

# Analisi aggregata per ore

Michele Carignani, Alessandro Lenzi

27 febbraio 2014

## 1 Generazione dei grafi orari

Per prima cosa i dati sono stati aggregati per ora. I dati originali del dataset **MItoMI** sono nel formato:

```
timestamp \t SourceId \t DestId \t Stregth
```

e sono stati suddivisi in 24 files (uno per ogni ora) e aggregati, per cui ogni file contiene (al massimo<sup>1</sup>) un record per ogni nodo nel formato:

```
SourceId \t DestId:Strength [\t DestId:Strength]
```

A questo punto i pesi sugli archi (sopra chiamati **Strength**) sono stati riscalati rispetto alla somma dei valori della stella uscente di un nodo, ottenendo la probabilità di transire dal nodo  $i$  al nodo  $j$ , ovvero:

$$sumStrength_i = \sum_{j \in FS(i)} Strength_{ij}$$

$$probability_{ij} = \frac{Strength_{ij}}{sumStrength_i}$$

## 2 Ricerca delle componenti fortemente connesse

## 3 Risultati

### 3.1 Statistiche

Per ogni ora, delle probabilità (totale non per nodi di partenza) media delle probabilità (totale)

$$MeanProb_h = \sum_{(i,j) \in Edges_h} \frac{probability_{i,j}}{|Edges_h|}$$

variancz

### 3.2 Componenti Fortemente Connesse

#### 3.2.1 Taglio 1

#### 3.2.2 Taglio 2

#### 3.2.3 Taglio 1

---

<sup>1</sup>poichè certi nodi possono non avere chiamate in uscita in una certa fascia oraria.

Fascia Oraria	0	1	2	3	4	5	6	7	8	9	10	
Numero CFC	•	•										
Dimensioni CF	•	•										

Figura 1: CFC scoperte con taglio  $AVG = 0.00291$