

Laboratorio di Algoritmi

Progetto “Controspionaggio” (luglio 2026)

Nota: La scadenza del progetto è fissata per martedì 7 luglio **compreso**.

Nota: Si consiglia di consultare sulla pagina web il documento che riporta le avvertenze utili per lo svolgimento del progetto. Si consiglia anche di verificare di tanto in tanto gli aggiornamenti a questo documento, che potranno riportare risposte ai dubbi degli studenti e correzioni di eventuali errori.

Il problema Le attività di spionaggio e controspionaggio si fondano spesso sulla combinazione e rielaborazione di informazioni frammentarie, raccolte da agenti diversi e comunicate dall’uno all’altro, direttamente o indirettamente. Per farlo, due agenti possono anche non incontrarsi mai in occasioni prefissate ed esplicite, ma limitarsi a frequentare regolarmente gli stessi posti fino a cogliere un’opportunità favorevole a scambiarsi informazioni.

Il progetto considera un insieme dato di *agenti* e un insieme di *posti*. Ogni agente visita una parte dei posti, e in particolare ciascuno di tali posti con una frequenza che è definita come il numero intero di visite compiute in un dato orizzonte temporale. Assumeremo, in modo semplificato, che se due agenti visitano lo stesso posto, ciascuno con la sua frequenza, allora quel posto dà un contributo potenziale allo scambio di informazioni che è proporzionale al prodotto delle due frequenze. Se si assume inoltre che le visite dei due agenti nei vari posti siano indipendenti, la forza complessiva della connessione potenziale fra loro è approssimabile con la somma di tali prodotti su tutti i posti visitati da entrambi. Non tutte le connessioni sono però abbastanza forti da essere significative: se due agenti condividono un solo posto visitato molto occasionalmente non sono davvero collegati. Consideriamo quindi potenzialmente collegati due agenti se e solo se hanno un numero di posti in comune e una forza complessiva sufficientemente grandi. L’insieme degli agenti e delle loro connessioni potenziali definisce una *rete informativa*.

Il primo obiettivo del progetto sarà un’analisi a scopi statistici della rete. Valuteremo in modo opportuno per gli agenti la distribuzione della loro mobilità e per i posti quella della loro frequenza di visita. È naturale, infatti, supporre che gli agenti più mobili e i posti più visitati siano quelli che danno un contributo più importante alla rete informativa, per lo meno a parità di altre considerazioni.

Un altro aspetto rilevante sarà la determinazione di una *sottorete emergenziale*, che connetta tutti gli agenti e sia al tempo stesso la meno ridondante, cioè contenga una sola catena di agenti per lo scambio di informazioni fra ogni coppia di agenti, e la più efficace, cioè abbia la massima forza complessiva. Daremo per scontato che nella rete iniziale tutti gli agenti possano comunicare fra loro (direttamente o indirettamente), e quindi esista sempre una sottorete emergenziale che consenta di farlo, generalmente con catene più lunghe, e quindi in modo più lento. Per convenzione arbitraria, in ogni connessione potenziale considereremo i due agenti in ordine alfabetico. Inoltre, siccome possono esistere reti non ridondanti di uguale forza complessiva, per garantire un risultato univoco nella costruzione della rete emergenziale, fra connessioni potenziali di ugual forza sceglieremo sempre quella lessicograficamente minore, cioè quella che ha primo agente (e, in caso di ulteriore parità, secondo agente) alfabeticamente precedente.

Una volta individuata la migliore sottorete emergenziale, si considererà un secondo problema. Supporremo di voler danneggiare la sottorete stessa il più possibile, perché costituita da agenti nemici o perché si vuole valutarne la resilienza ad attacchi nemici. Il danneggiamento avverrà rimuovendo fisicamente alcuni agenti dalla

sottorete o neutralizzandoli con l'impedir loro di visitare i posti consueti. Supponendo che si possa colpire un dato numero di agenti, il problema di determinare il sottoinsieme più efficace è computazionalmente difficile. Lo risolveremo quindi euristicamente applicando una strategia *greedy*, che consiste nel determinare iterativamente l'agente la cui rimozione porta il massimo danno alla sottorete (definito nel seguito), fino a raggiungere il numero desiderato di agenti rimossi. Ad ogni rimozione di un agente, la sottorete emergenziale si aggiorna usando anche connessioni potenziali che prima non includeva, ma che facevano parte della rete informativa complessiva. In generale, questo riduce la sua forza complessiva. Nonostante l'uso di nuove connessioni, è possibile che a un certo punto la sottorete emergenziale risulti comunque sconnessa. In ogni caso, l'effetto della rimozione di un agente sarà misurato per prima cosa dal numero di spezzoni nei quali la sottorete emergenziale risulta divisa (che va massimizzato) e poi dalla somma dei valori di forza degli spezzoni di sottorete residui (che va minimizzata). Nel caso la rimozione di diversi agenti dia luogo a sottoreti equivalenti, rimuoveremo il primo agente in ordine alfabetico.

