



Data Analytics
Analisi MMORPG Travian^[1]
1819-2-F1801Q104, Data Analytics, 2018-2019
Università degli Studi di Milano Bicocca
Dipartimento di Informatica, Sistemistica e Comunicazione

Caselli Alex 807129
Olivadese Fabrizio 820864
Villa Giacomo 807462

8 Luglio 2019

Indice

1 Abstract	4
2 Travian	5
2.1 Che cos'è Travian?	5
2.2 Meccaniche di gioco	5
2.3 Alleanze e relazioni tra utenti	6
2.4 Il mondo di gioco	7
3 Dataset e Preprocessing	8
3.1 I dataset	8
3.2 Preprocessing	9
3.2.1 Grafi	9
3.2.2 Alleanze	10
4 Obiettivi di Analisi	12
4.1 Introduzione	12
4.2 Tempo di Attività	12
4.3 Stato del server	12
4.4 Trickster Detection	13
4.5 Commercio e Messaggi	13
4.6 Studio singola community	13
5 Analisi e Risultati	14
5.1 Introduzione	14
5.2 La rete	14
5.2.1 Orario attività degli utenti	14
5.2.2 Trend Nodi e Archi	15
5.2.3 Communities Trend	17
5.2.4 Grado Medio e Distribuzione di Grado	19
5.2.5 Average Shortest Path e Diametro	22
5.3 Trickster Detection Tool	24
5.3.1 Introduzione	24
5.3.2 Logica di funzionamento	24
5.4 Commercio e Messaggi	26
5.4.1 Introduzione	26
5.4.2 Reciprocità	26
5.4.3 Densità	27
5.4.4 Clustering Coefficient	28
5.4.5 Diametro e Shortest Path	29
5.4.6 Average Degree	30
5.4.7 In and Out Degree Distribution	30
5.5 Studio Singola Community	31
5.5.1 Introduzioni	31
5.5.2 User e Edge Trend	32
5.5.3 Analisi Strutturale 1	33

5.5.4	Analisi Strutturale 2	34
5.5.5	Ricerca dei leader	36
5.5.6	Relazioni con i diplomatici	40
5.5.7	Eliminazione dei leader	41
6	Possibili sviluppi futuri	43
	Bibliografia	44

1 Abstract

Le **relazioni sociali** tra esseri umani sono rappresentabili mediante tre macro-aree:

- Ostili
- Comunicazione
- Scambio

Ogni essere umano si rapporta ad un altro seguendo una o più di queste tipologie di relazioni, le quali possono anche evolvere e cambiare nel tempo. Un essere umano appartiene anche a un determinato **gruppo sociale**, sia questo una squadra, una compagnia, un gruppo di colleghi. Le relazioni all'interno di ogni gruppo sociale (identificabile come community) sono tendenzialmente positive (comunicazione e scambio).

Lo studio e l'analisi della rete di relazione di un gioco MMORPG (Massive Multiplayer Online Role-Playing Game) può essere un utile punto di partenza per capire e comprendere in che modo le relazioni tra players (essere umani) si strutturano ed evolvono. I gruppi sociali, in questo contesto, sono rappresentati dalle **alleanze** tra giocatori; lo studio di come un player si muove a livello di relazione intra ed extra comunitarie può risultare un utile parametro di studio.

Guidati da questa logica, ci siamo mossi in uno studio della rete in nostro possesso al fine di **studiare** il rapporto tra i player e lo stato del server durante il periodo temporale di interesse.

2 Travian

2.1 Che cos'è Travian?

Travian è un **videogioco per browser** di tipo Massively multiplayer online game, prevalentemente gratuito, ispirato al gioco da tavolo Coloni di Catan.

Il gioco, di origine tedesca è stato pubblicato in Italia il 26 ottobre **2005** ed attualmente ha server di gioco in Germania, Spagna, Paesi Bassi, Italia, Francia, Polonia, Portogallo, Brasile, Argentina, Cile, Messico, Turchia, Russia, Ucraina, Lituania, Bulgaria, Romania, Repubblica Ceca, Slovacchia, Ungheria, Slovenia, Bosnia ed Erzegovina, Croazia, Serbia, Cina, Hong Kong, Danimarca, Svezia, Finlandia, Norvegia, USA, Regno Unito, Emirati Arabi Uniti e Austria. Esistono anche alcuni server internazionali.

Tale gioco di strategia militare è ambientato in un mondo immaginario popolato da tribù di Romani, Galli e Teutoni. Ogni giocatore può scegliere liberamente uno di questi popoli; la scelta condiziona pesantemente la strategia all'interno del gioco perché ogni civiltà ha caratteristiche diverse. Nel corso del gioco non è possibile cambiare la tribù scelta^[2].

2.2 Meccaniche di gioco

Ogni giocatore parte con un **villaggio** in un determinato punto della mappa (solitamente alle periferie), nel villaggio sono disponibili varie strutture le quali permettono tutta una serie di operazioni che spaziano dal reclutare **eserciti**, aumentare la produzione di **risorse**, aumentare i bonus difesa, permettere il **commercio** e molto altro. (guida agli edifici)



Figura 1: Possibile villaggio di livello medio/alto

L'obiettivo è aumentare di livello gli edifici al fine di diventare competitivi dato il punteggio, inoltre sarà necessario reclutare un buon esercito per difendere e **attaccare** altri players e **conquistare** altri villaggi; infatti un player può avere sotto il suo controllo diversi villaggi. Per conquistare altri villaggi è necessario reclutare una speciale truppa (Senatore nel caso dei Romani), la cui abilità è abbassare la lealtà del villaggio attaccato, una volta a zero il

villaggio passa all'attaccante. Tale truppa è molto costosa sia in termini di tempo che in termini di risorse, pertanto è molto preziosa.

Dati scontri tra players di una certa esperienza, i vettori di attacchi possono spaziare da un numero minimo di 10/15 fino anche a 100 **differenti attacchi**; in questa documentazione non si entrerà nel dettaglio delle strategie offensive/difensive.

2.3 Alleanze e relazioni tra utenti

Le relazioni tra gli utenti, tolto l'attacco, possono essere di due tipologie:

- **Commercio**
- **Messaggio**

Il commercio è possibile mediante un edificio (il mercato), il messaggio può invece essere inviato indipendentemente da qualsiasi fattore. Gli utenti possono altresì unirsi in **alleanze** per fornirsi aiuto reciproco (unità militare e risorse) ed essere dunque più **competitivi**; all'interno di ogni Alleanza vi sono delle figure più centrali, le quali sono a libera discrezione della specifica Alleanza, possiamo però identificarle nelle seguenti figure:

- **Capo/i e vice-capo/i**: individui di rilievo, guidano l'Alleanza e sono simbolo dell'unità della stessa, spesso fanno parte dei players più forti data l'Alleanza.
- **Diplomatici/ambasciatori**: individui che dialogano con esponenti di altre Alleanza per fini pacifici o meno, hanno un canale preferenziale con i capi della proprio Alleanza.
- **Reclutatori**: individui che si occupano di reclutare nuovi player da inserire nella Alleanza.

Vi possono poi essere una miriade di possibili figure istituzionali data un Alleanza (generale della difesa, generale dell'attacco, moderatore forum ...).

Dettagli		Posizione	
Tag:	SPARTA	Gen. Difesa	Fabbro
Nome:	SPARTANI	Consigliere	Jekyll&Hyde
Posizione in classifica	2	Capo & Diplomatico	Marcoo
Punti	13944	Capo & Reclutatore	Petrus
Membri	48	Capo & Sitter/Forum	Daniele
» Val al Forum		Gen. Attacco	Bellal

Figura 2: Tipica pagina di un alleanza

Gli utenti che fanno parte di una stessa alleanza, sono dotati di un **forum** nel quale poter aprire vari topic e comunicare (un ulteriore strumento oltre al tradizionale messaggio).

2.4 Il mondo di gioco

Il mondo è strutturato come una scacchiera dove ogni villaggio è posizionato in una **cella**. Il tempo di attraversamento di ogni cella dipende dalle unità che viaggiano, alcune risultano essere più veloci (i cavalieri) e altre più lente (catapulte); il tempo di viaggio verrà vincolato dall'unità più lenta.



Figura 3: Mappa di gioco

Il tempo di attraversamento, ma più in generale i **tempi** del gioco, dipendendo dal **tipo di server**. Possono esserci server più tradizionale, con tempi di attesa lunghi e i cosiddetti **speed server** dove i tempi sono notevolmente ridimensionati come il mondo da noi analizzato.

3 Dataset e Preprocessing

3.1 I dataset

Per la rappresentazione della rete eravamo dotati di **30 dataset** per tipologia di interazione utente; dunque, dati attacchi, messaggi e commercio, possediamo un totale di un **90 file .graphml**. In ogni file, dato il giorno, sono presenti **solo gli utenti che in quel giorno hanno effettuato l'operazione** dato il file; è per tanto possibile che, dato un player il quale ha solo effettuato commercio, questo sia presente solo in un file, dato il giorno.

Gli attributi in nostro possesso:

- **Utente**: valore `id` univoco per ogni utente, noi abbiamo ipotizzato sia sequenziale e che quindi potesse fornire informazioni relative all'anzianità dell'utente sul server.
- **Archi**: attributo `timestamp` in formato `unix time` che denota quindi, dal giorno fino al secondo, il momento dell'interazione; inoltre possiamo considerare la presenza di un attributo auto imposto dalla suddivisione in file, ovvero una label identificante la tipologia di arco.

Il numero di archi e nodi non viene qui indicato in quanto oggetto di analisi e presente nelle sessioni successive. I vari giorni erano rappresentati mediante una struttura **multigrafo**, di conseguenza veniva modellato con un singolo e indipendente arco ogni interazione dati due players; ovviamente al crescere di interazioni cresceva il numero di archi.

Non siamo, purtroppo, dotati di informazioni riguardanti:

- **Punteggio utente**: denota il punteggio in funzione al numero di villaggi e al livello degli edifici.
- **Numero di villaggi**: dato un utente il numero di villaggi sotto il suo possesso.
- **Contenuto messaggi/numero di risorse/numerosità esercito**: data un'interazione sappiamo solo che questa ha avuto luogo.

I dati mancanti potevano essere di aiuto in funzione ad uno studio più approfondito della rete.

Siamo infine dotati di 30 file `.txt` i quali presentano, per ogni riga, gli identificativi degli utenti facenti parte di un'**alleanza**; il problema di questi file riguarda il fatto che tengono nota solo delle alleanze **presenti** e non di quelle dissolte.

Vi è dunque un problema di allineamento e **consistenza**, l'Alleanza 128 del primo giorno non sarà mai l'Alleanza 128 dell'ultimo giorno, per questo dataset è stata necessaria un'operazione di pre-processing importante al fine di tener traccia delle Alleanze dissolte. Alla fine abbiamo ottenuto un allineamento totale, come verrà spiegato nella prossima sezione.

3.2 Preprocessing

3.2.1 Grafi

Per quanto riguarda i grafi si è effettuato un'operazione identica per ognuno di questi: **compattare gli archi** al fine di trasformare il multigrafo in un grafo. Utilizzando in questa prima fase la libreria **Networkx** si è provveduto a iterare sugli archi e a riempire un **dizionario** di dizionari così strutturato:

```
{ FromPlayerId: { ToPlayerId: NTimes } }
```

Teniamo dunque traccia, per ogni player, le relazioni con altri utenti (fortemente direzionali) e il numero di volte in cui queste avvengono; ovviamente questa politica causa una modifica degli attributi degli archi. In seguito alla creazione della struttura sopracitata si è provveduto ad **iterarla** e quindi a creare gli archi i quali saranno, ora, così strutturati:

- **Label**: indica il tipo di arco (M, A, T: Messaggio, Attacco, Trade).
- **Weight**: indica il peso dell'arco, ovvero il numero di volte in cui questo è presente dato il giorno.

Dunque alla fine avremo sempre **90 file graphml**, però **compattati** e dunque con un numero inferiore di archi; ovviamente il numero di nodi sarà sempre lo stesso.

In seguito si è passata alla creazione di ulteriori due reti per ogni giorno:

- **Rete Alleanza**: rete contenente solo gli archi di messaggi e trade.
- **Rete Totale**: rete contenente attacchi, messaggi e trade.

Per la prima è stata comunque creato un grafo, questo è stato possibile mediante l'aggiunta di alcuni attributi sugli archi. Gli archi sono quindi dotati di:

- **MWeight**: numero di messaggi scambiati tra i due players.
- **TWeight**: numero di trade effettuati tra i due players.

Il procedimento di costruzione è stato simile a quanto già fatto, quindi si crea e riempie un vocabolario di vocabolari dal quale poi si ottiene il grafo. È stato ovviamente verificato, in maniera approfondita, che il numero di archi corrispondesse alla somma degli archi dei grafi di partenza; per il numero di nodi ovviamente non si tratta di una semplice somma in quanto due utenti possono aver scambiato messaggi e aver commerciato nello stesso giorno.

In ultimo, per quanto riguarda i grafi, è stato creata la **rete totale** contenente tutti gli archi; ancora una volta il procedimento di creazione è quanto sopracitato. In questo caso abbiamo mantenuto la struttura multigrafo in quanto utile per possibili visualizzazioni future al fine di differenziare gli archi in base alla label (m, t, a), che in questa fase risulta quindi fondamentale per differenziare gli archi tra di loro.

Ovviamente ad ogni iterazione, dato il giorno, si salvavano le nuove strutture al fine di poterle utilizzare successivamente.

