

Fisica dei Sistemi Complessi

Relazione di progetto

3D Network Visualizator

Giacomo Benvenuti
Enrico D'Angelo
Fabrizio Nuzzo

Sommario

La visualizzazione dei dati è lo studio della rappresentazione visuale dei dati stessi, lo scopo principale è quello di comunicare le informazioni contenute nei dati in maniera chiara ed efficace. Per trasmettere le informazioni sia l'estetica che la funzionalità della visualizzazione devono andare di pari passo, in modo da fornire la comprensione delle informazioni contenute in un dataset altrimenti complesso da comprendere. Lo scopo fondamentale della visualizzazione dei dati è quindi la comunicazione degli aspetti chiave del dataset in una maniera più intuitiva.

Il progetto svolto ha lo scopo di visualizzare le interazioni che avvengono nei diversi istanti tra i nodi di una rete. 3D-Network-Visualizator visualizza queste interazioni in maniera grafica. I nodi che compongono la rete sono disposti in uno spazio tridimensionale secondo le loro coordinate. Le interazioni vengono visualizzate come degli archi che vanno da un nodo all'altro dando una diversa ampiezza all'arco in base alla frequenza dell'interazione, per meglio capire a colpo d'occhio l'importanza relativa delle diverse interazioni. Infine gli istanti scorrono su una linea temporale dando la possibilità di visualizzare come in un video i cambiamenti che avvengono nelle interazioni tra i diversi nodi.

1 Il Problema

L'uomo per sua natura ha una spiccata capacità d'analisi critica ma questa deve essere supportata da una chiara ed appropriata visualizzazione dei dati.

Osservando direttamente un dataset che rappresenta una rete e le interazioni che avvengono tra i suoi nodi, anche se di dimensioni modeste, non è immediato rispondere a domande quali: “quali sono le interazioni che avvengono?”, “quali nodi sono quelli con maggiori o minori interazioni?”, “esistono delle interazioni ricorrenti tra certi nodi?” Che sono solo alcune delle domande di interesse per chiunque si trovi ad analizzare un dataset di questo tipo.

2 3D-Network-Visualizator

3D-Network Visualizator visualizza le interazioni tra i nodi di una rete in modo pratico e intuitivo. I nodi vengono disposti in uno spazio tridimensionale in base alle loro coordinate. La visualizzazione dei diversi istanti di tempo avviene in maniera fluida come se si stesse guardando un video. Le interazioni tra i nodi sono rappresentate con degli archi che collegano coppie di nodi. Le interazioni sono evidenziate con colori differenti, con archi di altezza differente in base alla frequenza dell’interazione che rappresentano. Ogni arco evidenzia anche la direzione dell’interazione e la quantità delle interazioni che avvengono in un dato istante.

3D-Network Visualizator permette quindi di comprendere a colpo d’occhio le caratteristiche delle interazioni. È facile vedere, ad esempio, quali siano le interazioni con maggiore frequenza, o se alcuni nodi hanno interazioni sporadiche o meno.

Il software permette anche di concentrare la visualizzazione su un sottoinsieme dei nodi, evidenziando maggiormente le interazioni che coinvolgono questi nodi e mettendo in secondo piano tutte le altre. La visualizzazione avviene in uno spazio tridimensionale ed è quindi possibile spostare l’angolo di visualizzazione della rete in tutte le dimensioni se questo può aiutare la visualizzazione della rete.

2.1 Specifiche di funzionamento

Le specifiche di funzionamento di 3D-Network Visualizator sono:

- i nodi che compongono la rete vengono posizionati in base alle loro coordinate.
- i nodi possono essere disposti in uno spazio piano o in uno spazio sferico.
- le interazioni tra i nodi vengono visualizzate come archi, l’altezza degli archi è direttamente proporzionale alla frequenza delle interazioni.

- deve essere evidenziata la direzione delle interazioni.
- deve essere evidenziata la quantità delle interazioni.

