# Sintesi sulla Struttura a Comunità nelle Reti: Web-Communities e Cut-Communities

Basato sulla dispensa di M. Di Ianni (2017/2018)

**Abstract**

Questa sintesi analizza il concetto di struttura a comunità nelle reti complesse, un fenomeno empirico che descrive la presenza di sottoinsiemi di nodi densamente interconnessi. Vengono introdotte e confrontate due definizioni formali di comunità: le **web-communities**, basate sulla connettività interna di ciascun nodo, e le **cut-communities**, definite in termini di tagli minimi del grafo. La discussione si concentra sul problema computazionale di partizionare un grafo in comunità, dimostrando che, mentre la partizione in cut-communities è trattabile in tempo polinomiale, decidere se un grafo ammetta una partizione in due web-communities è un problema **NP-completo**. Viene presentata in dettaglio la dimostrazione di quest'ultimo risultato attraverso una riduzione da 3-SAT.

## 1 Introduzione alla Struttura a Comunità

L'analisi sperimentale di reti reali, in particolare quelle sociali, rivela una caratteristica topologica fondamentale: la **struttura a comunità** (o *clustering*). Questa proprietà descrive l'organizzazione della rete in moduli o regioni in cui i nodi sono densamente collegati tra loro, mentre i collegamenti tra regioni diverse sono relativamente radi. L'identificazione di tali comunità ha importanti applicazioni, dalla classificazione dei nodi alla comprensione della dinamica dei flussi informativi, fino all'ottimizzazione di servizi distribuiti come i sistemi di raccomandazione.

## 2 Definizioni Formali di Comunità

Per formalizzare il concetto intuitivo di "regione coesa", sono state proposte diverse definizioni. La dispensa ne analizza due principali.

### 2.1 Web-Communities

Questa definizione si basa su una proprietà locale: ogni membro di una comunità deve avere più legami all'interno della comunità che all'esterno.

**Definizione 1** (Web-Community)**.** Dato un grafo $G = (V, E)$, un sottoinsieme proprio $C \subset V$ è una **web-community** se per ogni nodo $u \in C$ vale:

$$|N(u) \cap C| > |N(u) - C|$$

dove $N(u)$ è l'insieme dei vicini di $u$.

La definizione può essere rilassata nella sua versione "debole" (**weak web-community**) utilizzando la disuguaglianza non stretta ($\geq$).

### 2.2 Cut-Communities

Questa definizione si basa su una proprietà globale: una comunità è un insieme di nodi che può essere "separato" dal resto del grafo tagliando un numero minimo di archi.

**Definizione 2** (Cut-Community)**.** Dato un grafo $G = (V, E)$, una **cut-community** è un sottoinsieme proprio $C \subset V$ che definisce una partizione $\langle C, V - C \rangle$ tale che il numero di archi tra $C$ e $V - C$ (la misura del taglio) sia minimo rispetto a tutti i possibili tagli non banali del grafo.

# 3 Il Problema della Partizione in Comunità

Un problema fondamentale è decidere se un grafo possa essere partizionato in due comunità. La complessità di questo problema dipende drasticamente dalla definizione di comunità adottata.

## 3.1 Partizione in Cut-Communities

Il problema di trovare un taglio di misura minima in un grafo è computazionalmente trattabile e può essere risolto in tempo polinomiale (ad esempio, tramite algoritmi basati sul flusso massimo). Pertanto, decidere se un grafo è partizionabile in due cut-communities è un problema in **P**. Tuttavia, le partizioni ottenute tramite tagli minimi non sempre corrispondono a comunità significative. Spesso, il taglio minimo isola un singolo nodo ad alta connettività, producendo una comunità banale.

