatoria, entonces la probabilidad de que $F(x_1, ..., x_m) \neq 0$ es al menos

$$\left(\frac{q-d}{q}\right)^N.$$

Demostración. Vamos a ver que para todos los monomios de F, $X_1^{i_1} \cdots X_m^{i_m}$, en las condiciones del teorema, se tiene que

$$\prod_{i=1}^{m} \frac{q - i_i}{q} \leq \left(\frac{q-d}{q}\right)^N . \tag{4.4}$$

Por el corolario1.16 sabemos que si $LT(F) = X_1^{i_1} \cdots X_m^{i_m}$, entonces habrá al menos $(q - i_1) \cdots (q - i_m)$ elementos que hagan $F \neq 0$. Tendremos pues que

$$P_{\acute{e}xito} \geq \frac{(q - i_1) \cdots (q - i_m)}{q^m}.$$

Haciendo uso del lema que acabamos de enunciar sabemos que

$$\frac{\prod_{j=1}^{m}(q - i_j)}{q^m} = \frac{(q-d)^N q^{m-N}}{q^m} = \left(\frac{q-d}{q}\right)^N$$

como queríamos demostrar.

□

Haremos uso de esta proposición para dar dos nuevas cotas mejores, pero tan fáciles de calcular como P_{Ho}.

Proposición 4.4. Sea η' el máximo número de aristas j tales que los coeficientes f_{ij} son elegidos de forma aleatoria y que forman parte de un mismo flujo. Entonces:

$$P_{\acute{e}xito} \geq P_{Ho2} := \left(\frac{q - |R|}{q}\right)^{\eta'} . \tag{4.5}$$

Demostración. En el polinomio de transferencia a priori cada monomio $X_1^{i_1} \cdot \cdot \cdot X_m^{i_m}$ cumplirá

- $i_1, ..., i_m \leq |R|$, pues cada coeficiente puede aparecer como mucho una vez por cada flujo, esto es, como mucho puede ser $|R|$.
- $i_1 + ... + i_m \leq |R|\eta'$, pues el número de aristas cuyos coeficientes de codificación se eligen de forma aleatorio en cada flujo es como mucho η'.

Entonces, aplicando la proposición anterior para el polinomio de transferencia tendremos la desigualdad que queríamos probar. □

Por último, podemos dar una cota que si bien es más débil que P_{Ho2}, no necesita de la búsqueda de los flujos para hallar el valor exacto de η'.

Proposición 4.5. Definimos η'' como el número de aristas j tales que a_{ij} ó f_{ij} son coeficientes de codificación aleatorios. Entonces:

$$P_{éxito} \geq P_{Ho1} := \left(\frac{q - |R|}{q}\right)^{\eta''}. \tag{4.6}$$

Demostración. Como el número de aristas en un flujo es como mucho el número total de las aristas de la red, $|E|$, en el caso de codificación en red aleatoria será como mucho η'', con $\eta'' = |E|$ en el caso en que no se asignen coeficientes a priori. Aplicando la proposición 4.3 tenemos la siguiente cota: □

Ejemplo 4.6. En la figura 4.1 tenemos todos los coeficientes de codificación eligidos a priori con valor 1, salvo $f_{4,9}, f_{69}, f_{5,10}, f_{7,10}, f_{9,11}, f_{9,12}, f_{10,11}, f_{10,12}$ que es escogerán aleatoriamente.

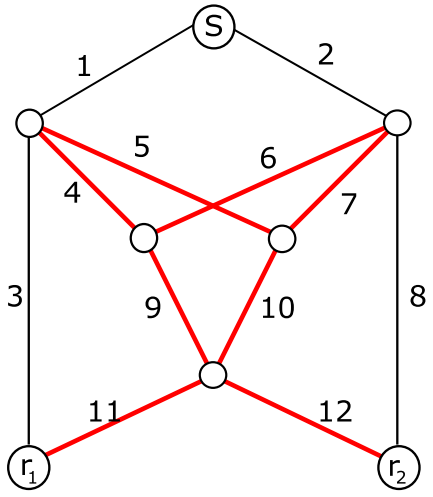


Figura 4.1: las aristas tales que f_{ij} se escoge aleatoriamente están remarcadas.

Entonces, $\eta = 8$, pues hay 8 coeficientes de codificación para elegir. Por otra parte, $\eta'' = 4$, pues tenemos las aristas 9 y 10 con $f_{i,9}$ y para $i = 4, 6$, $f_{i,10}$ para $i = 5, 7$, y las aristas 11 y 12 con $f_{i,11}$ y $f_{i,12}$ siendo $i = 9, 10$. Sin embargo, no pueden aparecer todas las aristas en el mismo flujo, luego

$\eta' = 2$. Así, tendremos las cotas:

$$P_{Ho} = \left(\frac{q-2}{q}\right)^8 \qquad P_{Ho1} = \left(\frac{q-2}{q}\right)^4 \qquad P_{Ho2} = \left(\frac{q-2}{q}\right)^2$$

donde P_{Ho2} nos proporciona una mejor aproximación, siempre que escojamos un cuerpo con $q > 2$.