Per una buona visualizzazione, ed analisi, della rete sono state implementate anche le seguenti specifiche:

- è possibile selezionare un sottoinsieme dei nodi durante la visualizzazione per concentrarsi sulle loro interazioni.
- l'angolo di visualizzazione nello spazio tridimensionale può essere spostato in tutte le dimensioni.
- è possibile variare il tempo impiegato per la visualizzazione di un istante di tempo (ha l'effetto di velocizzare o rallentare l'animazione).

3 Tecnologie utilizzate

Per la realizzazione di questo progetto sono state utilizzate diverse tecnologie:

Processing: è una libreria Java di sviluppo sotto licenza GNU LGPL. La libreria viene incontro a necessità di utilizzo più specifico dal punto di vista grafico. Le molteplici possibilità della libreria combaciano perfettamente con le nostre aspettative di rappresentazione grafica del progetto. Processing viene oggi utilizzato da studenti, artisti, designers, ricercatori e obbisti per produrre prototipi e prodotti finiti di livello professionale. Processing è stato inizialmente sviluppato da Ben Fry e Casey Reas nel 2001 mentre erano entrambi studenti sotto la supervisione di John Maeda al MIT Media Lab. Ulteriore sviluppo di Processing ha avuto luogo all'Interaction Design Institute Ivrea, Carnegie Mellon University, e alla UCLA, dove Reas è capo del Dipartimento di Design — Media Arts. Miami University, Oblong Industries, e la fondazione Rockefeller hanno contribuito finanziando il progetto.

XML: (eXtensible Markup Language). XML è un linguaggio di markup molto pratico per l'elaborazione da parte di un calcolatore, facilmente leggibile da parte dell'uomo, e indipendente dal linguaggio di programmazione usato per elaborarlo. Per questi motivi è stato scelto

Java linguaggio di programmazione orientato agli oggetti, indipendente dalla piattaforma. Grazie all'uso di Java lo stesso codice può essere eseguito su macchine Windows, Linux, MacOSX.

4 Funzioni Implementate

4.1 Creazione di una nuova rete

3D-Network-Visualizator permettere di creare una nuova rete che potrà poi essere visualizzata. In Figura 1 viene riportata l'interfaccia della creazione di una nuova rete alla quale è possibile accedere tramite il menù *File* → *crea una nuova rete*.

The screenshot shows a software window titled "3D-Network-Visualizator" with a menu bar containing "File", "Visualizza", and "About". The main area is a form for creating a new network. It includes the following elements:

- A text input field labeled "Nome della rete:".
- A section labeled "Nodi:" containing a text input field for "Nome del nodo:" and three text input fields for "Coordinata X:", "Coordinata Y:", and "Coordinata Z:".
- A button labeled "Inserisci il nodo".
- A section labeled "Interazioni:" containing a text input field for "Istante:", and two text input fields for "Nodo sorgente:" and "Nodo destinazione:".
- Below the interaction fields are three text input fields for "Tempo:", "Quantita':", and "Frequenza:".
- A button labeled "Inserisci l'interazione".
- A checkbox labeled "I nodi sono posizionati su una sfera".
- At the bottom, there are four buttons: "Crea una nuova rete", "Esporta la rete in formato xml", "Salva la rete non completata su disco", and "Carica una rete non completata da disco".

Figura 1: Form per la creazione di una nuova rete.

- Il campo “Nome della rete” indica sia il nome della rete che il nome del file che sarà prodotto.
- Il campo “Nome del nodo” indica l’etichetta del nodo che stiamo inserendo.
- I campi “coordinata X, Y e Z” indicano ovviamente il valore delle coordinate X, Y e Z.

Una volta inseriti i valori relativi ad uno nodo, cliccando sul pulsante *Inserisci il nodo*, le informazioni relative al nodo verranno memorizzate, i campi torneranno vuoti e sarà possibili inserire i dati relativi ad un altro nodo.