Infine, per permettere un corretto utilizzo delle strutture definite, è stato implementato un dizionario con il compito di mantenere i grafici data la giornata. La strutturazione di questo è dunque la seguente:

```
{ Day: (attackGraph, messageGraph, tradeGraph, allianceGraph, totalGraph) }
```

Per ogni giorno una lista contenente i grafi data la giornata, da questa poi verranno lanciate tutte le **analisi** che abbiamo condotto sulla rete.

3.2.2 Alleanze

Il lavoro di preprocessing per quanto riguarda le alleanze è stato, forse, uno dei task più complessi dell'intero progetto. I file .txt erano strutturati in maniera tale che, ad ogni riga, corrispondesse un'alleanza identificata mediante gli **id** degli utenti che vi facevano parte; il problema è costituito e causato dal fatto che se un'alleanza, durante il periodo, si **scioglie** questa viene semplicemente **cancellata** dal file, comportando un riduzione della dimensione del file.

Il problema è dunque questo “non allineamento” che vincolava ed impossibilitava tutta una serie di analisi: l'alleanza 118 al giorno 1 non sarà mai l'alleanza 118 al giorno 30. I concetti alla base del nostro ragionamento, che poi ha portato alla creazione dei file allineati, sono i seguenti:

- Un'alleanza che vede il 100% dei suoi componenti del giorno precedente non presenti al giorno considerato è da ritenersi **dissolta**.
- Un'alleanza che vede anche solo un componente presente al giorno considerato nella riga della stessa, **non è** da ritenersi **dissolta**.
- Le alleanze dissolte continuano ad esserlo nei giorni a venire, quindi la riga di queste continuerà a contenere la **stringa DISSOLVED**.
- Le alleanze che si **separano**, creando nuove alleanze e/o unendosi ad altre preesistenti causeranno lo scioglimento dell'alleanza (nel caso in cui nessun membro sia presente) e potranno comportare la creazione di nuove.
- Le nuove alleanze che verranno create nel tempo saranno **aggiunte** alla fine del file della giornata.

A livello di codice si valuta a “coppie” i file dunque, dati i numeri dei giorni, 1 con 2, 2 con 3, 3 con 4 e così via. In questa maniera le modifiche dato il giorno precedente si propagano sul giorno immediatamente successivo. Passiamo quindi a inserire, per ogni giorno date le coppie definite, le righe “splittate” in dizionari la cui chiave rappresenta una sorta di identificativo (sostanzialmente il numero della riga) di quella alleanza il quale sarà costante per tutta la durata del processo. La struttura è la seguente:

```
{ CommunityId: (IdPlayer1, IdPlayer2, ..., IdPlayerN) }
```

A questo punto performiamo un ciclo **for** su 4 livelli in quanto iteriamo le alleanze del giorno precedente, quelle del giorno attuale e per ognuna di queste iteriamo sui componenti delle alleanze date i due giorni; andiamo quindi a salvare il numero di utenti che

osservo al giorno odierno in funzione dell'alleanza in cui si trovano così da poter ricavare la percentuale di quelli presenti nella comunità di partenza del giorno precedente. A questo punto:

- Se la comunità del giorno precedente non ha nemmeno un utente questa viene dichiarata dissolta.
- Se la comunità del giorno precedente ha dei membri, questa verrà salvata nella posizione corretta con i membri del giorno che si sta considerando (potrebbero essercene di nuovi o di meno).

Ovviamente in funzione delle comunità dissolte si memorizzerà un valore che **quantifica** il numero di queste al fine di poter allineare il file del giorno odierno con quello del giorno precedente.

Una volta terminati i cicli **for** si andrà a verificare l'eventuale esistenza di **nuove comunità** procedendo poi con l'inserimento di queste in fondo alla file. Il risultato finale è dunque il seguente:

1	1 2
2	2002 2132 5309 4025 2860 1108 3440 896
3	2026 8951 2939 1694 9638 2857 5126 1920 7638 1470 2085 7846
4	1048
5	7595 6513 10085 8098 1640 10508 5658 2372 2622 7713 10568
6	1961 2600 11533 284 10460 7712 4245 2347 6690 3523 4298 862
7	3806
8	4507 7617 2263 3709 2862 4415 10258 4165 1420 7303 9703
9	5234 787 1129 804 4510 965 725 4370 805 1156 4403 4499 750
10	408 684 32 136 354 128 476 296 3250 1368 781 380 196 1104 1
11	2585 10676 8610 1283 3354 1676 11919 9795 6771 3816 9990 97
12	242
13	2833 4420 183 3413 2851 7994 418 1816 922 469 2194 4240
14	2970 1005
15	767 897 1022 449 2284 930 2314 2742 4775 384 994 1499 986 1
16	1845 6225 10775 5727 3518 4110 5630 10390 5693 4063 1384 93
17	2577 7475 1871 2769 11275 2223 10816 1490 10298 10277 10419
18	4004
19	2746 3044 3587 2640 2594 2679 10883 2110 2749 2984
	1 2
	1108
	1508 2939 5050 1694 1337 805 2857 7638 1470 7846 2135 653
	1048
	51 7595 2436 6431 8596 6513 10085 1640 552 7713 2035
	2600 12250 284 6010 7712 4245 7963 12055 2347 3523 4298 862
	DISSOLVED
	2263 3709 4165 7303
	5841 1156 4499 4696 7271 2085 4811 6840 4585 3388 6868 750
	408 32 128 476 296 11275 12371 380 10570 1857 62 9588 43
	2585 10676 8610 12230 1283 11851 11919 9795 6771 3816 9703
	DISSOLVED
	2833 4420 183 3413 2851 7994 418 1816 922
	2970
	DISSOLVED
	1845 1384 9355
	3778 4003 1490 2577 7475 9486 2769 2223 12162 10419 5629
	DISSOLVED
	2746 2640 2594 2110

Figura 4: Prime 19 alleanze del giorno 1 e del giorno 30

Sono stati effettuati **controlli** per quanto riguarda la veridicità delle azioni, in funzione dei membri delle alleanze e del numero di alleanze in funzione delle disciolte nei vari file. Si è deciso di mantenere nota di tutte le alleanze, indipendentemente dal numero di players di queste in quanto, come verrà spiegato successivamente, si è ritenuto importante performare un'analisi sulle comunità **disciolte** al fine di comprenderne la motivazione; eliminare dei dati non è stata ritenuta una corretta operazione.

La funzione che effettua questa operazione è denominata **Community reorganization in the period**, ed è presente nel notebook consegnato insieme a questa documentazione.

4 Obiettivi di Analisi

4.1 Introduzione

In questa sezione elenchiamo le motivazioni che stanno dietro alle **analisi** da noi effettuate, abbiamo realizzato vari studi che spaziano da analisi più strettamente legate al gioco e alle sue dinamiche per poi arrivare ad analisi più teoriche i cui risultati si è cercato di riadattare data la realtà di studio.

4.2 Tempo di Attività

Durante la fase di **preprocessing** sui grafi è stato **eliminato** l'attributo **timestamp** in quanto privo di significato data l'operazione di compattazione degli archi; tuttavia prima di effettuare questa è stata effettuata un'analisi sull'**orario** di attività dati i vari archi, sostanzialmente l'orario di questi.

Riteniamo utile comprendere l'**orario medio di punta** di attività (tenendo dunque in considerazione tutte quelle possibili), questo perché potrebbe rappresentare un momento per inserire pubblicità data la natura gratuita del titolo (tolti abbonamenti e bonus acquistabili con valuta in game).

4.3 Stato del server

Un ulteriore analisi da noi condotta riguarda lo **studio della numerosità degli archi e dei nodi** sull'intero periodo a nostra disposizione, in questo modo cerchiamo di dedurre lo stato del server e se questo risulta essere in una fase attiva con una crescita di utenti e azioni o in una fase passiva e di decrescita.

Dato il paper qui citato^[1], siamo consapevoli della natura **speed** del server, il che ci porta a pensare che comprendere il trend di nodi e archi possa essere semplice ed immediato. Abbiamo ritenuto anche utile comprendere il numero dei villaggi “barbari” ovvero di quei villaggi presenti nel server ma non appartenenti di fatto a nessun player; utilizzati principalmente per il saccheggio e per una facile conquista.

Un ulteriori analisi appartenente a tale famiglia riguarda lo studio del **trend delle community** al fine di identificare la numerosità di queste in termini di presenza, creazione e dissoluzione. Si è inoltre studiato, dato il fenomeno della dissoluzione, la numerosità di queste nel periodo e la componente in termini di utenti con supposizioni riguardanti gli stessi.

Siamo poi passati alle classiche, e forse anche più “tecniche”, analisi riguardanti **Average Degree, In/Out Degree Distribution, Average Shortest Path, Diameter Value**. Queste analisi, volte allo studio di ogni singola tipologia di interazione, hanno contribuito a fornire informazioni tecniche e strutturale data la rete.

4.4 Trickster Detection

Una grande problematica dei giochi appartenenti alla famiglia **MMORPG** riguarda il **multi-accounting**: stessa persona che gestisce due o più profili al fine di avvantaggiarne un altro (solitamente il suo/i main). Travian, così come altri giochi simili, ha meccanismi molto sviluppati in tal senso; ovviamente questi non sono di dominio pubblico. Abbiamo cercato, date una serie di considerazioni, di fornire un **tool** che potesse identificare comportamenti scorretti in tal senso.

4.5 Commercio e Messaggi

Influenzati dal fatto che le azioni sopracitate fossero di numerosità inferiore rispetto alla sola interazione offensiva, ci siamo prodigati nello studio di queste a livello di **community**. Abbiamo quindi valutato solo i membri allineati in quanto abbiamo ipotizzato che questi potessero essere più attivi e fornire più informazioni.

A questo punto abbiamo effettuato analisi di **reciprocità** e **densità** date le relazioni intra e extra community; siamo passati poi all'analisi del **clustering coefficient** per giungere infine alle già sopracitate analisi di **numerosità degli archi**, **Diametro**, **Average Shortest Path**, **Average Degree** e **In/Out Degree distribution**

4.6 Studio singola community

Al fine di comprendere meglio le dinamiche intra-community abbiamo deciso di concentrare il nostro studio su di una sola, nel tentativo di poterla poi confrontare con l'andamento medio precedentemente valutato; ma non solo: siamo interessati anche all'attività di ricerca dei leader e di fiancheggiatori di questi nel tentativo di dare una forma alla catena di comando di una data community.

Ci siamo dunque preoccupati, valutando relazioni di commercio e messaggi, di effettuare un'**analisi legata alla numerosità dei membri** e all'attività intra-community di questi, dunque se questi avevano interazioni interne, in funzione anche di quanti membri risultassero legati a quella community. Siamo poi passati alle analisi più strutturali con l'obiettivo di confrontare i risultati con quelli ottenuti precedentemente con la valutazione di tutte le attività intra-community; abbia perciò valutato **reciprocità**, **densità**, **clustering coefficient** e tutte quelle misure già citate nella sezione **4.5 Commercio e Messaggi** di questa documentazione.

Siamo poi passati all'analisi più interessante riguardante la leader detection, valutando i valori ottenuti con le varie misure di centralità studiate abbiamo cercato di identificare i potenziali leader dell'alleanza, per poi passare all'identificazione dei diplomatici ed infine abbiamo valutato le relazioni a livello di messaggi con gli utenti identificati come leader in funzione a tutti gli altri membri della community che non sono stati ritenuti centrali. In ultimo abbiamo provato ad eliminare gli utenti identificati come leader e abbiamo visualizzato nuovamente la community analizzando, a livello prettamente visivo, le significative **differenze**.

5 Analisi e Risultati

5.1 Introduzione

In questa sezione descriveremo come sono state effettuate le analisi in termini **tecnic**i avvicinandoci quindi al codice e alle logiche più strettamente informatiche per poi passare, passo dopo passo, allo **risultato** e allo studio di questo data la realtà di interesse.

5.2 La rete

Sulla rete, divisa per poi per tipologia di interazioni, abbiamo effettuato analisi di **attività**, **trend** nodi, archi, community e studio dei principali **indicatori** presentati durante il corso; procediamo con ordine.

5.2.1 Orario attività degli utenti

Ogni arco, prima della fase di preprocessing, era dotato di un attributo che indicava mediante tempo Unix il momento in cui l'attività denotata dallo stesso veniva effettuata. Abbiamo quindi provveduto prima della **cancellazione** di tale attributo a tener nota dell'ora di attività, valutando dunque le **24 ore** di una giornata; in tal senso il codice contenuto nei blocchi che prendono il nome di *Create compact <tipology> graph* andrà a incrementale il valore di una lista contenuta in un dizionario così strutturato:

```
{ Hour: (nAttack, nTrade, nMessage, nTotal) }
```

Il dizionario così creato verrà poi inserito, dato il Main, verrà poi trattato dalla funzione **saveActivityHour()** la quale effettua una media sui 30 giorni per tutte le attività sopracitate. I risultati poi verranno salvati su di un **file** che dal quale abbiamo estrapolato i seguenti grafici:

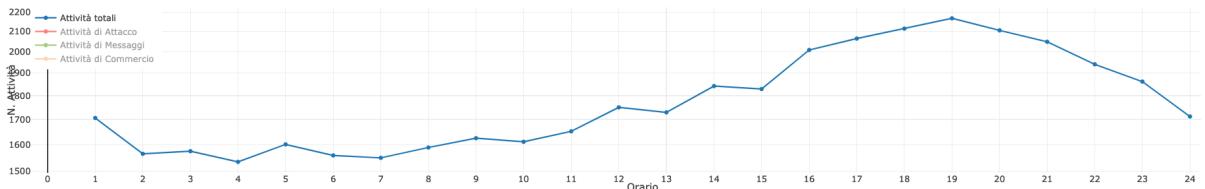


Figura 5: Media sui 30 giorni dell'orario di attività dei player

Abbiamo effettuato una media in quanto abbiamo considerato più utile e indicativo fornire un **andamento medio**, soprattutto nell'ottica di vendere tale informazione in funzione di un inserimento di eventuali banner pubblicitari all'interno del gioco. Possiamo quindi concludere che l'orario di **maggior attività si registra dalle ore 16:00 alle ore 21:00**, con valore di picco alle ore 19.00. Orario al quale gli adolescenti terminano i compiti e i lavoratori rientrano a casa e possono dedicare tempo al gioco.