A continuación, veamos un teorema en el que ordenamos las cotas que hemos obtenido:

Teorema 4.7.

$$P_{Ho} \leq P_{Ho1} \leq P_{Ho2} \leq P_{FP1} \leq P_{FP2} \leq P_{éxito}$$

Demostración. Para las cotas P_{Ho}, P_{Ho1} y P_{Ho2} basta ordenar los exponentes:

Por definición η es el número de coeficientes de codificación que vamos a elegir aleatoriamente, mientras que η'' corresponde al número de aristas j tales que a_{ij}, f_{ij} se escogen aleatoriamente, luego $\eta'' \leq \eta$. Por otra parte, η' se define a partir de las mismas aristas j que configuran η'', salvo que ahora se exige que sean el máximo de aristas j que pueden estar en el flujo, luego $\eta' \leq \eta''$, y por lo tanto, como $\frac{q-|R|}{q}$ es menor que 1:

$$\left(\frac{q-|R|}{q}\right)^{\eta} \leq \left(\frac{q-|R|}{q}\right)^{\eta''} \leq \left(\frac{q-|R|}{q}\right)^{\eta'}.$$

Para P_{FP1} y P_{FP2} es inmediato que $P_{FP1} \leq P_{FP2}$.

Veamos por último entonces que $P_{Ho2} \leq P_{FP1}$.

Tomando $N = \eta'$ y $d = |R|$, los monomios del polinomio de transferencia cumplen las condiciones de la proposición 4.3, en el cual habíamos probado que para todo monomio $X_1^{j_1} \cdots X_m^{j_m}$ se tiene

$$\prod_{i=1}^{m} \frac{q-j_i}{q} \geq \left(\frac{q-|R|}{q}\right)^{\eta'} = P_{Ho2},$$

desigualdad que entonces se cumple también para el monomio que hace mínimo $\prod_{i=1}^{m} \frac{q-j_i}{q}$, correspondiente a P_{FP1}, como queríamos ver.

□

Ejemplo 4.8. Veamos estas cotas para el ejemplo que habíamos desarrollado en el capítulo 3, en la figura 4.2

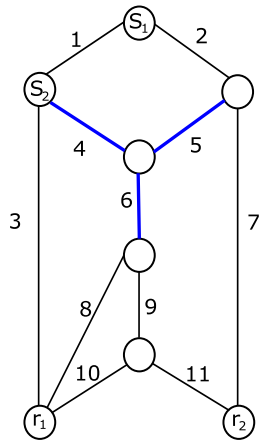


Figura 4.2: las aristas tales que f_{ij} se escoge aleatoriamente están remarcadas.

Fijaremos a priori

$$a_{11} = a_{12} = a_{23} = a_{24} = 1 \qquad y \qquad f_{ij} = 1$$

salvo $f_{4,6}$ y $f_{5,6}$. Por lo tanto la primera cota que tendremos será

$$P_{Ho} = \frac{(q-2)^2}{q^2}.$$

Denotaremos $a = f_{46}$ y $b = f_{5,6}$. Fijemos el orden monomial que sea, el monomio dominante contendrá o la expresión a^2 o ab. Supongamos que $ab \prec a^2$. Entonces:

$$P_{FP2} = \frac{(q-1)^2}{q^2}.$$

Tanto η' como η'' es igual a 1, luego

$$P_{Ho1} = P_{Ho2} = \frac{q-2}{q}.$$

Es fácil ver que el teorema 4.7 se cumple. Veamos para algunos valor de q las probabilidades que tenemos de éxito:

q	4	16	64
$P_{Ho} = \frac{(q-2)^2}{q^2}$	0,25	0,765625	0,9385
$P_{Ho2} = \frac{q-2}{q}$	0,5	0,875	0,96875
$P_{FP2} = \frac{(q-1)^2}{q^2}$	0,5625	0,8789	0,96899

Observamos que a mayor valor de q la probabilidad de éxito al asignar aleatoriamente dichos coeficientes de codificación tiene cotas próximas 1.

4.2. Aproximación combinatoria

Abordaremos esta vez a partir de métodos de combinatoria la búsqueda de una cota para la probabilidad de éxito del problema de codificación en red aleatoria, utilizando como apoyo el algoritmo 3.1. Asumamos en esta sección que ninguno de los coeficientes de codificación son conocidos a priori.

4.2.1. Cotas de flujo

Dado un problema de codificación en red, consideramos un sistema de flujos de tamaño h arbitrario $\mathcal{F} = \{F_1, ..., F_{|R|}\}$. Una vez hemos asignado de forma aleatoria los coeficientes de codificación, vamos a revisar el algoritmo 3.1 de forma que si en alguna etapa del mismo algún conjunto de vectores de codificación global no genera $\mathbb{F}_q^h$ la elección aleatoria de los coeficientes no será adecuada, y el algoritmo nos devolverá error. Sólo cuando se hayan completado todas las etapas con éxito podremos garantizar que las asignaciones resultan una solución para el problema.