- Il campo “Istante” indica l’identificativo dell’istante a cui si intende aggiungere un’interazione.
- Il campo “Nodo sorgente” indica l’etichetta del nodo dal quale ha inizio l’interazione.
- Il campo “Nodo destinazione” indica l’etichetta del nodo obiettivo dell’interazione.
- Il campo “Tempo” indica l’etichetta dell’istante. Può rappresentare, ad esempio, il tempo reale in cui sono state prese le misurazione.
- Il campo “Quantità” indica la quantità delle interazioni.
- Il campo “Frequenza” indica la frequenza delle interazioni.

Anche in questo caso cliccando sul pulsante *Inserisci l’interazione* tutte le informazioni inserite vengono memorizzate e i campi vengono liberati per poter inserire una nuova interazione.

Il checkbox con etichetta “I nodi sono posizionati su una sfera” indica ovviamente, come dice l’etichetta, se i nodo sono o meno posizionati su una sfera.

Il pulsante *Salva la rete non completata su disco* permette di memorizzare in un file i progressi finora fatti nella creazione di una rete. Non viene generato un file di input per il visualizzatore, questa funzionalità è pensata per permettere all’utente di creare una rete di input in più tempi.

Il pulsante *Carica una rete non completata da disco* permette invece di caricare una rete incompleta, memorizzata precedentemente, e di continuare con l’inserimento dei nodi e delle interazioni.

Una volta che tutti i dati sono stati inseriti si può cliccare sul pulsante *Esporta le rete in formato xml* per creare un file di input per il visualizzatore. A questo punto è possibile visualizzare la rete appena creata.

Il pulsante *Crea una nuova rete* invece serve per annullare tutte le modifiche fatte finora e ricominciare daccapo con la creazione di una nuova rete.

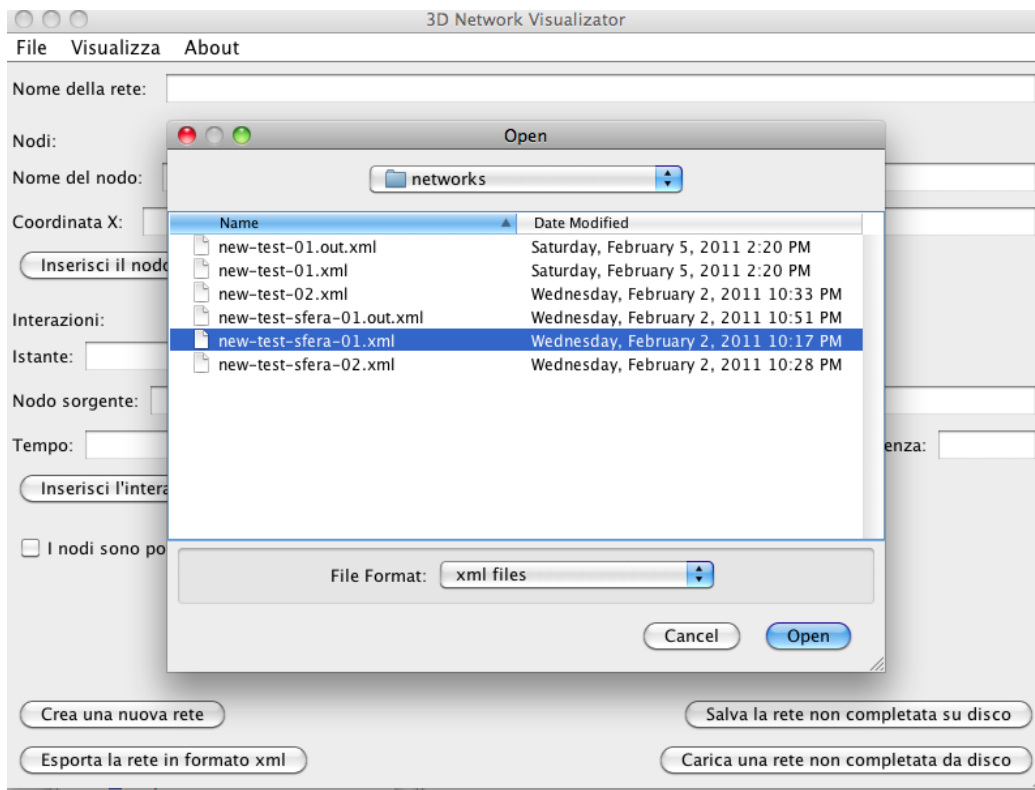


Figura 2: File dialog per importare una nuova rete.