5.2.2 Trend Nodi e Archi

Nel tentativo di comprendere il trend di utenti e di archi nel grafo ci siamo concentrati sullo studio della loro **numerosità** in funzione del **giorno**. Il numero di archi, di qualsiasi tipologia, tiene conto dell'effettiva numerosità iniziale, misurazione effettuata dunque prima di compattare questi; per quanto riguarda il **numero di utenti** abbiamo calcolato tale valore sul grafo completo, il quale racchiude tutte le tipologie di archi.

In questo modo siamo sicuri di aver incluso in tale analisi **tutti i possibili nodi**, anche quelli non attivi e che magari vengono solo attaccati. In tal senso abbiamo anche cercato di definire una numerosità per quelli che sono i **villaggi abbandonati**, ovvero quei villaggi utilizzati dai players come un villaggio da razziare o facilmente conquistabile.

Abbiamo quindi valutato quei villaggi che, dato il giorno, venivano interessati **solo dall'attacco** in ingresso; ovviamente tale misura include anche quei villaggi che appartengono a players che hanno **smesso di giocare**. Tuttavia, stando alle regole Travian, questi diventeranno dopo un periodo di 7 giorni effettivamente villaggi barbari, l'utente non potrà quindi riappropriarsi dello stesso. Forniamo ora i risultati dell'analisi per poi passare alle considerazioni:

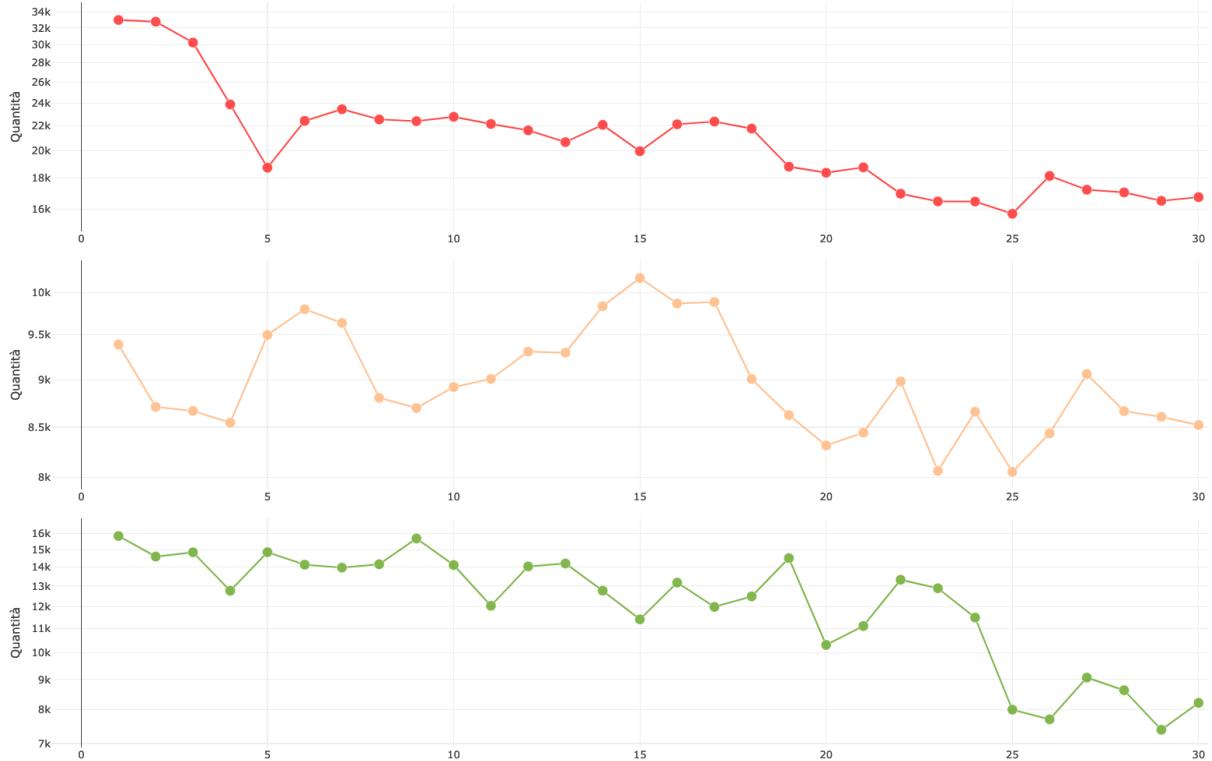


Figura 6: Numero di attacchi, commerci e messaggi nel periodo

Analizzando prima il numero di archi ci possiamo rendere conto di come sia l'**attacco** l'**attività principale** e di come, anche nei picchi, i messaggi non riescano a superare il numero di azioni offensive. Questo è sicuramente giustificato dalla **tipologia di gioco** ma

soprattutto dal server. Difatti lo **speed server** fa sì che i movimenti sulla griglia di gioco siano più veloci, di conseguenza gli attacchi possono essere molto più dinamici e le truppe più lente (come le catapulte) possono muoversi con un tempo molto ridotto. Possiamo comunque notare una tendenza di “**decadenza**” in quanto notiamo una **diminuzione generale di tutte le attività**.

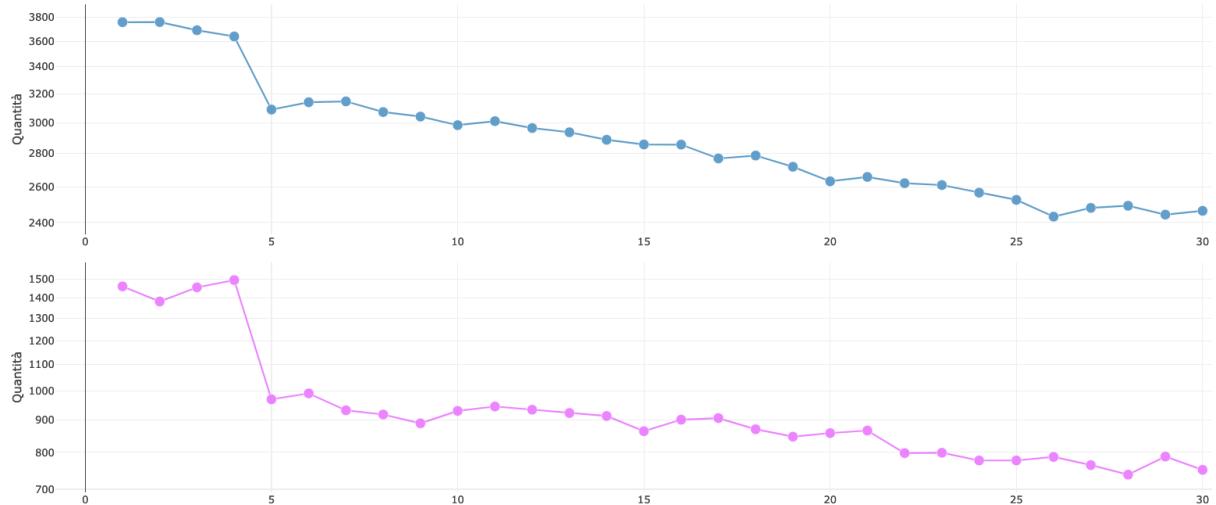


Figura 7: Numero di utenti e “*barbari*”

Le supposizioni sopra, dato il numero di archi, vengono ulteriormente confermate dal **numero di utenti** e il **numero di utenti inattivi** in piena e continua decrescita lungo tutto il periodo. Vi è una **perdita** importante di utenti che si attesta intorno al **34%**. Un appunto particolare va fatto sul **quinto giorno**, il quale ha una decrescita importante e significativa, tutta in una sola volta. Tale situazione è presente anche nell’analisi successiva, non siamo riusciti a trovare una spiegazione a tale fenomeno.

Forniamo in ultimo una visione complessiva di quanto mostrato precedentemente, al fine di mostrare una **visione generale della numerosità**:

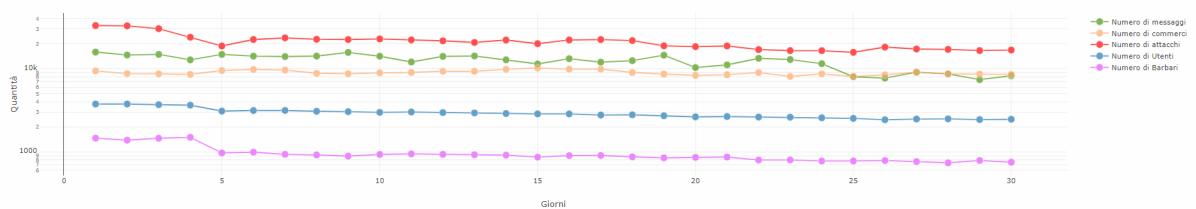


Figura 8: Numerosità a confronto

`userNumberTrend` e `edgeNumberTrend`.

5.2.3 Communities Trend

Dati i trend precedentemente mostrati abbiamo voluto indagare riguardo le **communities** osservando se la numerosità di queste seguisse un andamento simile. Per effettuare tale analisi abbiamo valutato i file delle community ottenuto in seguito alla fase di pre-processing precedentemente mostrata; dunque siamo in possesso di informazioni riguardanti il **numero effettivo di communities attive**, il numero di quelle **dissolte** e quello delle **create**. Per questa analisi è stata creata una struttura dati che prende il nome di **communityAffiliation**; sostanzialmente un dizionario la cui struttura è la seguente:

$$\{ \text{Day: } \{ (\text{CommunityId: } ((\text{User1}, \text{User2}, \dots, \text{UserN}), \text{nUser}, \text{nActive}) \} \}$$

Teniamo quindi traccia, dato il giorno e la community, dei **membri** appartenenti a questa, del numero di utenti nella stessa e di quelli attivi (un utente per essere attivo deve avere un outdegree nella rete totale almeno maggiore uguale a 1).

I **risultati** da noi ottenuti sono i seguenti:

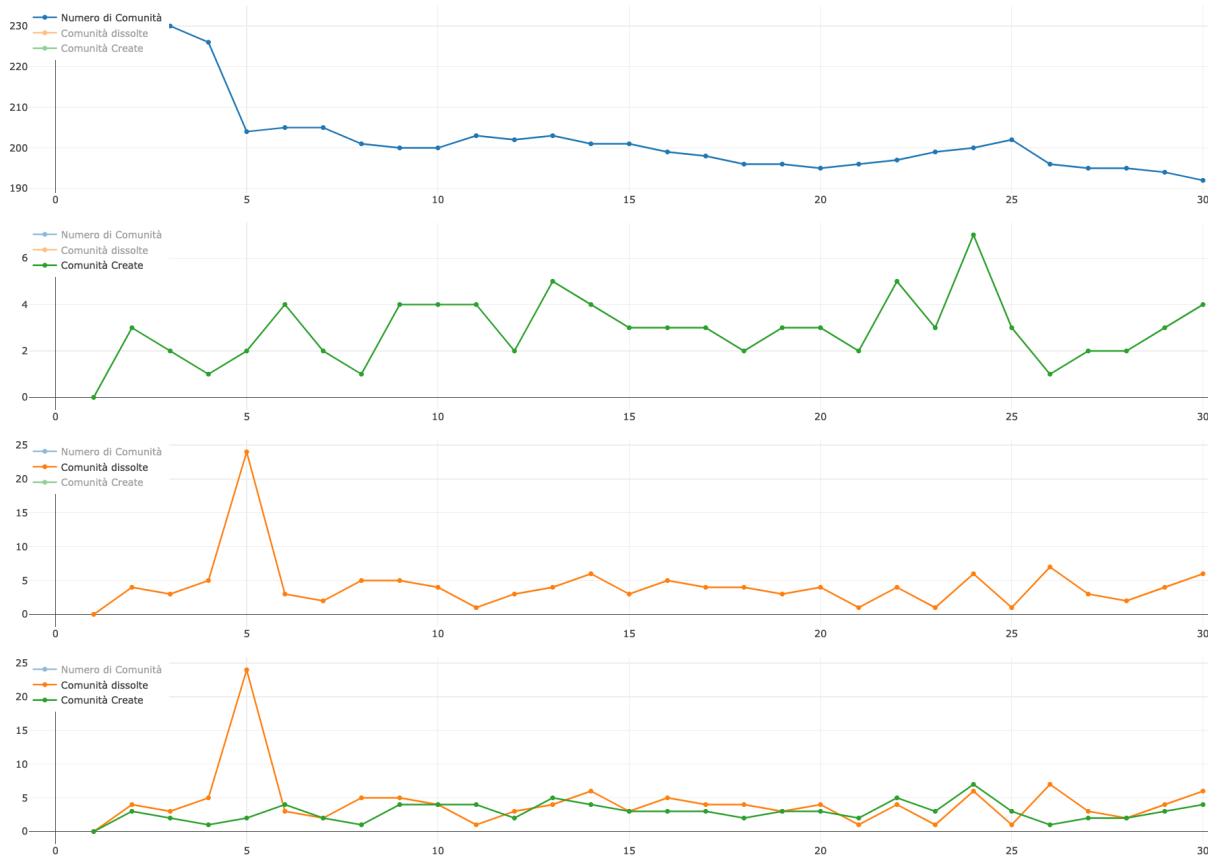


Figura 9: Trend numero communities totali, create, dissolte e confronto tra create e dissolte

Procediamo con ordine:

- Il **numero di communities diminuisce** durante il periodo, passando da un iniziale 232 a 192; vi è quindi una descrescita pari al 17,3%. Un ulteriore conferma che il server procede verso la fase finale di attività. Mediamente un'Alleanza è formata da 10 membri.
- Il **numero di communities create** segue un andamento **oscillatorio**, con picchi più o meno alti che sembrano relativamente indipendenti dal periodo (osserviamo una tendenza di crescita verso la fine di questo). Dal singolo grafico non otteniamo informazioni significative.
- Il **numero di communities dissolte** risulta più interessante rispetto a quello delle communities create, notiamo il picco al giorno cinque già presente nelle precedenti analisi per il quale non siamo riusciti a trovare una spiegazione. Possiamo osservare una **tendenza sicuramente più costante**, tolte poche eccezioni, rispetto a quanto visto nel grafico precedente.
- Il ultimo il **confronto tra numero create e dissolte** è relativamente **a favore delle dissolte**, elemento necessario alla decrescita mostrata nel primo grafico. Mediamente al giorno si creano 3 communities e se ne dissolvono 4.