Notación 4.9. Denotamos para cada arista $j \in \mathcal{F}$ por $\mathcal{R}_{\mathcal{F}}(j)$ al número de receptores cuyo flujo contiene a la arista j. Para todas las aristas del sistema de flujos se tendrá $\mathcal{R}_{\mathcal{F}}(j) \leq |R|$.

Modificamos el algoritmo ligeramente de forma que para cada arista j en el sistema de flujos se eligen de forma aleatoria todos los coeficientes f_{ij}, en la red extendida.

La inicialización del algoritmo 3.1, es decir, los valores iniciales de $C_1, ..., C_{|R|}$, $B_1, ..., B_{|R|}$, no dependían de ninguna elección aleatoria, y no se cambia con la modificación introducida.

Recordemos que en la etapa del algoritmo correspondiente a la arista j, distinta de las $01, ..., 0h$ añadidas en la red extendida, se buscan los flujos que contienen a j, de los cuales hay $k = \mathcal{R}_{\mathcal{F}}(j)$ flujos. En esta parte de la iteración hay que definir los coeficientes f_{ij} con $i, j \in \mathcal{F}$ e $i \in in(j)$ en la red extendida. S sea $k' = |in(j)|$. El mismo razonamiento hecho en la proposición 3.18 nos lleva a deducir que de los $q^{k'}$ posibles, hay como mucho $\mathcal{R}_{\mathcal{F}}(j)q^{k'-1} - (\mathcal{R}_{\mathcal{F}}(j) - 1)$ elecciones fallidas. Esta elección nos proporciona de forma inmediata nuevas cotas:

Proposición 4.10. Dado un problema de codificación en red aleatoria,

1.
$$P_{éxito} \geq P_{CF2} := \prod_{j \in \mathcal{F}} \left(\frac{q - \mathcal{R}_{\mathcal{F}}(j)}{q} + \frac{\mathcal{R}_{\mathcal{F}}(j) - 1}{q^{|in(j)|}} \right) \tag{4.7}$$

2.
$$P_{éxito} \geq P_{CF1} := \prod_{j \in \mathcal{F}} \frac{q - \mathcal{R}_{\mathcal{F}}(j)}{q} \tag{4.8}$$

Demostración. Es consecuencia directa de la probabilidad de éxito en cada etapa en la proposición 3.18. □

La cota P_{CF1} aunque es menos precisa, es más sencilla de calcular únicamente obteniendo un sistema de flujos arbitrario.

En las cotas P_{CF1} y P_{CF2} podemos reemplazar $\mathcal{R}_{\mathcal{F}}(j)$ por :

$$r_{\mathcal{F}}(j) = |\{C_v^{j-1} - \{in(j) \cap F_v\}\} \ / \ j \in \mathcal{F}|$$

correspondiente al número de cortes de la etapa $j-1$ tales que que al quitar la arista que precede a j son distintas. Por lo tanto $r_{\mathcal{F}}(j) \leq \mathcal{R}_{\mathcal{F}}(j)$ y se obtienen las cotas:

$$P_{CF4} := \prod_{j \in \mathcal{F}} \left(\frac{q - r_{\mathcal{F}(j)}}{q} + \frac{r_{\mathcal{F}(j)-1}}{q^{|in(j)|}} \right) \tag{4.9}$$

$$P_{CF3} := \prod_{j \in \mathcal{F}} \frac{q - r_{\mathcal{F}}(j)}{q} \tag{4.10}$$

La particularidad de éstas reside en que no solo importa el sistema de flujos escogido, sino que el número $r_{\mathcal{F}(j)}$ puede variar dependiendo del orden ancestral fijado en la red.

Ejemplo 4.11. En este ejemplo contemplaremos para la misma red dos órdenes ancestrales sobre las aristas diferentes. Siguiendo los pasos del algoritmo 3.1, buscaremos $r_{\mathcal{F}}(j)$ para cada j y estudiaremos las cotas P_{CF3} y P_{CF4} para ambos órdenes.