4.2 Importare una rete

Tramite il menu *File* → *Importa una nuova rete* è possibile caricare una file di input che rappresenta una rete e le interazioni tra i suoi nodi.

Una volta importato il file, il programma inizierà a visualizzare la rete con una breve animazione introduttiva, finita l'animazione sarà possibile far partire la simulazione tramite il tasto *play* posizionato nella barra dei comandi (Figura 7).

In qualsiasi momento è possibile mettere in pausa la simulazione tramite il tasto *pause*, in questo modo sarà possibile soffermarsi sulla visualizzazione dell'istante corrente ed analizzare con cura tutte le interazioni.

Uno slider posto di fianco alla barra dei comandi (Figura 7) mostra l'avanzamento della simulazione.

Tramite il campo situato a fianco lo slider è possibile cambiare la durata di visualizzazione di un istante, questa opzione permette di velocizzare o rallentare l'animazione.

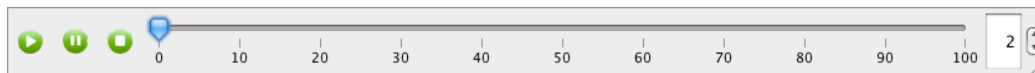


Figura 3: Slider per controllare l’animazione delle interazioni.

4.3 Opzioni di Visualizzazione

Tramite il menu Visualizza è possibile accedere a diverse opzioni di visualizzazioni tra cui: 1. Visualizza spazio 3d di riferimento (attiva di default) tramite questa opzione è possibile nascondere la griglia di riferimento dell’animazione dove sono posizionati i nodi. Questa opzione è utile per ridurre al minimo gli elementi presenti nel riquadro. 2. Visualizza solo archi entranti (disattivata di default) questa opzione è molto utile se vogliamo visualizzare solo le connessioni e il traffico in ingresso; questa opzione è utilizzabile in accoppiata alla funzione partizione in modo da visualizzare solo le connessioni e il traffico in ingresso di determinati nodi. 3. Visualizza gli archi di tutti i nodi (attiva di default) Questa opzione permette di nascondere in qualsiasi istante gli archi della rete, utile per visualizzare la posizione dei nodi sulla griglia; questa opzione è utilizzabile in accoppiata alla funzione partizione.

4.4 Interazione con una partizione della rete

Durante i primi test della nostra applicazione abbiamo notato che era difficile concentrare l’attenzione su una ristretta cerchia di nodi, appunto una partizione della rete. L’attività della rete cambia ogni istante e la possibilità che le connessioni interessino tutti i nodi non è remota. Uno scenario del genere è sicuramente utile per avere un quadro completo della situazione in un dato istante ma se la nostra analisi è rivolta soltanto a determinati nodi tutte le altre interazioni tra i nodi della rete risultano superflue.

Per ovviare a questo problema è stata implementata la funzione partizione; Questa funzione è disponibile in qualsiasi momento dell’animazione; per utilizzarla basta portarsi con il cursore sul nodo/i che ci interessano e selezionarli premendo il tasto “ctrl”.