A questo punto ci siamo interessati soprattutto alle **communities dissolte** (il cui numero totale è 127), nel tentativo di comprenderne la numerosità in termini di membri e in funzione di questa cercare di capire, dato l'id utente da noi considerato una sorta di “data di nascita”, la tipologia di utenti che ne faceva parte. Riteniamo utile fare luce su questa dinamica al fine di poter, eventualmente, migliorare determinate dinamiche di gioco come quelle che riguardano la creazione delle Alleanze.

Valutando l'intero periodo osserviamo i seguenti dati:



Figura 10: Numerosità communities sciolte, età players nelle “single-communities”

Possiamo fin da subito notare come il **74% delle communities sciolte** sia composta da **un solo utente**, mentre solo una piccola parte risulti essere rappresentativa di communities più numerose (che rappresentano con più probabilità il fenomeno della fusione di communities). Abbiamo quindi che probabilmente tali **utenti sono stati sconfitti** da altri o hanno **abbandonato** il gioco, in quanto abbiamo anche osservato che dati gli utenti delle “single-communities” solo il 5% di questi risulta essere presente in una communities nei giorni successivi.

Inoltre, date le “single-communities”, **più della metà è composta da utenti giovani** (id maggiore di 8000); i quali ipotizziamo essere più **inesperti** delle meccaniche di gioco che, data l'analisi, non sono state chiarite in fase di creazione dell'Alleanza.

Ipotizziamo dunque che le meccaniche di Alleanza e logiche di queste **non siano spiegate in maniera sufficientemente chiara** mediante all'interno del gioco; in caso contrario un utente sarebbe più incline a entrare in **communities già formate** o comunque a tentare di crearne una ma, una volta fallito, tentare di entrare in una già presente. Tale situazione porta a una possibile **perdita** in termini di denaro, questo perché un utente che partecipa ad un'Alleanza può interagire con altri utenti diventando dunque un utente attivo. La continua attività e interazione può portare l'utente ad affezionarsi al gioco e a spendere denaro per poter continuare a giocare con dei bonus e vari abbonamenti.

In ultimo abbiamo voluto indagare sul **giorno cinque** in maniera analoga a quanto fatto su tutto il periodo:



Figura 11: Numerosità communities sciolte, età players nelle “single-communities”

Per quanto riguarda la numerosità osserviamo un'accentuazione sul fenomeno delle *“single-communities”* già osservato precedentemente; per l'età invece è tutto più equamente distribuito sempre però a favore degli utenti young (40.9%).

Sarebbe quindi interessante provare a creare una sorta di **“tutorial”** da presentare all'utente nel caso in cui questo sia interessato a creare un'alleanza, **rieffettuando le analisi** siamo convinti che **i risultati sarebbero diversi**.

Le funzioni che effettuano queste analisi sono: `communitiesNumberTrend()`, `dissolvedCommunitiesTrend()`, `creationCommunitiesTrend()`, `nUserDissolved()`, `blackDay()`, `informationAboutCommunity()`. Tutte quante si basano sul **dizionario** precedentemente mostrato in questa sezione.

5.2.4 Grado Medio e Distribuzione di Grado

Passando poi ad **analisi più classiche** e legate alla **struttura della rete** abbiamo valutato il grado medio e cercato di incrociare le informazioni ottenute da questo con la **distribuzione di grado** al fine di comprendere l'andamento all'interno della rete.

Riteniamo utile quest'analisi in quanto potrebbe ulteriormente confermare quanto già intuito per ciò che riguarda le attività principali all'interno della rete.

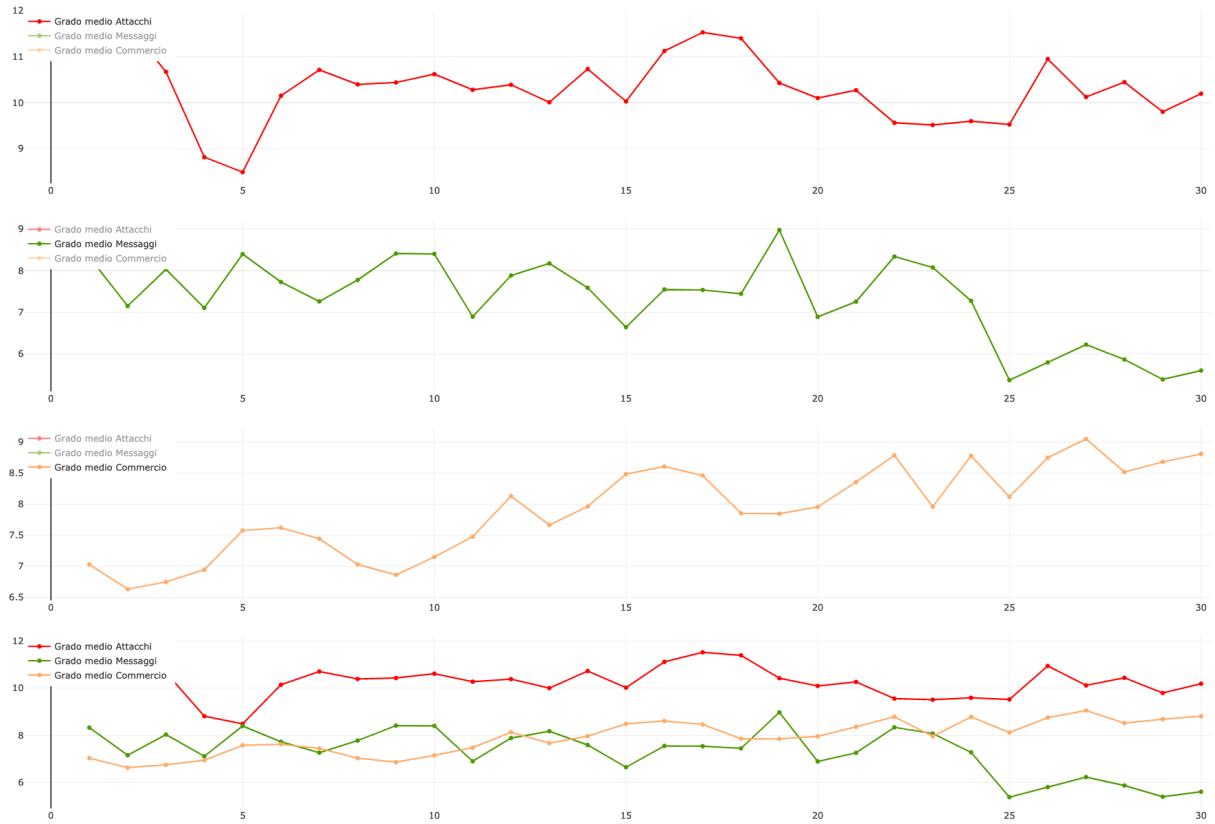


Figura 12: Grado medio attacchi, messaggi, trade e confronto tra questi

Come ci aspettavamo l'attacco è l'attività predominante all'interno della rete, tolto il picco negativo al giorno quattro e cinque, abbiamo una **media, a utente, di 10 archi** data questa tipologia di interazione; situazione motivata ancora una volta dal velocità di movimento superiore del server. L'**andamento positivo e costante**, legato alla continua **diminuzione di players**, ci fa pensare che quelli che rimangono **continuino con un atteggiamento offensivo**; pensiamo dunque che i **players eliminati durante il periodo analizzato non fossero così attivi**.

Per quanto riguarda invece i **messaggi** notiamo una tendenza relativamente **costante** nei primi 23 giorni, in seguito notiamo una profonda **decrescita**. Questa situazione può essere causata dalla **diminuzione continua e costante delle comunità** e del numero di utenti.

In ultimo i **commerci**, con grande sorpresa, sembrano **aumentare considerevolmente nonostante la decrescita degli elementi nella rete**. Questo ci fa pensare che gli utenti rimasti cerchino in tutti i modi di aumentare le proprie possibilità di vittoria con un incremento del commercio al fine di avere più risorse da utilizzare per rafforzare i propri villaggi e i propri eserciti.

In funzione dell'analisi sul grado medio ci siamo poi interessati a valutare la **distribuzione di grado per tutte le tipologie di interazione**; abbiamo deciso di dividere

in degli intervalli per rendere più semplice la visualizzazione dei risultati.

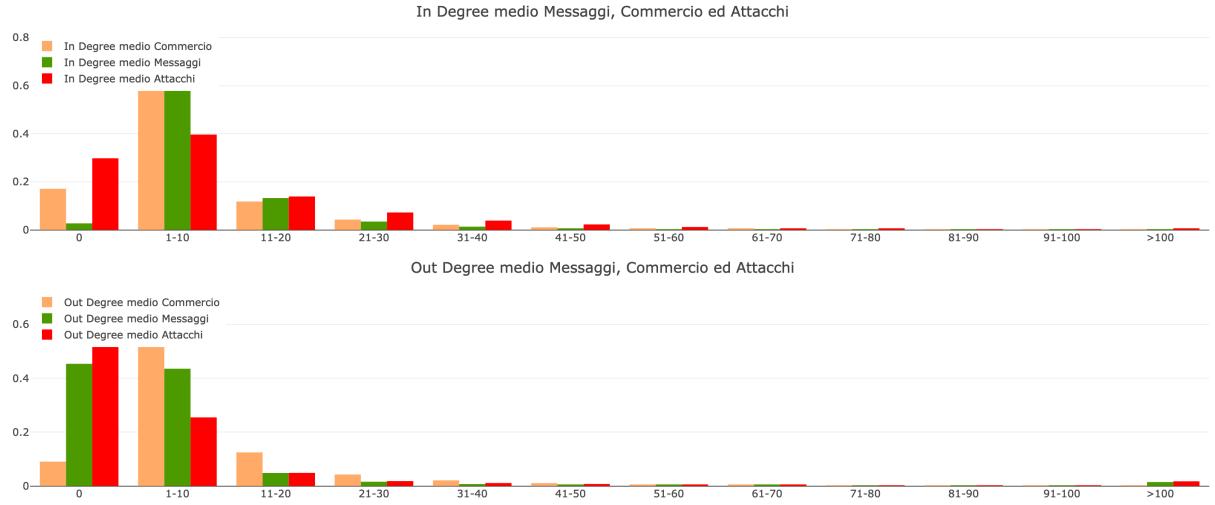


Figura 13: Distribuzione di grado delle interazioni tra gli utenti

Possiamo osservare come, per quanto riguarda i messaggi, sia molto più probabile non inviare messaggi rispetto a non riceverne alcuno; ipotizziamo dunque la presenza di nodi i quali tendono ad inviare un grande quantitativo di messaggi. Tali nodi potrebbero rappresentare, all'interno delle communities, i nodi leader; sono dunque necessarie ulteriori analisi al fine di confermare tale eventualità.

Per quanto riguarda il commercio, osserviamo una tendenza più equilibrata in funzione dell'aumento di quantità di scambi; ciò denota che più la quantità di commercio è elevata più questa, pare, essere controbilanciata. Analisi riguardanti la reciprocità di tale collegamento aiuteranno a definire in maniera più chiara la situazione.

Per quanto riguarda l'attacco osserviamo un tendenza belligerante a un numero limitato di utenti, difatti il la percentuale media riguardante un outdegree pari a zero è molto superiore rispetto alla probabilità media di un indegree pari a 0. Notiamo tuttavia che, dati i vettori di attacco tra utenti veterani, questi siano relativamente limitati; difatti questi solitamente si aggirano sui 10/15 attacchi come minimo, data la realtà di interesse notiamo come queste intervalli presentino una probabilità relativamente. Dunque ipotizziamo che l'inizio della fase finale non sia stato registrato dato il periodo di riferimento.