Empecemos con el orden ancestral fijado en la parte izquierda de la figura 4.3. Tomamos el sistema de flujos $\mathcal{F} = \{F_1, F_2\}$, con $F_1 = \{(1,4),(2,3,6)\}$ y $F_2 = \{(1,5),(2,3,7)\}$. Para la inicialización tomamos los conjuntos $C_1^0 = C_2^0 = \{01, 02\}$, las aristas añadidas en la red extendida, que unen respectivamente χ_i con s, $i = 1, 2$. a través de de la siguiente tabla escontraremos los $r_{\mathcal{F}}(j)$ para $j \in \mathcal{F}$:

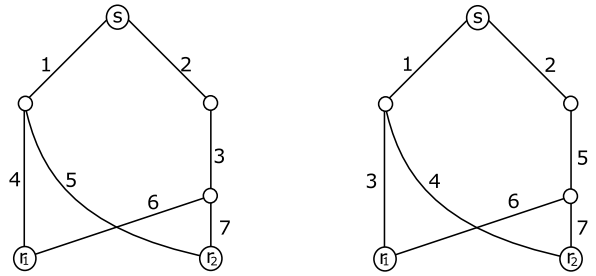


Figura 4.3: red en la que se han fijado dos ordenes monomiales distintos

j	$\|in(j)\|$	C_1^j	C_2^j	$r_{\mathcal{F}}(j)$
1	2	$\{02,1\}$	$\{02,1\}$	1
2	2	$\{1,2\}$	$\{1,2\}$	1
3	1	$\{1,3\}$	$\{1,3\}$	1
4	1	$\{3,4\}$	$\{1,3\}$	1
5	1	$\{3,4\}$	$\{3,5\}$	1
6	1	$\{4,6\}$	$\{3,5\}$	1
7	1	$\{4,6\}$	$\{5,7\}$	1

Por lo tanto tendremos para este orden monomial que

$$P_{CF3} = P_{CF4} = \left(\frac{q-1}{q}\right)^7.$$

Por otra parte, sigamos ahora el orden ancestral de la parte derecha de la figura 4.3. Esta vez tendremos el sistema de flujos $\mathcal{F} = \{F_1, F_2\}$ con $F_1 = \{(1,3),(2,5,6)\}$ y $F_2 = \{(1,4),(2,5,7)\}$. Haciendo la misma inicialización para los conjuntos C_1^0 y C_2^0 tendremos esta vez:

j	$\|in(j)\|$	C_1^j	C_2^j	$r_{\mathcal{F}}(j)$
1	2	$\{02,1\}$	$\{02,1\}$	1
2	2	$\{1,2\}$	$\{1,2\}$	1
3	1	$\{2,3\}$	$\{1,2\}$	1
4	1	$\{2,3\}$	$\{2,4\}$	1
5	1	$\{3,5\}$	$\{4,5\}$	2
6	1	$\{3,6\}$	$\{4,5\}$	1
7	1	$\{3,6\}$	$\{4,7\}$	1

Por lo tanto tendremos:

$$P_{CF3} = \left(\frac{q-1}{q}\right)^6 \left(\frac{q-2}{q}\right)$$

y

$$P_{CF4} = \left(\frac{q-1}{q}\right)^6 \left(\frac{q-2}{q} + \frac{1}{q}\right) = \left(\frac{q-1}{q}\right)^7$$

Siendo cotas diferentes para los órdenes monomiales dados. En particular, bajo la primera elección obtenemos una mejor cota.

A continuación, compararemos las cotas que acabamos de dar con las que habíamos dado en la sección anterior. Consideramos un orden monomial en el que el polinomio de transferencia tenga monomio dominante, pues de lo contrario la existencia del monomio dominante no está garantizada.

Proposición 4.12.

$$P_{CF1} \leq P_{FP2}.$$

Demostración. Elijamos una arista $j \in \mathcal{F}$, y etiquetemos las indeterminadas del polinomio de transferencia correspondientes a $\{a_{ij}\}$, $\{f_{ij}\}$ por $W_1, ..., W_{m'}$. Denotamos $Y_j^{\mathcal{R}_\mathcal{F}(j)} = W_1^{k_1} \cdots W_{m'}^{k_{m'}}$, y entonces el monomio del polinomio de transferencia correspondiente al sistema de flujos dado es

$$\prod_{j \in \mathcal{F}} Y_j^{\mathcal{R}_\mathcal{F}(j)}.$$

Por otra parte, la desigualdad 4.4 aplicada con $N = 1$ a los exponentes $(k_1, ..., k_{m'})$ de $Y_j^{\mathcal{R}_\mathcal{F}(j)}$, da

$$A(j) := \prod_{i=1}^{m'} \frac{q - k_i}{q} \geq \frac{q - \mathcal{R}_\mathcal{F}(j)}{q}.$$

Si se tiene un orden monomial tal que $\prod Y_j^{\mathcal{R}_\mathcal{F}(j)}$ es el monomio dominante, entonces

$$P_{FP2} = \prod_{j \in \mathcal{F}} A(j) \geq \prod_{j \in \mathcal{F}} \frac{q - \mathcal{R}_\mathcal{F}(j)}{q} = P_{CF1}$$

como queríamos ver. □

Proposición 4.13.

$$P_{Ho2} \leq P_{CF1}.$$

Demostración. Antes que nada, recordemos que η' es el máximo número de aristas j en un mismo flujo tales que $\{f_{ij}\}$ y $\{a_{ij}\}$ se escogen aleatoriamente.