In questo screenshot possiamo notare come la funzione partizione faciliti l’analisi di questa parte del network formata da 3 nodi nell’istante 6

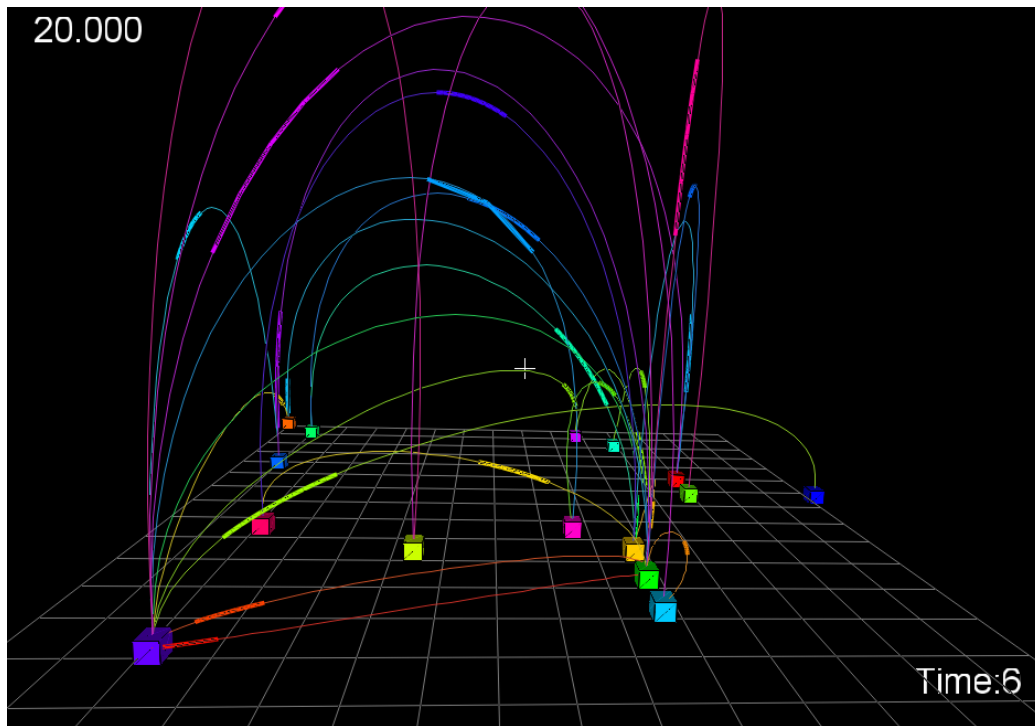


Figura 4: esempio di visualizzazione senza partizione.

4.5 Tipologie di struttura

4.5.1 Piano

4.5.2 Sfera

5 Il dataset

Il tipo di dataset che viene visualizzato contiene tutte le informazioni che riguardano i nodi che compongono la rete, e istante per istante, tutte le interazioni che avvengono tra i nodi. Gli attributi che riguardano i nodi sono:

l'identificativo È un identificativo numerico che viene utilizzato all'interno del dataset per indicare uno specifico nodo. Ogni nodo possiede un identificativo univoco.

l'etichetta L'etichetta svolge circa lo stesso ruolo dell'identificativo ma è più adatta per essere memorizzata dalle persone. gli identificativi servono al software per distinguere e identificare i diversi nodi, per questo