5.2.5 Average Shortest Path e Diametro

Al fine di comprendere l'interconnessione reale date le varie tipologie di relazione abbiamo effettuato degli studi sullo **shortest path medio** e sul **diametro** per ognuna di queste. Date le analisi già effettuate **ci aspettiamo che sia tutto a favore dell'attacco**.

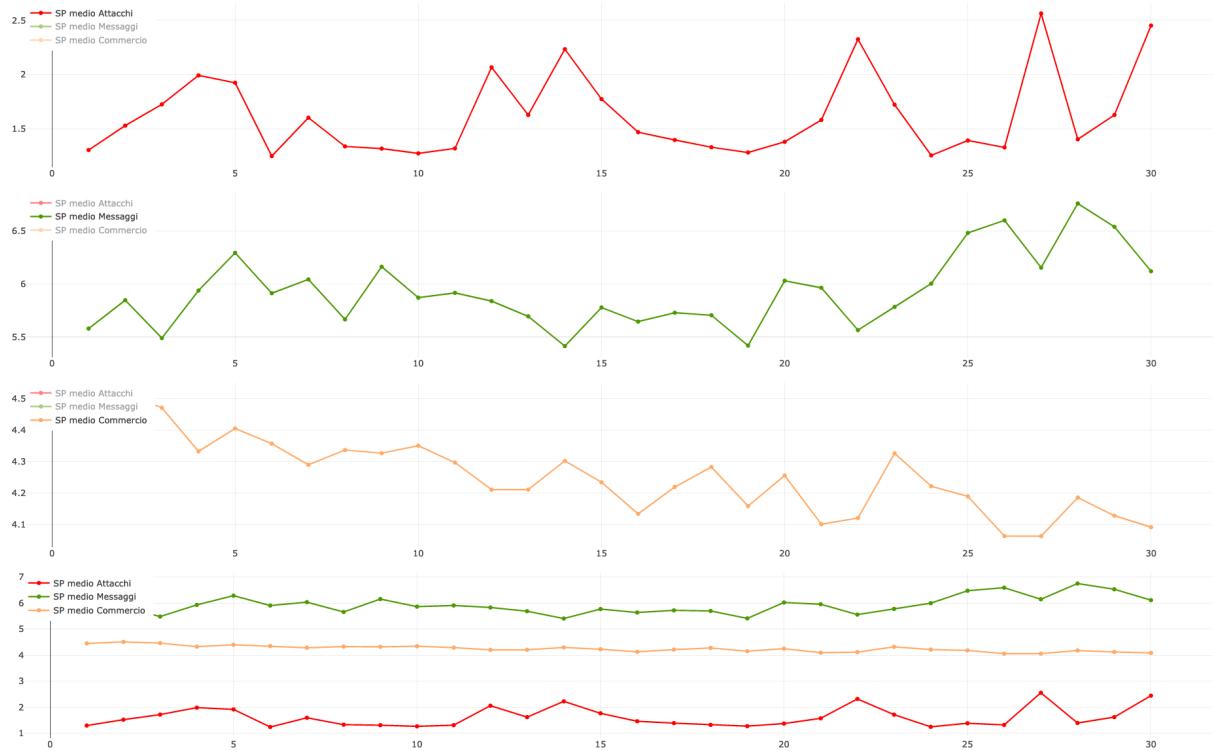


Figura 14: Shortest path attacchi, messaggi, trade e confronto tra questi

Lo **shortest path dell'attacco** presenta dei picchi che tuttavia non sono di grande interesse, difatti abbiamo un valore medio che si **aggira intorno al 2**; gli utenti quindi sono molto **bellicosi** e “vicini” dal punto di vista offensivo. L’andamento risulta essere costante se messo a confronto con i valori delle altre interazioni.

Per i **messaggi** osserviamo un trend di significativa crescita verso gli ultimi giorni del periodo sotto esame; questo però in corrispondenza del 30esimo giorno presenta un segno di decrescita la quale poi non possiamo sapere se continua a causa della manca di informazioni. Tuttavia possiamo notare che il tutto si allinea alla decrescita di grado medio registrata precedentemente: **gli utenti comunicano di meno e di conseguenza le loro distanze aumentano**.

Discorso analogo ma ribaltato per ciò che riguarda il **commercio**, di fatti la **crescita di grado** registrata **si ripercuote con una decrescita della distanza** tra utenti dato questo collegamento.

L'ultimo grafico mostra un semplice confronto tra le diverse tipologie.

Abbiamo quindi a questo punto valutato il **diametro** al fine di verificare la distanza massima che viene registrata date le diverse tipologie di relazione, in funzione poi dell'average shortest path potremmo inferire sulla effettiva distanza massima al fine di capire se questa rappresenti un caso isolato o meno.

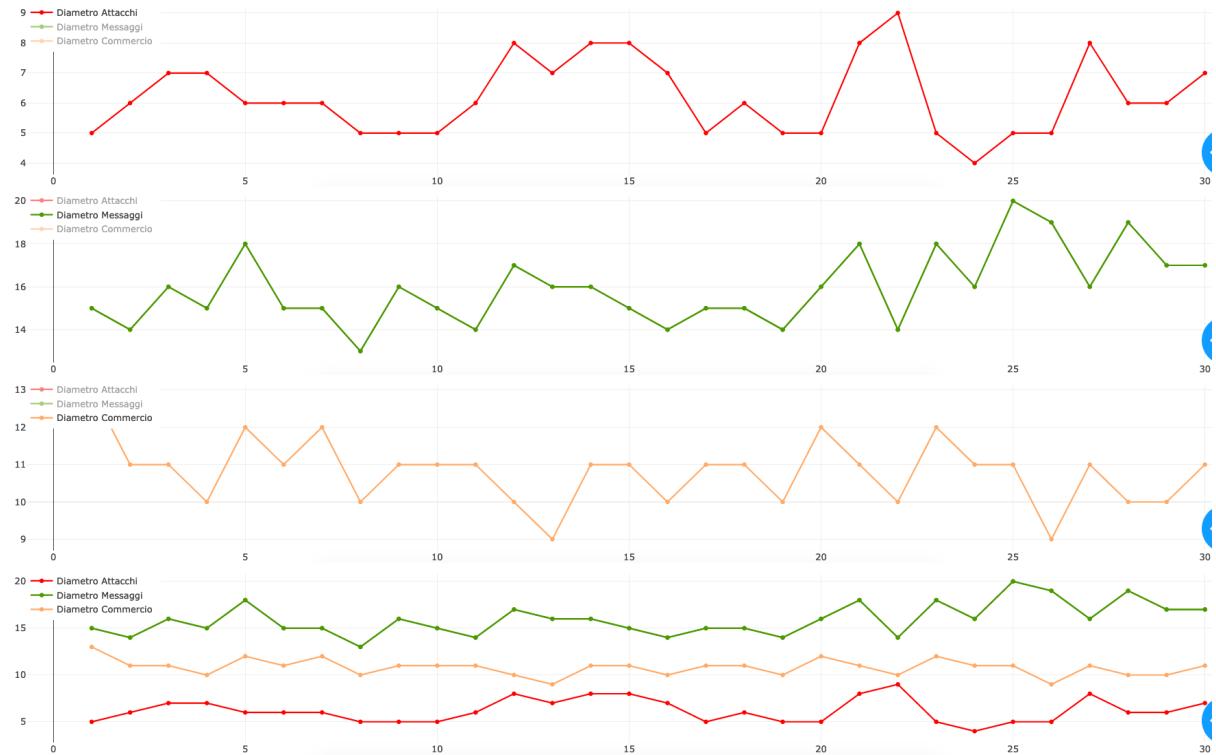


Figura 15: Diametro attacchi, messaggi, trade e confronto tra questi.

Per ciò che concerne gli **attacchi** possiamo notare come vi siano **situazioni di catena**, immaginiamo quindi (dati i picchi positivi) che utenti attaccati, attacchino altri per rimbattere delle risorse perse precedentemente; questa tendenza sembra ripercuotersi creando dunque una vera e propria **catena di attacchi**. Questo fenomeno ci sembra relativamente raro; difatti i giorni in cui si registrano i picchi più alti (come il 22esimo giorno) a livello di shortest path medio ci si aggira intorno a 2.4/2.5. Questi valori sono poi quelli più elevati data la misurazione.

Diverso discorso per quello che riguarda i **messaggi**, qui il diametro è molto più elevato toccando punte pari a 20 verso gli ultimi giorni; questi poi presentano in generale valori più alti i quali erano previsti data l'aumento dello shortest path e la relativa diminuzione di grado proprio negli ultimi giorni di osservazione. I valori del diametro sono quindi strettamente legati all'andamento di questa tipologia di interazione e denotano un allontanamento dei players dato questa tipologia di relazione.

In ultimo il **commercio** il quale negli ultimi giorni si avvicina molto ai valori registrati anche per l'attacco; ulteriore prova che questa interazione acquisti molto importanza con il trascorrere del tempo. Vige comunque una **catena di commerci**.

5.3 Trickster Detection Tool

5.3.1 Introduzione

Il **multi account** è una pratica molto diffusa in questa tipologia di Browser game, l'idea è quella di creare uno o più **account fintizi** (solitamente comunque solo uno per essere meno "sospetti") al fine di poter **privilegiare mediante un continuo commercio i villaggi principali**. Ovviamente il player che si crea un secondo account utilizzerà anche un nome fintizio.

Travian è **molto severo** per quanto concerne questi comportamenti, dato il regolamento vi sono un paio di regole riguardanti le sanzioni di tale azioni:

- **§ 1.1** Ogni giocatore ha diritto al possesso ed all'uso di un unico account per server. Un account deve essere sempre giocato a proprio beneficio. Gli account che esistono solo per il beneficio di altri account saranno bannati permanentemente.
- **§ 7.3** I multiaccount nei server speed e i multiaccount con una popolazione inferiore a 100 possono essere cancellati senza ban'avvertimento.

Figura 16: Regolamento Travian^[3]

Questa **problematica è molto sentita** anche a livello di forum; sono molti i topic aperti, più o meno recenti, riguardanti la stessa.

Ovviamente Travian cerca di implementare controlli per contrastare questo fenomeno, i dettagli di questi non sono di dominio pubblico per ovvie ragioni. Abbiamo quindi provato ad **implementare un meccanismo che possa aiutare all'identificazione dei players più sospetti** fornendo un aiuto a dei moderatori umani del server.

5.3.2 Logica di funzionamento

Il tool è pensato per lavorare in maniera continuativa nel tempo, al fine di poter monitorare il comportamento degli utenti; lo scopo è pertanto quello di **identificare i players sospetti** che iniziano ad operare durante il **periodo di osservazione**, abbiamo comunque tentato di poter ampliare il più possibile al fine di poter identificare anche gli account creati precedentemente al periodo analizzato. Le supposizioni alla base di tutto sono le seguenti:

1. **Un cheater** avrà sicuramente un giorno in cui si farà solo commercio verso un determinato utente. Questo perché per ogni nuovo utente sul server sono disponibili dei giorni di protezione dagli attacchi di altri players.

2. Se vi è un giorno in cui **si ha fatto solo trade** si è sospetti e si inizia a valutare tutte le tipologie di relazioni sul periodo.
3. **Subire tanti attacchi** può scagionare: subisco un attacco, anziché perdere risorse le invio nella speranza poi di poterle riavere.
4. **Inviare e ricevere messaggi** diminuisce il sospetto.
5. **Ricevere risorse da colui a cui si ha inviato** riduce il sospetto.
6. **Effettuare tanti attacchi aumenta** il livello di sospetto: non mi basta produrre e basta attacco per fare risorse ed invio all'account principale.
7. **Inviare a più utenti** riduce il livello di sospetto.

Abbiamo quindi osservando i players che nel periodo avevano **un giorno in cui commerciavano e basta**. Questa operazione dunque è stata effettuata per tutti i giorni e tutti i players. Abbiamo dunque creato una **struttura** così formata:

$$\{ \text{Day}: \{ (\text{idSuspected}: \{ \text{idBeneficiary}: \text{qtn} \}) \} \}$$

A questo punto abbiamo “compattato” tale struttura la quale non è più divisa per giorni ma, per ogni sospettato, ci mostrerà con chi ha commerciato e in che quantità durante il periodo. Sostanzialmente un dizionario come quello precedente ma con un livello di embedded in meno e con il numero di volte in cui ho visto l'utente nel periodo:

$$\{ \text{idSuspected}: (\text{nTimes}, \{ \text{idBeneficiary}: (\text{nTimes}, \text{qtn}) \}) \}$$

A questo punto abbiamo **eliminato** i players che risultavano osservati solo una volta sull'intero periodo e valutato gli altri. Abbiamo quindi definito una funzione per attribuire ad ognuno dei sospettati un **indice di colpevolezza**; la funzione è la seguente:

$$EOTC = \text{Numero archi in uscita di trade}$$

$$EITC = \text{Numero archi in ingresso di trade}$$

$$EOMC = \text{Numero archi in uscita di messaggio}$$

$$EIMC = \text{Numero archi in ingresso di messaggio}$$

$$EOAC = \text{Numero archi in uscita di attacco}$$

$$EIAC = \text{Numero archi in ingresso di attacco}$$

$$PS = \frac{1}{PS}$$

$$\text{GuiltyIndex} = \frac{(EOTC * PS * 1.3 + EIMC * 0.3 + EOAC * 0.4) - (EITC * 5 + EOMC + EIAC * 0.08)}{EIAC * 0.08 + EITC + EIMC + EOAC + EOMC * 2 + EOTC}$$

Figura 17: Funzione di Guilty e legenda

Ovviamente la misura è utile solo in questo ambito ed è **frutto di osservazioni sperimentali**. Dati dunque i vari indici, si filtrano quelli maggiori di una certa soglia (0.3) e si passa ad esaminare solo questi nodi valutando, ai fini di aiutare un umano, anche i giorni in cui effettua i trade e in cui riceve gli attacchi; in questa maniera, osservando se

vi è corrispondenza tra queste relazioni, si può inferire in maniera più corretta osservando anche gli altri valori.