Conservando la notación de la demostración anterior, tendremos que el monomio $\prod_{j\in\mathcal{F}} Y_j^{\mathcal{R}_{\mathcal{F}}(j)}$ es el monomio del polinomio de transferencia correspondiente al sistema de flujos $\mathcal{F}$. Podemos aplicar el corolario 1.16, por el cual tendremos una probabilidad de que el polinomio no se anula de al menos

$$\prod_{j\in\mathcal{F}} \frac{q-\mathcal{R}_{\mathcal{F}}(j)}{q}.$$

Por otra parte, los monomios del polinomio de transferencia cumplen:

- $\mathcal{R}_{\mathcal{F}}(j) \leq |R|$.
- $\sum_{j\in\mathcal{F}} \mathcal{R}_{\mathcal{F}}(j) \leq |R|\eta'$

y por la proposición 4.3 tendremos que

$$P_{CF1} = \prod_{j\in\mathcal{F}} \frac{q-\mathcal{R}_{\mathcal{F}}(j)}{q} \geq \left(\frac{q-|R|}{q}\right)^{\eta'} = P_{ho2}.$$

□

4.2.2. Cota de Balli, Yan y Zhang

Las últimas cotas que vamos a estudiar en este trabajo se obtienen a partir del algoritmo 3.1, pero con una diferencia sustancial: ahora el algoritmo se basa en los vértices que visita y no en las aristas.

Se supone, como en la sección anterior, que asignaremos todos los coeficientes de codificación de forma aleatoria. Procederemos al igual que con la cota de flujos, una vez hayamos asignado los valores de forma aleatoria, comprobaremos a través de una versión modificada del algoritmo 3.1 la viabilidad de dichos coeficientes, siguiéndolo a través de los vértices. La inicialización la haremos como estamos acostumbrados, tomando la red extendida, con los emisores $\chi_1, ..., \chi_h$ y las aristas 01, ..., $0h$. Buscamos un sistema de flujos $\mathcal{F}$ de tamaño h sobre la red extendida, al cual asignaremos $f_{ij} = 0$ a aquellos pares de aristas que no formen parte de $\mathcal{F}$. Vamos a detallar el algoritmo:

Notación 4.14. Dado un sistema de flujos $\mathcal{F} = \{F_1, ..., F_{|R|}\}$, $V(F_i)$ será el conjunto de todos los vértices por las que pasa el flujo F_i. De la misma forma, $V(\mathcal{F})$ es el conjunto de vértices por los que pasa el sistema de flujos.

Para poder recorrer la red a través de los redes hemos de fijar previamente un orden ancestral sobre el conjunto de los vértices $V(\mathcal{F})$, de la misma forma que fijábamos el de las aristas. A partir del orden ancestral dado para el conjunto de los vértices, escogemos un orden ancestral relativo para las aristas, anteponiendo aquellas que comienzan en el primer vértice de $V(\mathcal{F})$, y así sucesivamente.

Finalmente, tomamos los conjuntos $C_1 = ... = C_{|R|} = \{01, ..., 0h\}$ con sus correspondientes conjuntos de vectores de codificación global: $B_1 = ... = B_{|R|} = \{X_1, ..., X_h\}$.

Veremos a continuación en qué consiste cada etapa del algoritmo 3.1 modificado:

Supongamos que hasta el paso $v-1$ los conjuntos B_l generan $\mathbb{F}_q^h$. Tomemos ahora un $v \in V(\mathcal{F}) - \{\chi_1, ..., \chi_h\}$ siguiendo el orden ancestral. Definimos para cada $l = 1, ..., |R|$ los conjuntos:

- $I_l := in(v) \cap F_l$
- $J_l := out(v) \cap F_l$

y aplicamos la siguiente actualización al conjunto C_l^{v-1}:

$$C_l^v = (C_l^{v-1} - I_l) \cap J_l$$

Para que $B_l^v = c_g(C_l^v)$ genere $\mathbb{F}_q^h$ habrá que escoger tantos f_{ij} como pares $(i, j) \in I_l \times J_l$. En nuestro caso, al haber escogido antes los coeficientes aleatoriamente, habrá que comprobar si con esos f_{ij} los conjuntos B_l^v generan $\mathbb{F}_q^h$. Cuando el vértice de la etapa es un receptor $I_l = J_l = \emptyset$, finalizando el algoritmo con éxito si los B_l finales, con $l = 1, ..., |R|$, generan $\mathbb{F}_q^h$.

Para comprobar el éxito de las asignaciones de los coeficientes de codificación en cada etapa veamos antes el siguiente resultado, similar al lema 3.17:

Lema 4.15. Sean k, μ y $h \in \mathbb{Z}$ con $1 \leq k < h$ y tales que $k + \mu < h$. Sea $\{b_1, ..., b_h\}$ una base de $\mathbb{F}_q^h$ y sean $b'_{k+1}, ..., b'_{k+\mu}$ vectores de $\mathbb{F}_q^h$ tales que

$$V = \mathbb{L}(\{b_1, ..., b_k, b'_{k+1}, ..., b'_{k+\mu}\})$$

tiene dimensión $k + \mu$.