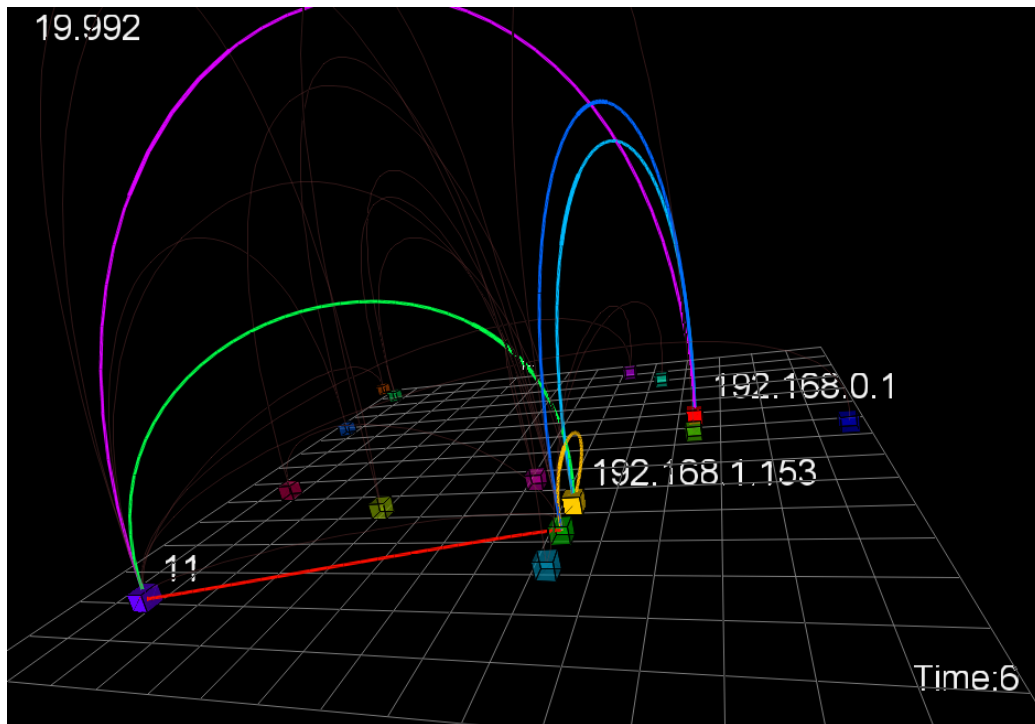


Figura 5: esempio di visualizzazione con partizione.

motivo sono degli indici numerici. Per una persona è invece molto più facile memorizzare degli identificativi alfanumerici, e questo è lo scopo delle etichette. Inoltre quando i nodi rappresentano delle entità fisiche esistenti, come ad esempio delle città, è utile che le label rispecchino il nome di tale entità al fine di aiutare la comprensione del fenomeno. L'etichetta di un nodo non deve necessariamente essere univoca.

coordinate Sono le coordinate cartesiane dei nodi.

Gli attributi che riguardano le interazioni sono:

la sorgente È l'identificativo del nodo da cui ha origine l'interazione.

l'obiettivo È l'identificativo del nodo obiettivo dell'interazione.

quantità Rappresenta la quantità del flusso che costituisce l'interazione. Cosa sia effettivamente il flusso che costituisce l'interazione dipende dal tipo di rete in esame. Può per esempio trattarsi di pacchetti scambiati da diversi elaboratori collegati via rete, o di automobili che si spostano da un luogo ad un altro.