## 3.2 Relazione tra le Definizioni

Esiste un legame tra le due nozioni. Si può dimostrare che se una partizione $\langle C, V - C\rangle$ è un taglio minimo e gli insiemi $C$ e $V - C$ hanno cardinalità maggiore di uno, allora sia $C$ che $V - C$ sono anche *weak web-communities*. Il problema è che gli algoritmi per il taglio minimo non offrono garanzie sulla dimensione delle partizioni risultanti.

## 3.3 NP-Completezza del Partizionamento in Web-Communities

A differenza del caso precedente, decidere se un grafo possa essere partizionato in due web-communities è un problema computazionalmente intrattabile.

**Teorema 1.** *Decidere se un dato grafo $G = (V, E)$ può essere partizionato in due web-communities è un problema* ***NP-completo****.*

*Dimostrazione (schizzo).* Il problema è chiaramente in NP. Per dimostrarne la NP-durezza, si costruisce una riduzione polinomiale dal problema **3-SAT**. Data una formula booleana $f$ in forma 3-CNF, si costruisce un grafo $G_f$ tale che $G_f$ è partizionabile in due web-communities se e solo se $f$ è soddisfacibile.

La costruzione di $G_f$ prevede i seguenti elementi:

- **Nodi "Verità" (T e F):** Due nodi speciali, T (True) e F (False), che devono necessariamente appartenere a comunità diverse in qualsiasi partizione non banale. Se finissero nella stessa comunità, tutti gli altri nodi del grafo sarebbero costretti a unirsi a loro, invalidando la partizione.
- **Gadget per le Variabili:** Per ogni variabile booleana $x_i$, si costruisce un sottografo (gadget) contenente i nodi $x_i$ e $\bar{x}_i$. La struttura del gadget, sfruttando nodi ausiliari di grado 2 e un lemma preliminare, forza i nodi $x_i$ e $\bar{x}_i$ a trovarsi in comunità opposte. L'appartenenza del nodo $x_i$ alla comunità di T viene interpretata come l'assegnazione $x_i = \text{true}$.
- **Gadget per le Clausole:** Per ogni clausola $c_j = (l_1 \vee l_2 \vee l_3)$, si introduce un nodo $c_j$ collegato ai tre nodi letterali corrispondenti ($x_i$ o $\bar{x}_i$) e a nodi ausiliari che lo forzano ad appartenere alla comunità di T. Affinché il nodo $c_j$ soddisfi la condizione di web-community (avere più vicini interni che esterni), è necessario che almeno uno dei suoi tre vicini letterali si trovi anch'esso nella comunità di T.

**Logica della Riduzione:** Una partizione valida di $G_f$ in due web-communities $\langle C_T, C_F\rangle$ (con $T \in C_T$ e $F \in C_F$) induce un'assegnazione di verità per le variabili di $f$: $x_i$ è 'true' se il nodo $x_i$ è in $C_T$, 'false' altrimenti. Poiché ogni nodo clausola $c_j$ deve trovarsi in $C_T$ e per farlo deve avere almeno un vicino letterale in $C_T$, ciò garantisce che ogni clausola della formula $f$ sia soddisfatta.

Viceversa, se esiste un'assegnazione di verità che soddisfa $f$, è possibile costruire una partizione valida di $G_f$ collocando i nodi letterali nelle comunità di T o F in base all'assegnazione. Poiché ogni clausola è soddisfatta, ogni nodo $c_j$ avrà almeno un vicino letterale nella comunità di T, soddisfacendo così la condizione di web-community. La costruzione del grafo è polinomiale nella dimensione della formula, completando la dimostrazione. □

## 4 Conclusioni

La dispensa illustra come la formalizzazione di un concetto apparentemente intuitivo come quello di "comunità" possa portare a problemi computazionali di diversa natura. Mentre la nozione di cut-community, basata su una proprietà globale, è algoritmicamente trattabile ma non sempre significativa, la definizione più naturale di web-community, basata su proprietà locali, conduce a un problema di partizionamento NP-completo, evidenziando la complessità intrinseca nell'analisi strutturale delle reti.