Il progetto Il progetto richiede la stesura di un programma che legga da un file di testo le informazioni relative agli agenti, ai posti e alle frequenze di visita dei posti da parte degli agenti. La prima riga del file conterrà il numero n degli agenti e il numero k di agenti da rimuovere, che sarà sempre strettamente inferiore a n . Le righe seguenti riportano i nomi degli agenti, identificati da parole singole, tutte diverse, di al più 30 caratteri¹. Per esempio:

```
6 2
Bruno
Diego
Fabio
Ada
Elena
Carla
```

indica che vi sono $n = 6$ agenti, che $k = 2$ dovranno essere rimossi e che i nomi degli agenti sono: Bruno, Diego, Fabio, Ada, Elena e Carla.

Dopo i nomi degli agenti, il file riporta in una riga il numero p dei posti e le due soglie intere π e ϕ (rispettivamente sul numero di posti condivisi e sulla loro forza complessiva), che determinano l'esistenza di collegamenti potenziali fra agenti. Seguono le informazioni relative ai posti: ogni riga riporta il nome di un posto, che è una parola singola di al più 30 caratteri². Il nome di ciascun posto j è seguito da uno spazio bianco singolo e dal numero di agenti che visitano il posto stesso. Seguono, alternativamente, il nome di ogni agente i e la frequenza della visita f_{ij} , tutti separati da spazi bianchi singoli. Per esempio:

```
7 2 4
Emporio 3 Diego 2 Elena 2 Fabio 1
Bazar 2 Ada 2 Bruno 2
Faro 2 Elena 1 Fabio 2
Caffe' 3 Bruno 1 Carla 2 Diego 2
Galleria 2 Bruno 1 Diego 3 Darsena 3 Carla 1 Diego 2 Elena 1
```

¹Questa indicazione serve a dimensionare le stringhe in modo semplice. Non va intesa come un invito a considerare costante la lunghezza dei nomi stessi nelle analisi di complessità. Tanto meno va considerato limitato (e quindi costante) il numero degli agenti in base al limitato numero di nomi generabili per la lunghezza indicata.

²Valgono per i nomi dei posti le stesse indicazioni fornite per i nomi degli agenti.

Archivio 3 Ada 1 Bruno 2 Carla 1

indica che ci sono 7 posti e due agenti sono potenzialmente collegati se e solo se visitano gli stessi 2 posti con frequenze tali che la somma dei due prodotti sia pari almeno a 4. Il primo posto è l'Emporio, che è visitato da 3 agenti: Diego ed Elena lo visitano due volte durante l'orizzonte temporale, mentre Fabio lo visita una volta sola.

Il programma dovrà per prima cosa produrre due analisi. Quella sulla mobilità degli agenti si aprirà con una riga che contiene solo la parola chiave **Agenti:**. Seguirà la stampa dei nomi degli agenti in ordine decrescente rispetto al numero di posti visitati. In caso di parità, si osserverà l'ordine decrescente per la somma delle frequenze di visita sui posti. In caso di ulteriore parità, varrà l'ordine alfabetico. La stampa dovrà riportare su ogni riga il nome di un agente, il numero di posti visitati e la frequenza totale di visita, separati da spazi bianchi singoli. Nell'esempio considerato:

```
Agenti:  
Diego 4 9  
Bruno 4 6  
Carla 3 4  
Elena 3 4  
Ada 2 3  
Fabio 2 3
```

indica che Diego visita 4 posti (Emporio, Caffè', Galleria e Darsena) per un totale di $2 + 2 + 3 + 2 = 9$ visite. Anche Bruno visita 4 posti, ma per un totale di 6 visite, e così via. Per Carla ed Elena e per Ada e Fabio vale l'ordine alfabetico.