Nella **demo** che verrà presentata, si mostrerà l'**output** del tool.

5.4 Commercio e Messaggi

5.4.1 Introduzione

Data la numerosità limitate di queste due tipologie di relazioni, ci siamo interessati a studiarle al fine di comprenderne la natura e l'evoluzione nel periodo. Per compiere questo studio ci siamo quindi concentrati solo sui nodi i quali risultavano allineati con una qualche alleanza, abbiamo deciso di includere anche quelli in communities più piccole (non abbiamo imposto quindi vincoli) in quanto riteniamo corretto valutare tutti i players che sono più probabilmente i più attivi.

Abbiamo dunque provveduto da prima ad **etichettare i nodi date le loro alleanze**, poi abbiamo eliminato i non allineati e ci siamo quindi concentrati solo sugli archi di **commercio e messaggio**. Le analisi sono volte a uno studio e a un confronto, tra le relazioni di messaggio e commercio, **sia a livello interno che esterno alle community**; questo ci permetterà di capire come ci si relaziona in questi due contesti.

5.4.2 Reciprocità

Come prima analisi abbiamo osservato il **valore della reciprocità** al fine di comprendere la natura degli scambi commerciali e di messaggi, dunque sui primi ci aiuterà a capire la tendenza ad uno scambio equo (io do a te, tu dai a me), sui secondi ci fornirà informazioni sulle natura delle conversazioni (**monodirezionalità** vs. **bidirezionalità**).

Partendo dalla reciprocità del commercio otteniamo i seguenti valori:

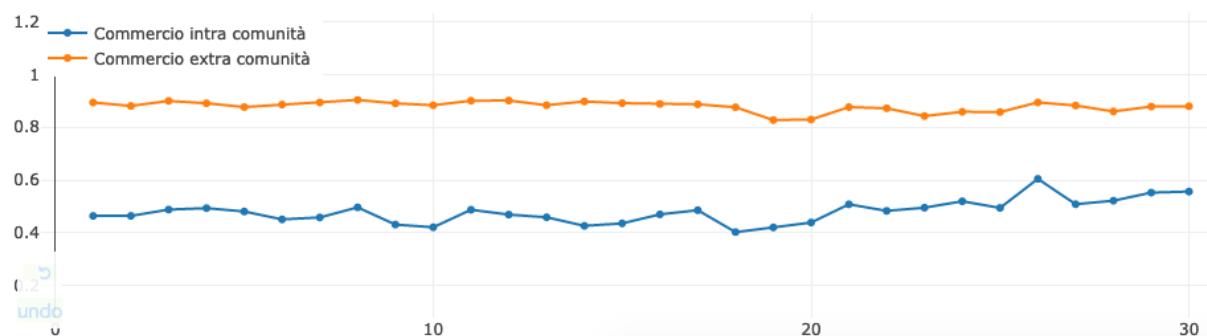


Figura 18: Reciprocità commercio intra ed extra community

Possiamo notare come, a livello **extra-community**, la reciprocità presenti valori molto alti con media intorno al **0.9**; questo significa che vi è la **presenza di un mercato**

globale che vede uno scambio equo tra i partecipanti. Possiamo affermare dunque che tra utenti di community diverse vi è la tendenza di commerciare mediante le offerte che si possono creare tramite l'edificio mercato, permettendo così uno scambio.

La situazione nelle community pare essere rovesciata, osserviamo valori di reciprocità mediamente sotto allo **0.5**; questo significa che **nelle community vi è una tendenza assistenzialista** che quindi vede determinati players aiutare altri e non ricevere nulla in cambio. Ipotizziamo possa essere una tattica che permette ai **players più forti di aiutare i players più deboli** inviando loro un certo quantitativo di risorse.

Per quanto riguarda i **messaggi** abbiamo questa situazione:

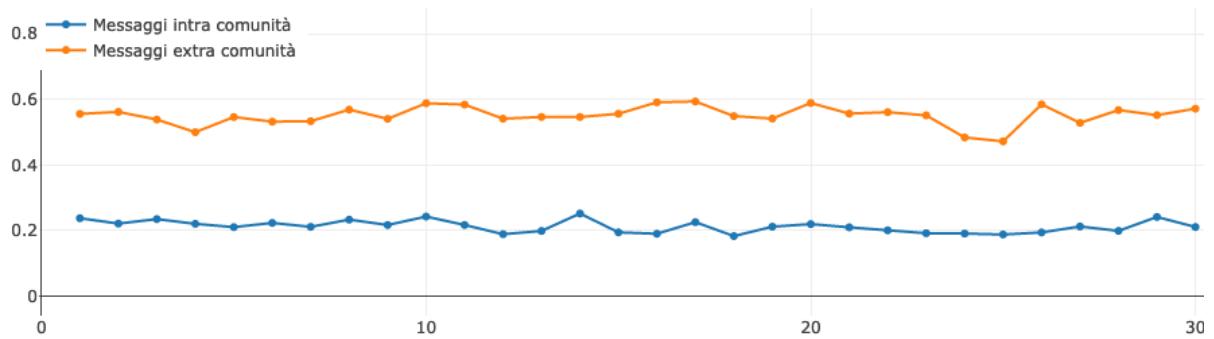


Figura 19: Reciprocità messaggi intra ed extra community

Tra players di **community diverse** osserviamo una **reciprocità** che si aggira intorno allo **0.55/0.58**; dunque si tende a rispondere ai messaggi di utenti esterni alla propria community. Difatti **il messaggio, in questo caso, rappresenta l'unica possibilità** di comunicazione tra gli utenti.

Internamente a una community possiamo invece notare un indice di reciprocità molto basso; mediamente pari a **0.20**. Questo ci fa pensare che il messaggio sia utilizzato come un **avviso monodirezionale** da parte dei membri più centrali della community (i capi unità ad esempio) verso tutti gli altri. Questa apparente poca comunicazione è **motivata dalla presenza di un forum integrato per ogni alleanza**, in questo è possibile ed avviene tutta la comunicazione intra-community.

5.4.3 Densità

Al fine di cercare una misura dei fenomeni registrati nella precedente analisi abbiamo deciso di effettuare, in maniera analoga a quanto fatto sopra, un **studio della densità**.

I valori per quanto riguarda il commercio sono i seguenti:

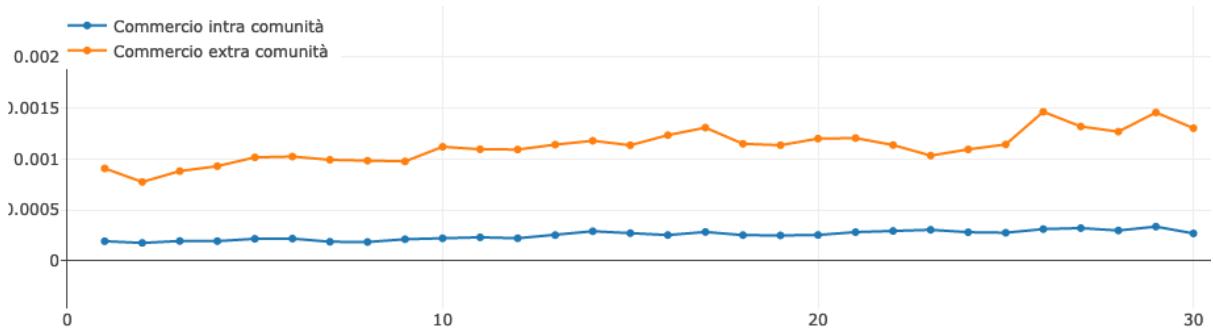


Figura 20: Densità commercio intra ed extra community

Valutando prettamente il confronto tra le misure, possiamo affermare che il **commercio** avviene principalmente tra **players non appartenenti alla stessa alleanza**; dunque il commercio è principalmente equo e il fenomeno di assistenzialismo registrato all'interno delle community è molto limitato e poco presente.

Stessa analisi per quanto riguarda i messaggi:

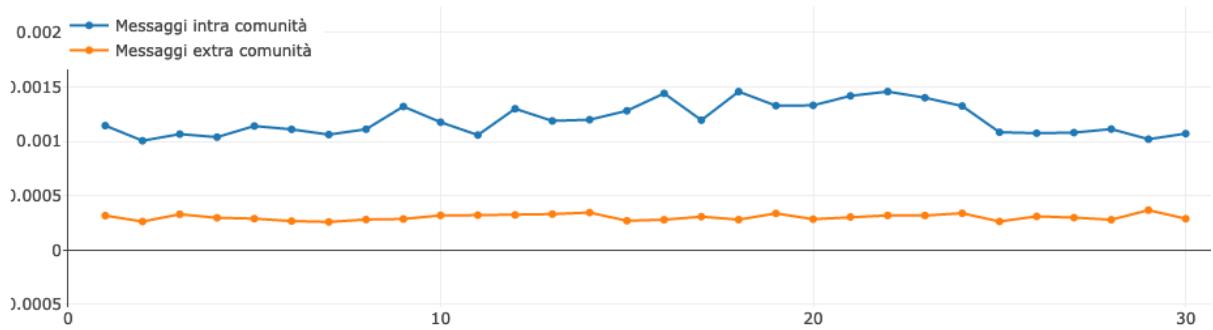


Figura 21: Densità commercio intra ed extra community

Come ci aspettavamo, la **densità è minore dati utenti di community diverse**, dunque la comunicazione tra le diverse alleanze è effettuata da pochi individui; tale analisi ci ha suggerito la possibilità di poter **identificare i diplomatici**. Internamente alla community, come ci aspettavamo, abbiamo valori più alti. La **tendenza monodirezionale** del messaggio è dunque quella predominante.

5.4.4 Clustering Coefficient

Analisi tutta interna alle community. Abbiamo deciso di analizzare la probabilità con cui i vicini di un nodo si relazionano tra di loro al fine di comprendere la **presenza di strutture a stella** le quali possono essere sfruttate per inferire, mediante le diverse misure di centralità studiate, sui **possibili leader**.

I dati da noi ottenuti sono i seguenti:

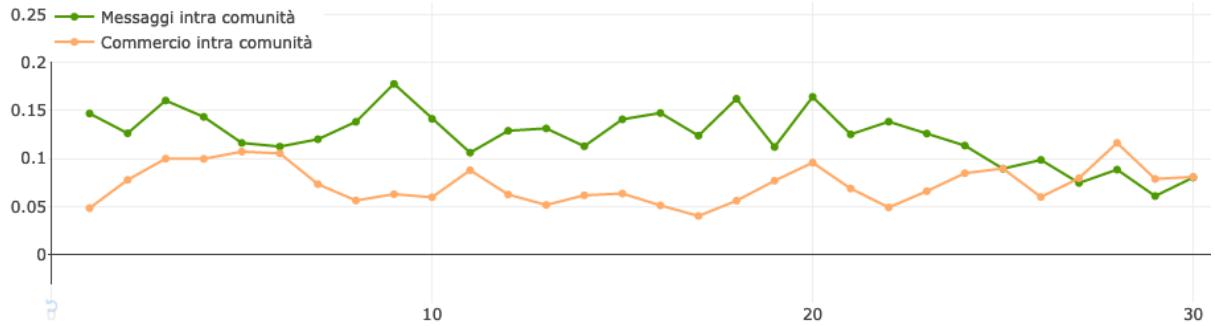


Figura 22: Clustering Coefficient dei messaggi e dei commerci intra community

Possiamo osservare che la **tendenza assistenzialista** è nuovamente certificata dal fatto che il **clustering coefficient** su questa tipologia di relazione oscilla intorno al **10%**. I messaggi invece presentano si valori più alti, ma comunque sia molto dimensionati con media del 13%/14%; ipotizziamo dunque che sia relativamente semplice identificare i leader valutando, tra le altre misure, la centralità di grado rispetto ai messaggi.

5.4.5 Diametro e Shortest Path

Dalle analisi emerge come il **diametro intra-community** per i messaggi, come ci aspettavamo, **sia minore rispetto a quello extra-community**. Questa situazione è probabilmente dovuta alla maggiore densità registrata all'interno delle community a causa della loro stessa natura. Il diametro del commercio segue invece un andamento alternato ma mantiene comunque lo stesso trend a favore dei nodi intra-community, anche se non in modo così evidente come visto per i messaggi.