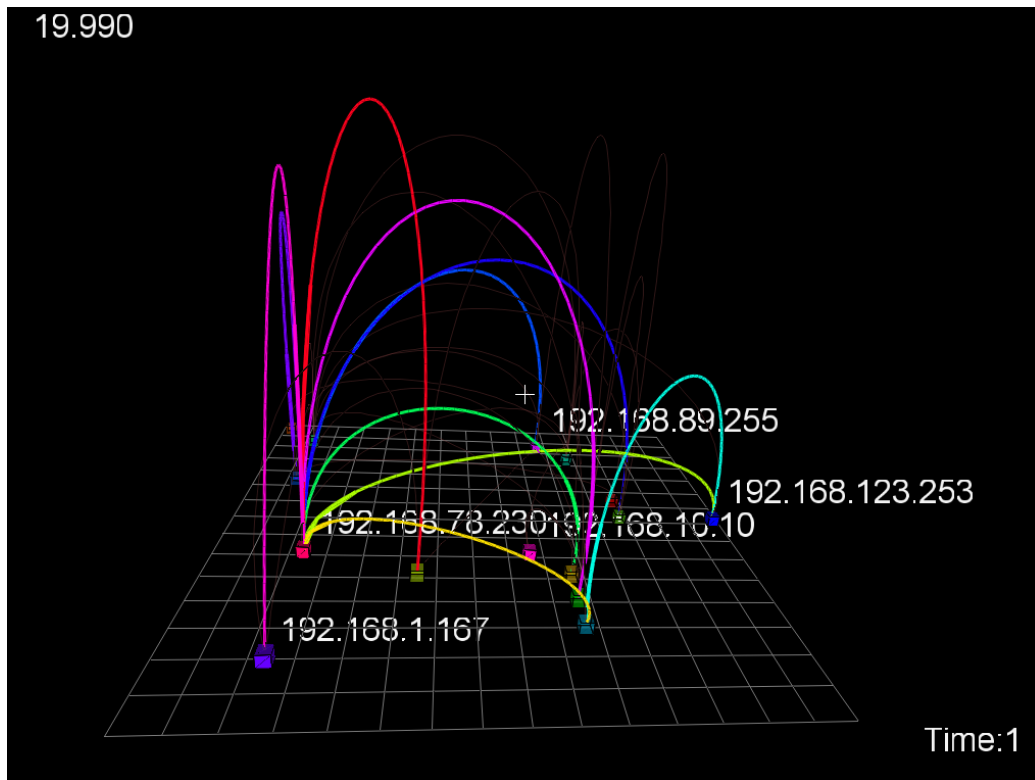


Figura 6: esempio di visualizzazione con partizione.

frequenza È la frequenza che caratterizza l'interazione.

Le informazioni che riguardano le interazioni sono naturalmente suddivise per istanti. Anche gli istanti hanno degli attributi:

valore È l'identificativo dell'istante. Gli identificativi degli istanti devono essere univoci e consecutivi.

etichetta Anche in questo caso le etichette servono per meglio comprendere il fenomeno. Le etichette possono infatti rispecchiare il tempo fisico, l'ora, in cui sono state prese le misurazioni che costituiscono le informazioni riguardanti le interazioni.

5.1 Il file network XML

L'input del nostro visualizzatore è un file che contiene la descrizione della topologia di una rete e le interazione che avvengono tra i suoi nodi. Come

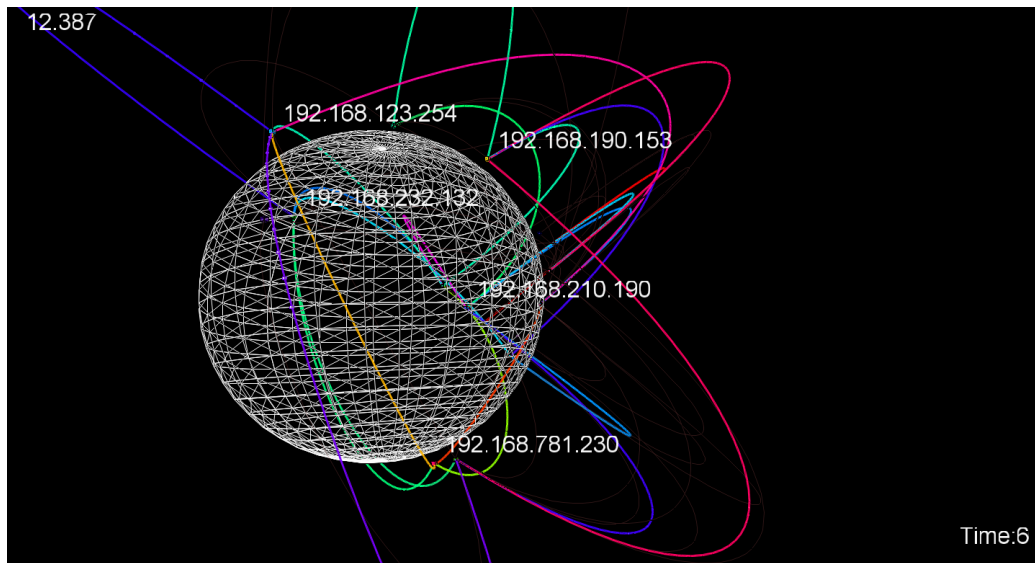


Figura 7: esempio di visualizzazione con partizione.