Dado $c \in \mathbb{F}_q^h$, el número de elecciones de $(a_{k+1}, ..., a_{k+\mu})$ con $a_i \in \mathbb{F}_q$ tales que

$$c + a_{k+1}b_{k+1} + ... + a_h b_h \in V$$

es exactamente q^μ.

Demostración. Escribimos $b'_j = \beta_{j,1}b_1 + ... + \beta_{j,h}b_h$ para $j = k+1, ..., k+\mu$, con $\beta_{j,l} \in \mathbb{F}_q$. Como V es de dimensión $k+\mu$,la siguiente matriz tiene rango $k+\mu$:

$$B = \left(\begin{array}{ccc|ccc} 1 & ... & 0 & \beta_{k+1,1} & ... & \beta_{k+\mu,1} \\ & ... & & & ... & \\ & & & & & \\ 0 & ... & 1 & \beta_{k+1,k} & ... & \beta_{k+\mu,k} \\ 0 & ... & 0 & \beta_{k+1,k+1} & ... & \beta_{k+\mu,k+1} \\ & ... & & & ... & \\ 0 & ... & 0 & \beta_{k+1,k+\mu} & ... & \beta_{k+\mu,k+\mu} \\ \hline 0 & ... & 0 & \beta_{k+1,k+\mu+1} & ... & \beta_{k+\mu,k+\mu+1} \\ & & & & & \\ 0 & ... & 0 & \beta_{k+1,h} & ... & \beta_{k+\mu,h} \end{array}\right)$$

donde se supone por comodidad que hemos reordenado $\{b_{k+\mu}, ..., b_h\}$ la matriz A formada por las $k+\mu$ primeras filas cumple $\det A \neq 0$.

El vector $c \in \mathbb{F}_q^h$ será: $c = c_1b_1 + ... + c_hb_h$. Denotemos

$$\tilde{c} = c + a_{k+1}b_{k+1} + ... + a_hb_h =$$

$$c_1b_1 + ... + (c_{k+1} + a_{k+1})b_{k+1} + ... + (c_h + a_h)b_h.$$

Para que $\tilde{c} \in V$, la matriz $(B|\tilde{c})$ ha de seguir teniendo rango $k+\mu$. Tomando la matriz A definimos las matrices

$$B_j = \left(\begin{array}{ccccccc} & & & & & & c_1 \\ & & & & & & . \\ & & & A & & & . \\ & & & & & & c_{k+\mu} + a_{k+\mu} \\ \hline 0 & ... & 0 & \beta_{k+1,j} & ... & \beta_{k+\mu,j} & c_j + a_j \end{array}\right)$$

para $j = k+\mu+1, ..., h$. Entonces, cada matriz B_j ha de tener determinante igual a cero, es decir

$$0 = \det B_j = (c_j + a_j) \det A + (\text{ términos que no dependen de } a_j).$$

Esto es, para cada elección de $(a_{k+1}, ..., a_{k+\mu})$ se tiene un único valor de a_j tal que $\tilde{c} \in V$. Por lo tanto existirán exactamente q^μ formas de elegir los coeficientes a_j de forma que $\tilde{c} \in V$. □

Proposición 4.16. La probabilidad de que en un vértice $v \notin R$ los conjuntos $B_1^v, ..., B_{|R|}^v$ generen $\mathbb{F}_q^h$, suponiendo que lo análogo es cierto para los vértices anteriores, es

$$\prod_{i=1}^{|J_l|} \frac{q^{|J_l|} - q^{i-1}}{q^{|J_l|}} = \prod_{i=1}^{|J_l|} \left(1 - \frac{1}{q^{|J_l|-i+1}}\right) \geq 1 - \frac{1}{q-1} \tag{4.11}$$

Demostración. El que la probabilidad es $\prod_{i=1}^{|J_l|} \frac{q^{|J_l|}-q^{i-1}}{q^{|J_l|}}$, se deduce del lema anterior de forma análoga a la demostración de la proposición 3.18.

Para la desigualdad vamos a probar primero que

$$\prod_{i=1}^{n}(1-x^i) - (1-\sum_{i=1}^{n} x^i) \geq 0$$

cuando $x \in (0,1)$. Para $n = 2$ se prueba inmediatamente.
Supongamos que $\prod_{i=1}^{n}(1-x^i) - (1-\sum_{i=1}^{n} x^i) = \alpha \geq 0$. Entonces:

$$\prod_{i=1}^{n}(1-x^i)(1-x^{n+1}) = (1-\sum_{i=1}^{n} x^i + \alpha)(1-x^{n+1}) =$$

$$1-\sum_{i=1}^{n} x^i + \alpha - x^{n+1} + x^{n+1}(\sum_{i=1}^{n} x^i) - x^{n+1}\alpha =$$

$$= [1-\sum_{i=1}^{n+1} x^i] + x^{n+1}\sum_{i=1}^{n} x^i + (1-x^{n+1})\alpha \geq 1-\sum_{i=1}^{n} +1x^i$$

pues $x^{n+1} \in (0,1)$.