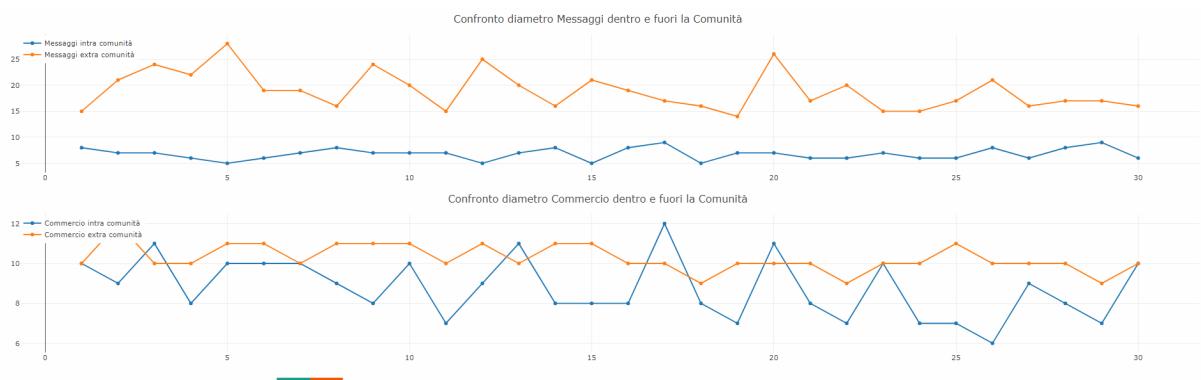


Figura 23: Diametro dei messaggi e dei commerci intra ed extra community

Lo **Shortest Path** riconferma ulteriormente l'ipotesi mostrando un cammino minimo di lunghezza inferiore tra membri della stessa community rispetto a membri di comunità differenti Infatti alla base del concetto di community possiamo trovare come i nodi che fanno parte siano molto legati tra loro rispetto all'esterno.

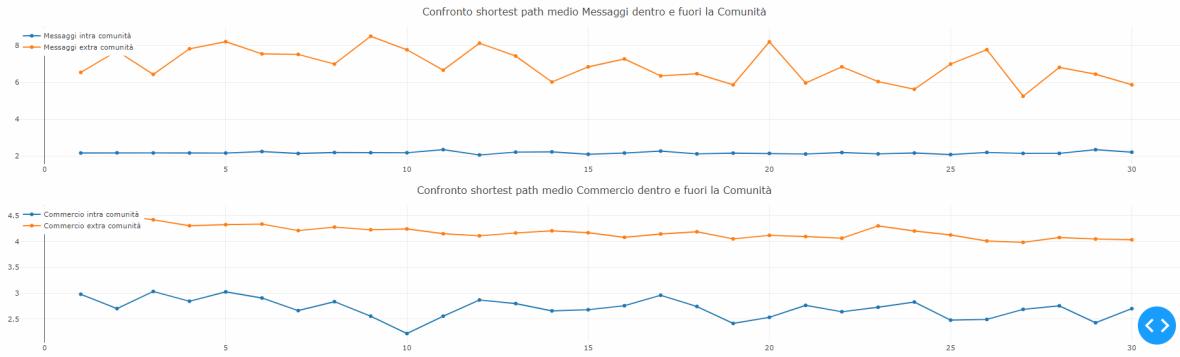


Figura 24: Shortest Path dei messaggi e dei commerci intra ed extra community

5.4.6 Average Degree

La minor densità dei messaggi extra comunità si ripercuote sul anche **grado medio dei messaggi**, che mostra una **maggior attività internamente alle comunità** nelle quali ci aspettiamo un livello maggiore di comunicazione tra i membri. I **diplomatici** possono esistere ma hanno rapporti limitati rispetto ai membri totali di una community. Analogi discorsi, seppur ribaltato, per il **commercio**. La **tendenza assistenzialista** è quindi limitata seppur presente, la presenza di un mercato tramite il quale è possibile scambiarsi risorse rende il commercio con utenti sconosciuti molto più probabile.

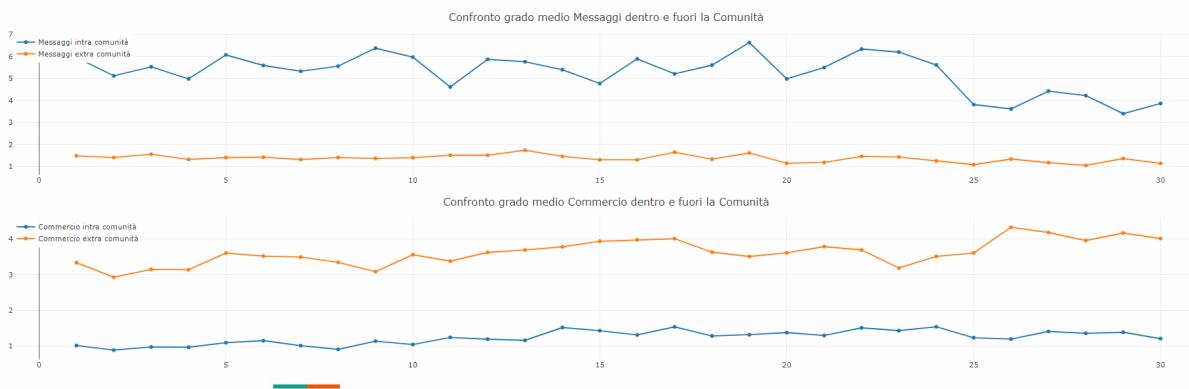


Figura 25: Average Degree dei messaggi e dei commerci intra ed extra community

5.4.7 In and Out Degree Distribution

Per il **commercio extra comunità** vi è un **andamento bilanciato** tra ingresso ed uscita dove più il numero di archi aumenta più la probabilità è simile. Non considerando coloro che non commerciano, la tendenza è rilegata ad un numero limitato. All'interno delle **community** osserviamo invece che coloro che commerciano **sono una minoranza**

e le quantità sono più rilegate all'intervallo 1-10; vi è però un andamento meno simile tra ingresso ed uscita.

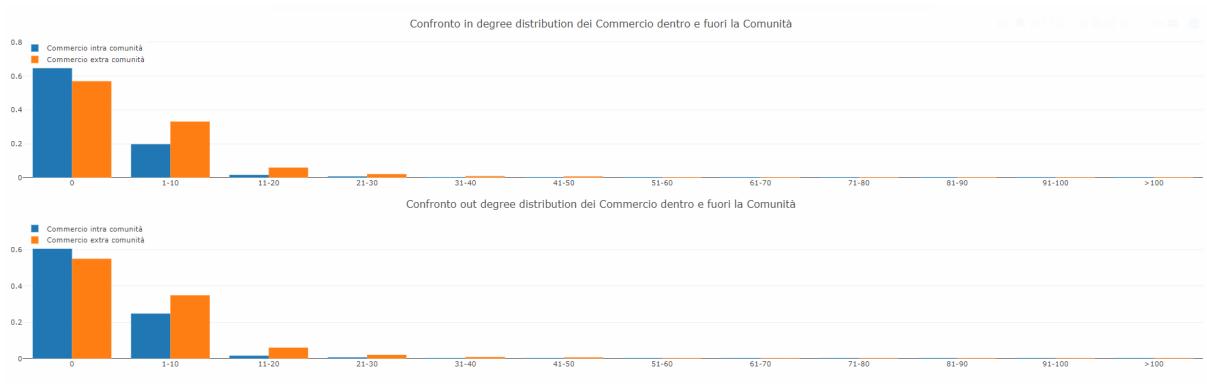


Figura 26: In ed Out Degree distribution dei commerci intra ed extra community

A conferma della **natura di avviso dei messaggi**, osserviamo come la probabilità di non inviare nessun messaggio sia molto superiore a quella di non riceverne alcuno. Questo andamento si ripercuote in tutti gli intervalli definiti. **Tra comunità diverse** è anche in questo caso chiara la **tendenza alla reciprocità** la dove i messaggi sono effettivamente scambiati. Osserviamo che **gli utenti di comunità diverse che parlano sono una minoranza**, essi potrebbero essere i **diplomatici**.

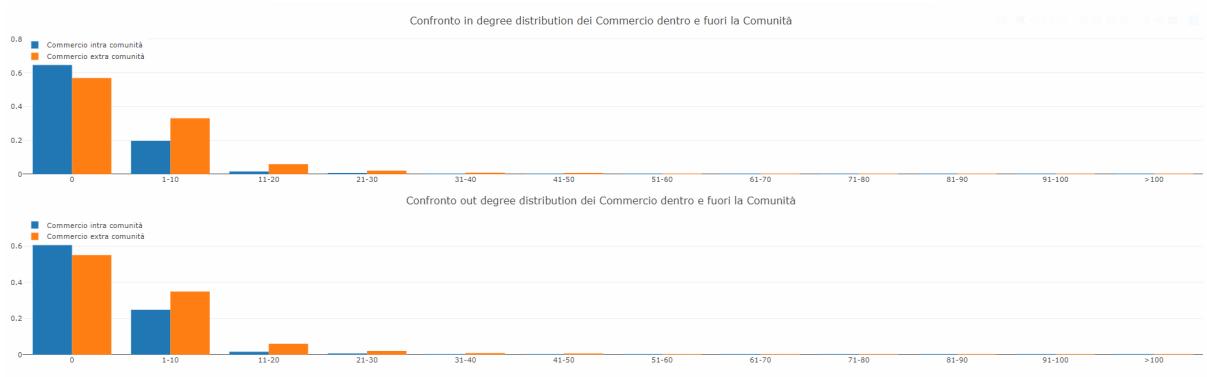


Figura 27: In ed Out Degree distribution dei messaggi intra ed extra community

5.5 Studio Singola Community

5.5.1 Introduzioni

Date le analisi effettuate sull'intera rete date tutte le alleanze, abbiamo deciso di **concentrare il nostro studio su di una singola community** al fine di poter valutare la

situazione interna in maniera più precisa e puntuale. Inoltre, grazie a questi studi, possiamo effettuare analisi volte a identificare gli elementi centrali quali i capi e i diplomatici, effettuando poi analisi sulla comunicazione degli stessi.

Ci siamo concentrati sull'**alleanza più grossa al primo giorno** che risultasse **presente su tutto il periodo**; abbiamo dunque estrapolato dalla rete di messaggi e commercio le relazioni interne a tale community.

5.5.2 User e Edge Trend

Data la rete dell'alleanza estrapolata dal grafo abbiamo analizzato il **numero di utenti** al fine di differenziare il numero di utenti segnati per quella alleanza (dato il file delle alleanze), **gli utenti attivi** nella rete e quelli **attivi internamente nell'alleanza**. I risultati ottenuti sono i seguenti:

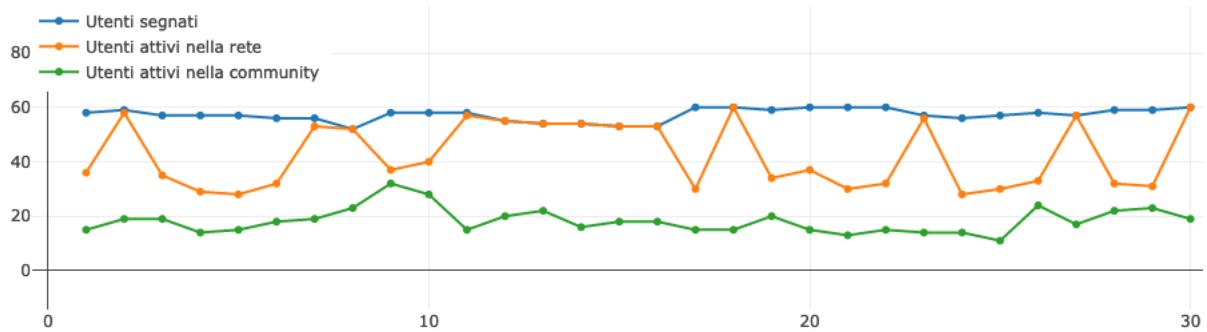


Figura 28: Numero utenti segnati, presenti nella rete e attivi nelle relazioni intra community

Possiamo notare come il numero di utenti che interagiscono in maniera attiva, mediante messaggi e commerci, siano sempre inferiori rispetto al numero di utenti attivi nella rete e segnati. Dunque possiamo osservare come **solo una parte degli utenti sia attiva nelle comunicazioni e nel commercio interni**. Ipotizziamo che in questa minoranza siano presenti gli elementi centrali.

Non possiamo, date queste analisi, definire la community più o meno attiva, non avendo ulteriori informazioni riguardanti l'utilizzo del forum.

Abbiamo anche valutato il numero di archi (messaggi e trade) sull'intero periodo ottenendo i seguenti dati:

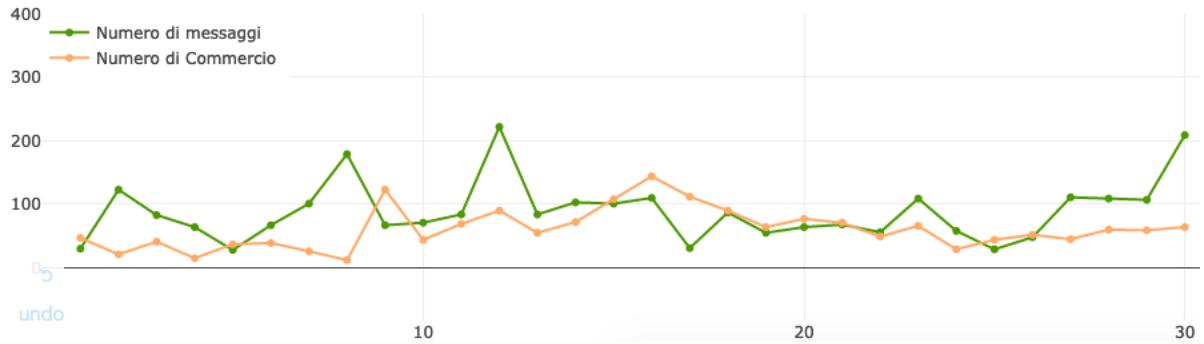


Figura 29: Numero archi dentro la community

Possiamo osservare come, mediamente, **il numero di messaggi scambiato sia superiore rispetto al numero di scambi commerciali effettuati**. Mediante lo studio di indici quali la reciprocità sarà possibile comprendere la tipologia di comunicazione.