formato del file di input abbiamo scelto l'XML per diversi motivi. In primo luogo per la facilità con cui può essere processato da un calcolatore. In secondo luogo per la sua portabilità su diverse piattaforme e linguaggi di programmazione, in questo modo lo stesso file di input può essere utilizzato facilmente anche da altri software scritti in altri linguaggi di programmazione per altre piattaforme software. Infine per la sua caratteristica di essere facilmente leggibile anche dall'uomo. Il principale punto a sfavore di questo formato è la scarsa compattezza, nel senso che oltre alle informazioni relative alla rete il file contiene anche molte meta-informazioni che servono appunto per facilitare l'elaborazione del file e facilitarne la sua leggibilità. Questo significa che per file di input che rappresentano reti di grandi dimensioni, ad esempio con migliaia di nodi e centinaia di interazioni, molto spazio per la memorizzazione del file su disco sarà occupato dalle meta-informazioni. In questi casi un formato binario sarebbe stato sicuramente più compatto, ma avrebbe perso tutte le caratteristiche positive elencate sopra.

Qui di seguito viene riportato un file di input di esempio:

```
<?xml version="1.0" encoding="UTF-8"?>
<network>
  <static-data>
    <network-name>new-test-02</network-name>
    <nodes-list>
      <node id="0" label="192.168.0.1" x="2" y="0" z="6" />
```

```

    <node id="1" label="192.168.0.2" x="3" y="0" z="8" />
  </nodes-list>
  <flat>true</flat>
</static-data>
<dynamic-data>
  <instant value="0" label="0">
    <interaction source="1" target="0" quantity="1" frequency="0.6080425" />
  </instant>
  <instant value="1" label="1">
    <interaction source="1" target="0" quantity="8" frequency="1.31866568" />
  </instant>
</dynamic-data>
</network>

```

La struttura del file di input è divisa in due parti principali:

static-data: Questa parte contiene le informazioni statiche della rete, cioè la sua topologia, è composta dai seguenti elementi:

- network-name: il nome della rete
- nodes-list: l'elenco dei nodi presenti nella rete
- flat: indica se la rete deve essere visualizzata in uno spazio piano o sferico

Ogni nodo in nodes-list è descritto con i seguenti attributi:

- id: è l'identificativo univoco del nodo
- label: è l'etichetta del nodo
- x, y e z: sono le coordinate del nodo.

dynamic-data: Questa parte contiene tutte le informazioni dinamiche della rete, cioè le interazioni che avvengono tra i suoi nodi. È strutturata come una lista di istanti. Ogni istante è strutturato come una lista di interazioni. Ogni interazione è descritta con i seguenti attributi:

- source: è l'identificativo del nodo dal quale parte l'interazione
- target: è l'identificativo del nodo obiettivo dell'interazione
- quantity: rappresenta la quantità di dati scambiati durante l'interazione
- frequency: è la frequenza dell'interazione

6 Conclusioni

Lo studio e l'implementazione del programma 3D Network Visualizator ci ha dato la possibilità di scoprire i vantaggi derivati da una efficace ed adatta visualizzazione dei dati. Spesso per effettuare analisi non basta avere statistiche dettagliate e indici ricavati da grandi quantità di dati per il semplice motivo che viene a mancare una visione d'insieme che potrebbe portare un risultato d'analisi inaspettato. Inizialmente il nostro dominio applicativo erano le reti di calcolatori dislocate geograficamente che scambiano pacchetti di dati. Il dominio in questione per quanto può essere banalmente descrivibile per esempio tramite un grafico bidimensionale, esclude alcune informazioni che potrebbero essere importanti in caso di analisi approfondite. Una fra tutte la frequenza con cui i nostri ipotetici calcolatori si scambiano pacchetti; come rendere visibile questa informazione? L'introduzione della terza dimensione è stata essenziale per poter aggiungere elementi. Prendendo in esempio la frequenza con cui i nodi si scambiano i pacchetti, la nostra soluzione è stata di curvare il link in modo da creare un arco in relazione all'aumento della frequenza di comunicazione. Questo tipo di approccio è risultato di sicuro impatto per avere un'idea dei nodi che hanno un scambio maggiore dei pacchetti nella rete in un dato istante.