Ahora basta tomar $x = \frac{1}{q}$, y se tiene

$$\prod_{i=1}^{n}(1-\frac{1}{q^i}) \geq 1-\sum_{i=1}^{n}\frac{1}{q^i} \geq 1-\sum_{i=1}^{\infty}\frac{1}{q^i} = 1-\frac{1}{q}\cdot\frac{1}{1-\frac{1}{q}} = 1-\frac{1}{q-1}$$

como queríamos demostrar. □

Proposición 4.17. Sea $\mathcal{I} = \{v \in V \ / \ out(v) \cap \mathcal{F} \neq \emptyset\}$ el conjunto de los puntos de $V(\mathcal{F})$ por los cuales hacemos actualizaciones en el algoritmo. Entonces:

$$P_{éxito} \geq P_{Balli2} := \prod_{v\in\mathcal{I}}\left(1-\frac{\rho(v)}{q-1}\right) \tag{4.12}$$

$$\geq P_{Balli1} := \left(1-\frac{|R|}{q-1}\right)^{|\mathcal{I}|} \tag{4.13}$$

$$\geq P_{Balli0} := \left(1-\frac{|R|}{q-1}\right)^{|V|.} \tag{4.14}$$

Demostración. Por la proposición 4.16 tendremos que, si en cada etapa v tenemos $\rho(v)$ receptores cuyo flujo pasa por el vértice v, la probabilidad de éxito al final será

$$P_{éxito} \geq 1 - \sum_{i=1}^{\rho(v)} \left(1 - \prod_{i=1}^{|J_l|} \frac{q^{|J_l|} - q^{i-1}}{q^{|J_l|}} \right) \geq 1 - \sum_{i=1}^{\rho(v)} (1 - (1 - \frac{1}{q-1})) = 1 - \frac{\rho(v)}{q-1}.$$

Las otras desigualdades se deducen de $\rho(v) \leq |R|$ $\forall v \in V(\mathcal{F})$ y $|I| \leq |V|$. □

Si $q = |R| + 1$ la cota de Balli carece de sentido para P_{Balli1} y P_{Balli0}, pues ambas serán iguales a 0, e incluso si algún $w \in \mathcal{I}$ tienen $\rho(W) = |R|$, ésta también se anulará. Veamos esto en un ejemplo, y comparemos la cota de Balli con las que ya hemos estudiado.

Ejemplo 4.18. Dada la red ya extendida de la figura 4.4, buscaremos las cotas vistas a lo largo de este capítulo, y las compararemos, para distintos valor de q.

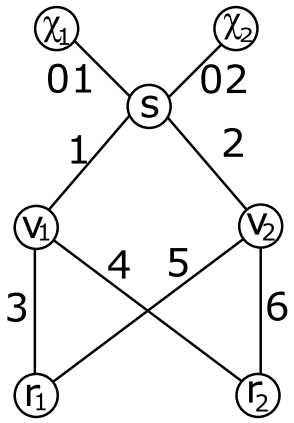


Figura 4.4: red ya extendida

Los coeficientes que se elegirán aleatoriamente son $f_{13}, f_{14}, f_{25}, f_{26}$.

De esta forma, tendremos que $\eta = \eta'' = 4$ y $\eta' = 2$, pues no habrá más que 2 aristas de entre las 4 posibles en un mismo flujo. Entonces:

$$P_{Ho} = P_{Ho1} = \left(\frac{q-2}{q} \right)^4 \qquad P_{Ho2} = \left(\frac{q-2}{q} \right)^2.$$

Fijamos un orden monomial tal que $f_{13}f_{14}f_{25}f_{26}$ sea el monomio dominante del polinomio de transferencia. Entonces:

$$P_{FP2} = \left(\frac{q-1}{q}\right)^4.$$

Para las cotas de flujos, tomamos el sistema de flujos $\mathcal{F} = \{F_1, F_2\}$ tal que $F_1 = \{(1,3),(2,5)\}$ y $F_2 = \{(1,4),(2,6)\}$. Entonces tendremos $\mathcal{R}_{\mathcal{F}}(j) = 1$ para $j = 3,4,5,6$, con $|in(j)| = 1$, correspondientes a los f_{ij} que se van a escoger de forma aleatoria. Entonces las cotas de flujo son:

$$P_{CF2} = P_{CF1} = \left(\frac{q-1}{q}\right)^4$$

Por último para las cotas de Balli fijamos un orden ancestral sobre el conjunto de los vértices, tal que

