

Trajectory Clustering Algorithms - GCMP vs DSC vs CTM.

Federico Naldini

Alma Mater Studiorum - Università di Bologna, Cesena.

federico.naldini3@studio.unibo.it

18/10/2019

Framework per l'identificazione di gruppi di oggetti che hanno viaggiato assieme per un certo periodo di tempo.

Si basa sulla divisione dell'area di viaggio degli oggetti in una griglia di celle, per ogni cella vengono raggruppate le traiettorie che sono passate da lì.

A questo punto le celle vengono identificate come le transazioni, le traiettorie come item e tramite l'algoritmo *Apriori* si generano e si restituiscono in output tutti i *frequent closed itemset*.

GCMP

Framework dedicato per il riconoscimento di *Co-movement patterns* in maniera distribuita.

DSC

Framework che, dato un insieme di traiettorie, riconosce e clusterizza le *sub-trajectories* estratte da queste.

CTM

Basandosi su *frequent itemset mining*, individua oggetti che hanno viaggiato assieme per un insieme di istanti non continui.

GCMP

Utilizza un algoritmo di clustering *density-based* o *distance-based*.

DSC

Impiega una variante pesata di *LCSS* che definisce un range spazio-temporale.

CTM

Utilizza un criterio di raggruppamento basato sulla divisione dell'area in cui si muovono gli oggetti in celle.

GCMP

Apriori Enumeration:
consente il pruning
degli insiemi di
oggetti che non
rispettano i criteri di
L-Consecutivness e
G-connection.

DSC

Impiega una variante
pesata di *LCSS* che
definisce un range
spazio-temporale.

CTM

Nativamente ignorato.
Tuttavia è possibile
aggiungere una
dimensione temporale
alle celle, rendendole
di fatto cubi.

GCMP

- **M**: numero minimo di elementi.
- **K**: numero minimo di istanti.
- **L**: lunghezza minima sottosequenze consecutive.
- **G**: massimo intervallo tra un istante e il successivo.

DSC

- ϵ_{sp} : tolleranza spaziale.
- ϵ_t : tolleranza temporale.
- **K**: limite inferiore al voting per rappresentante.
- α : soglia di coesione per i cluster.

CTM

- **MinSize**: numero minimo di elementi per un raggruppamento.
- **MinSup**: minimo di percorsi assieme in x celle.
- **MinCoh**: minimo coesione per un itemset.

GCMP

Unificazione dei campionamenti temporali in scala (una volta sola per dataset).

DSC

Unificazione dei campionamenti temporali in scala.
Costruzione istogramma *equi-depth* sul tempo e partizionamento dei dati in *buckets* basati su questo (una volta sola per dataset).

CTM

Calcolo dell'area in cui si muovono gli oggetti e generazione del reticolo di celle.
Eventuale generazione di una misura univoca per il tempo e divisione in intervalli (una volta sola per dataset).

Comparison: GCMP and CTM = Swarm

Co-Movement pattern in cui i vincoli temporali sono praticamente assenti, rimangono solamente quelli spaziali che possono essere mappati come segue:

$$\mathbf{M} \longleftrightarrow \mathbf{minSize}$$
$$\mathbf{eps} \longleftrightarrow \mathbf{minCoh} \text{ (valido solo in locale)}$$

Rilassando ogni vincolo temporale al massimo, il risultato ottenuto dalla ricerca di un pattern swarm può avvicinarsi molto all'output di *CTM*.

Comparison: DSC and CTM

E' impossibile rilasciare completamente vincolo temporale su DSC, data l'implementazione dell'algoritmo e il parizionamento dei dati in bucket di stessa densità.

Si può provare a impostare una dimensione temporale sulle celle di *CTM* coincidente con l'istogramma individuato per DSC, inoltre per determinare le sottosequenze si può provare a impostare come criterio di partizionamento delle traiettorie un cambio del vicinato.

Tuttavia DSC rimane molto più preciso di CTM e molti itemset individuati da CTM sarebbero invece scartati da DSC(vedi esempio) .

GCMP

- Assenza di pruning sulla base del principio Apriori su criterio spaziale.
- Coesione solamente locale agli snapshot, globalmente viene utilizzato un *density connected criteria*.

DSC

- Poca flessibilità.
- L'idea di fondo è differente.

CTM

- Nessun supporto alla continuità negli istanti temporali.
- Molto legato alla divisione della mappa in celle, raggruppamenti che stanno sui bordi potrebbero non essere riconosciuti.