La seconda analisi riguarda la frequenza di visita dei posti e si aprirà con una riga che contiene solo la parola chiave **Posti:**. Seguirà la stampa dei nomi dei posti in ordine decrescente rispetto al numero di agenti visitanti. In caso di parità, si osserverà l'ordine decrescente sul numero totale di visite, e poi eventualmente l'ordine alfabetico. Ogni riga riporterà il nome del posto, il numero di agenti visitatori e quello di visite, separati da spazi singoli. Nell'esempio considerato:

```
Posti:  
Caffè' 3 5  
Emporio 3 5  
Archivio 3 4  
Darsena 3 4  
Bazar 2 4  
Galleria 2 4  
Faro 2 3
```

indica che il Caffè' viene visitato da 3 agenti (Bruno, Carla e Diego) per un totale di $1 + 2 + 2 = 5$ visite. Lo stesso vale per l'Emporio, che però è alfabeticamente successivo, mentre Archivio e Darsena hanno 3 agenti visitatori, ma con sole 4 visite complessive.

Avendo costruito la rete, ne stamperemo una breve caratterizzazione, costituita da una prima riga contenente la sola parola chiave **Rete:** e una seconda riga costituita dal numero di agenti, dal numero di connessioni potenziali e dalla somma dei valori della forza di ciascuna connessione potenziale, separati da spazi singoli. Nell'esempio considerato:

```
Rete:  
6 6 31
```

a indicare che la rete comprende 6 agenti, con 6 connessioni potenziali e una forza complessiva pari a 31. Nel caso specifico, le singole connessioni potenziali, che non è richiesto stampare, sono: **Ada Bruno** con forza 6, **Bruno Carla** con forza 4, **Bruno Diego** con forza 5, **Carla Diego** con forza 6, **Diego Elena** con forza 6, **Elena Fabio** con forza 4. Si noti l'ordine alfabetico imposto sui due agenti nella descrizione delle connessioni.

A questo punto, si dovrà stampare la sottorete emergenziale. La prima riga riporterà la parola chiave **Sottorete:**, seguita da uno spazio bianco e dalla somma dei valori della forza delle connessioni che ne fanno parte. Seguiranno, riga per riga, le connessioni potenziali, in ordine lessicografico (come specificato più sopra), ciascuna costituita dai nomi dei due agenti, separati da uno spazio bianco e dalla forza della connessione, separata da uno spazio bianco. Nell'esempio considerato:

```
Sottorete: 27
Ada Bruno 6
Bruno Diego 5
Carla Diego 6
Diego Elena 6
Elena Fabio 4
```

indica che la sottorete ha forza pari a $6+5+6+6+4 = 27$ ed è composta dalle connessioni potenziali **Ada Bruno** con forza 6, **Bruno Diego** con forza 5, **Carla Diego** con forza 6, **Diego Elena** con forza 6, **Elena Fabio** con forza 4. Sostanzialmente, solo una connessione potenziale risulta ridondante e quella da cancellare per ottenere la sottorete emergenziale più forte è **Bruno Carla**, con forza 4. Si noti l'ordine lessicografico fra le connessioni, oltre che fra i due agenti di ognuna.

Per concludere, bisognerà stampare il sottoinsieme di agenti rimossi. La stampa si aprirà con una riga che contiene solo la parola chiave **Rimozione:**. Le righe seguenti riporteranno i nomi di ciascuno degli agenti, nell'ordine con cui l'algoritmo greedy li ha scelti. Ogni riga riporterà il nome, uno spazio bianco, il numero di spezzoni in cui la sottorete emergenziale è divisa (1 se tutti gli agenti rimasti possono ancora comunicare), un altro spazio bianco e la forza della sottorete dopo la rimozione dell'agente. Naturalmente, solo le connessioni potenziali fra gli agenti non rimossi contribuiscono alla forza della sottorete. Nell'esempio considerato:

```
Rimozione:
Diego 2 14
Bruno 3 4
```

indica che l'agente con impatto massimo sulla sottorete emergenziale è **Diego**, la cui rimozione spezza l'intera rete informativa in due parti e fa scendere la forza della sottorete emergenziale (ovviamente, anch'essa divisa in due spezzoni) da 27 a 14. A questo punto, l'agente con impatto massimo sulla nuova sottorete emergenziale è **Bruno**, la cui rimozione spezza l'intera rete informativa e la sottorete in tre parti e fa scendere la forza della sottorete da 14 a 4. È facile vedere, infatti, che l'unica connessione potenziale rimasta è quella fra **Elena** e **Fabio**, la cui forza è pari a 4.

Chiarimenti

In questa sezione saranno riportate le risposte a domande e dubbi.