$$s < v_1 < v_2 < r_1 < r_2.$$

Tomemos el conjunto $\mathcal{I} = \{v_1, v_2\}$ según la definición, con $\rho(w) = 2$ para cada $w \in \mathcal{I}$. No consideraremos que s esté en $\mathcal{I}$ porque no tenemos que escoger los coeficientes de codificación $f_{i,s}$ siendo $i \in in(s)$. Entonces:

$$P_{Balli2} = P_{Balli1} = \left(1 - \frac{2}{q-1}\right)^2$$

$$P_{Balli0} = \left(1 - \frac{2}{q-1}\right)^5.$$

Demos ahora valores a q para las distintas cotas:

q	3	4	16	64
P_{Ho}	0,0123	0,0625	0,5862	0,8807
P_{Ho_2}	0,1111	0,25	0,7656	0,9385
$P_{FP2} = P_{CF1}$	0,1975	0,3164	0,7725	0,8389
P_{Balli1}	0	0,1111	0,7511	0,9375
P_{Balli0}	0	0,0041	0,4889	0,8510

Como $q = 3$ hace que las cotas de Balli sean cero, podemos mejorarla tomando la siguiente expresión

$$P_{Balli'} = \prod_{w \in \mathcal{I}} \left(1 - \sum_{i=1}^{\rho(v)} \left(1 - \prod_{i=1}^{|J_l|} \frac{q^{|J_l|} - q^{i-1}}{q^{|J_l|}}\right)\right)$$

que si bien es más pesada, nos proporcionará algo mayor que cero.

Para $w \in \mathcal{I}$ tendremos que $|J_l| = 2$ y que además $\rho(w) = 2$, siendo $w = v_1$ y v_2. Entonces:

$$P_{Balli'} = \left(1 - 2\left(1 - \prod_{i=1}^{2} \frac{q^2 - q^{i-1}}{q^2}\right)\right)^2 = 0{,}0343.$$

La cual nos da una mejor aproximación que P_{Ho}, pero menos fina que P_{Ho2}, patrón que se repetirá con P_{Balli1} a medida que q toma valores más grandes.

4.3. Nombre de las cotas

En esta última sección explicaremos brevemente el por qué de los nombres de las cotas que hemos citado.

En primer lugar hemos utilizado P_{Ho}, la cual fue introducida por Ho en el artículo [8]. Para las cotas P_{FP}, simplemente se ha conservado el subíndice proveniente del término en inglés de la huella de un ideal, *footprint*, pues son obtenidas a partir del corolario 1.16.

Para las cotas de flujos, P_{CF}, simplemente nos hemos quedado con las iniciales de la misma. Más información sobre esta cota se encuentra en el artículo [9].

Y la última cota, P_{Balli}, fue introducida por Balli, Yan y Zhang, y a ellos les debe su nombre. Se encuentra con más detalle en los artículos [2] y [12].

Bibliografía

[1] Rudolf Ahlswede, Ning Cai, Shuo-Yen Robert Li, and Raymond W Yeung. Network information flow. *Information Theory, IEEE Transactions on*, 46(4):1204–1216, 2000.

[2] Huseyin Balli, Xijin Yan, and Zhen Zhang. On randomized linear network codes and their error correction capabilities. *Information Theory, IEEE Transactions on*, 55(7):3148–3160, 2009.

[3] David Cox, John Little, and Donal O'shea. *Ideals, varieties, and algorithms*, volume 3. Springer, 1992.

[4] Reinhard Diestel. *Graph theory*. Springer-Verlag Berlin and Heidelberg GmbH & amp, 2000.

[5] William Fulton. *Curvas algebraicas*. Reverté, 1972.

[6] Olav Geil and Tom Hoholdt. Footprints or generalized bezout's theorem. *Information Theory, IEEE Transactions on*, 46(2):635–641, 2000.

[7] Olav Geil and Casper Thomsen. *Aspects of random network coding*. World Scientific, 2011.

[8] Tracey Ho, Muriel Médard, Ralf Koetter, David R Karger, Michelle Effros, Jun Shi, and Ben Leong. A random linear network coding approach to multicast. *Information Theory, IEEE Transactions on*, 52(10):4413–4430, 2006.

[9] Sidharth Jaggi, Peter Sanders, Philip Chou, Michelle Effros, Sebastian Egner, Kamal Jain, Ludo MGM Tolhuizen, et al. Polynomial time algorithms for multicast network code construction. *Information Theory, IEEE Transactions on*, 51(6):1973–1982, 2005.

[10] Ralf Koetter and Muriel Médard. An algebraic approach to network coding. *IEEE/ACM Transactions on Networking (TON)*, 11(5):782–795, 2003.

[11] Douglas Brent West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.

[12] Zhen Zhang et al. On the limiting behavior of random linear network codes. In *2009 Workshop on Network Coding, Theory, and Applications*, pages 1–5, 2009.