5.5.3 Analisi Strutturale 1

L'analisi della **reciprocità** conferma la nostra ipotesi, il **commercio** presenta una reciprocità in generale **inferiore** rispetto ai **messaggi** in quanto all'interno della community esiste più un **approccio assistenzialista** verso chi ha più bisogno invece di uno scambio equo come con l'esterno attraverso il mercato il che porta ad essere una azione prevalentemente unidirezionale il commercio all'interno della community.



Figura 30: Reciprocità di messaggi e commerci all'interno della community

Coerentemente con quanto appena analizzato, analizzando la **densità** vediamo che **i messaggi hanno una densità maggiore** in quanto si tratta dell'**attività principale** all'interno della community; Essi vengono infatti probabilmente utilizzati per coordinamento tra utenti.

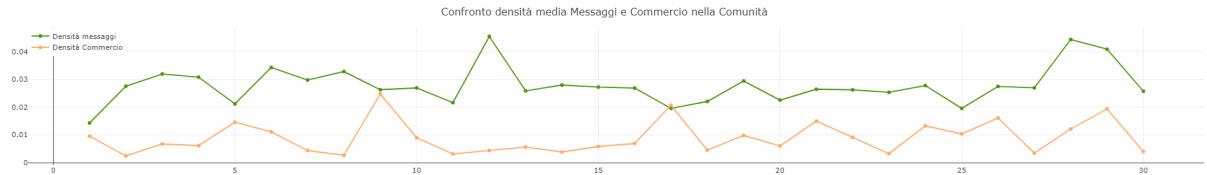


Figura 31: Densità di messaggi e commerci all'interno della community

Dall'analisi del Coefficiente di **Clustering** vediamo che per quanto riguarda i messaggi, riscontriamo **valori più alti rispetto al commercio**; è presente quindi una tendenza alla comunicazione privata, che siano i leader che dialogano con i sottoposti? Per quanto riguarda invece il **commercio** osserviamo un'ulteriore prova di **assistenzialismo** e **soccorsa** da certi utenti, probabilmente più avanzati ed esperti, ad altri.

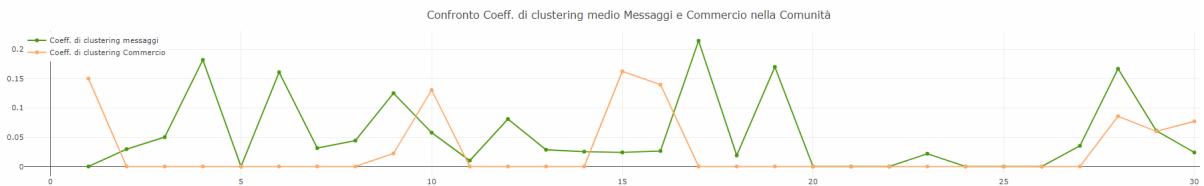


Figura 32: Coefficiente di Clustering di messaggi e commerci all'interno della community

5.5.4 Analisi Strutturale 2

Analizzando **Diametro** e **Shortest Path** all'interno della community ci aspettiamo **utenti particolarmente vicini tra loro**, e questo ci viene confermato dal diametro che si aggira intorno ad un valore di circa 2 volte il cammino minimo e con valori di Shortest Path quasi sempre intorno al 2 ed diametro mai inferiore al 3.

Per il commercio i valori sono più bassi, tale situazione non è dovuta ad una elevata interconnessione quanto più a un invio 1:1; situazioni di catena nel commercio sono più rare.

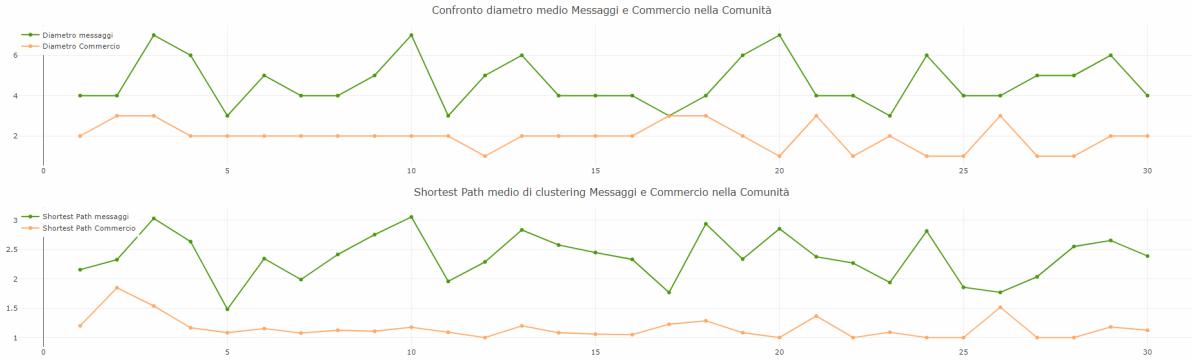


Figura 33: Diametro e Shortest Path di messaggi e commerci all'interno della community

Nel confronto messaggi e commercio osserviamo che **i messaggi sono generalmente più presenti** rispetto al commercio. Tale risultato rispecchia le considerazioni fatte nello studio delle relazioni intra-comunitarie.



Figura 34: Grado Medio di messaggi e commerci all'interno della community

Abbiamo poi valutato la distribuzione di grado per quanto riguarda i messaggi e il commercio all'interno della community sull'intero periodo; i risultati ottenuti sono i seguenti risultati:

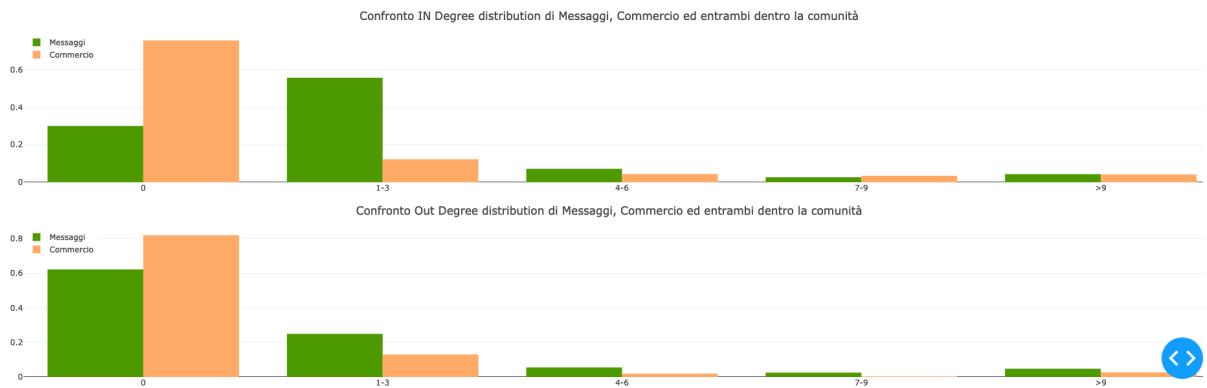


Figura 35: Distribuzione di grado

Possiamo subito notare come vi sia la solita tendenza all’inviare pochi messaggi ma a ricevne molti in ingresso; tale trend è stato già notato data l’analisi delle realtà intra ed extra community. Pensiamo dunque che in quella percentuale (pari al 24,4%) si possano trovare coloro che possimo andare ad identificare come leader. Possiamo notare come invece, all’aumentare della quantità di messaggi scambiati, le distribuzioni di grado siano molto simili; tale situazione pensiamo sia dovuta alla comunicazione tra i leader e i loro collaboratori.

Per quanto riguarda i messaggi possiamo notare una tiepida tendenza a favore della ricezione, dunque pare sia più probabile non inviare nulla che non ricevere nulla (82,07% versus 75,63%). Questa situazione pare distribuirsi poi sugli altri intervalli di quantità. Difatti al crescere delle quantità è più probabile ricevere grandi quantitativi piuttosto che inviarne; questo conferma che anche in questa alleanza è presente una tendenza all’assistenzialismo da parte di players, ipotizziamo, più forti nei confronti di altri meno.

5.5.5 Ricerca dei leader

A questo punto abbiamo deciso di metterci alla ricerca degli **elementi più centrali della rete** al fine di identificare i **capi dell’Alleanza**; abbiamo inoltre tentato di identificare i **diplomatici**.

Per effettuare queste ricerche ci siamo appoggiati alle **misure di centralità** presentate durante il corso, ad ogni studio di queste assoceremo la motivazione per il quale abbiamo deciso di utilizzare una certa misura. Per la ricerca dei **leader** abbiamo identificato come tali coloro che nel periodo risultavano per più volte i primi due date tutte le misure di centralità utilizzate; per l’identificazione dei **diplomatici** invece abbiamo valutato la edge betweenness dati i messaggi che collegavano i suddetti nodi ad altri di altre community. Di questi abbiamo selezionato i 4 che si presentavano più volte nelle prime posizioni.

Nelle immagini che andremo a proporre vedono già i nodi che riteniamo essere quanto sopra descritto, questa scelta è motivata dal fatto che pensiamo sia una più semplice rappresentazione rispetto al fornire i dati numerici.

Outdegree Date le precedenti considerazioni riguardanti l’esistenza intra-community di una tendenza all’utilizzo del messaggio in senso unidirezionale, ci siamo concentrati sulla centralità di grado (valutando solo l’outdegree) per ciò che riguarda i messaggi. Abbiamo incluso nelle considerazioni anche il commercio in quanto reputiamo che un leader debba essere anche mediamente più forte rispetto ai suoi sottoposti, tenendo in considerazione la tendenza assistenzialista, abbiamo deciso di valutare per entrambe i collegamenti, andando a premiare in funzione di questi. Di seguito alcuni **screen dei vari giorni con i nodi più grandi** a indicare un maggior indice di grado:

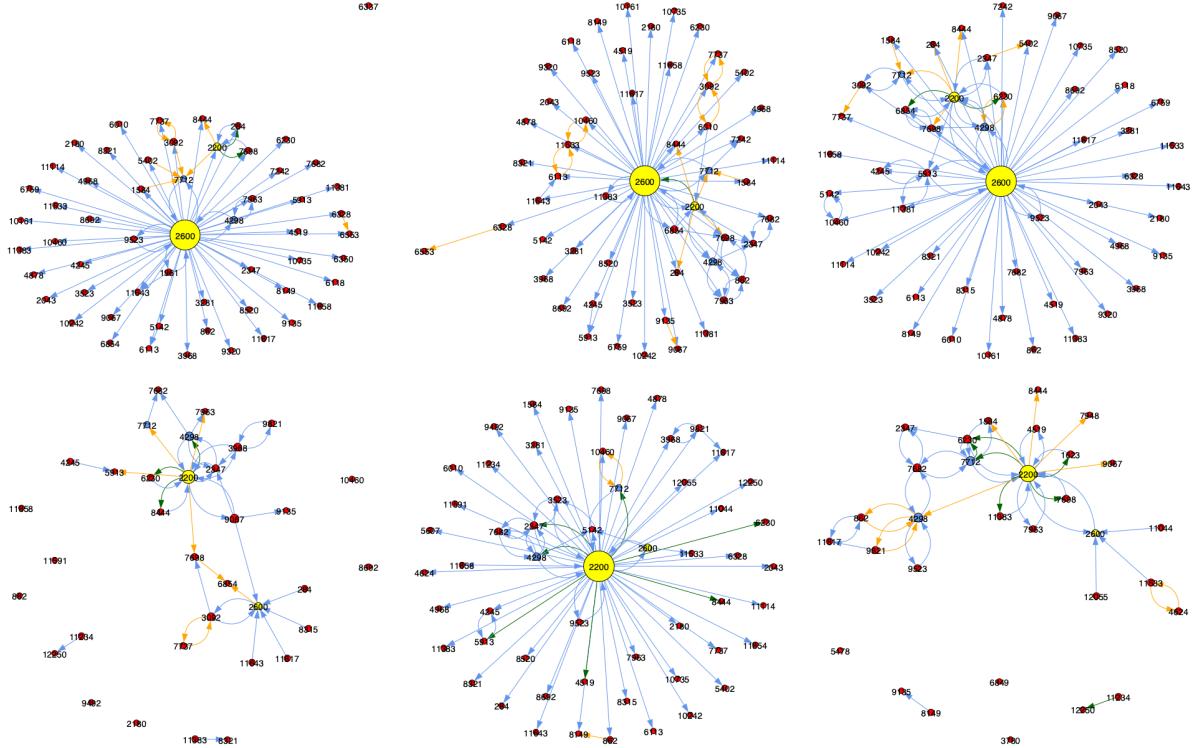


Figura 36: La community nei giorni 11, 13, 14, 19, 23 e 29

Gli archi blu denotano i **messaggi**, gli archi arancioni denotano il **commercio** gli archi verdi denotano la presenza di entrambi.

La possibilità di visualizzare la rete di una community conferma le nostre precedenti supposizioni riguardanti la tipologia di interazione intra-comunitario. Osserviamo un utilizzo dei **messaggi** dai leader ai sottoposti a forma di avviso in **broadcast** mentre il **commercio** è, nella maggior parte dei casi, **monodirezionale**; vi sono tuttavia **realità di scambio equo** come nei giorni 29 e 13 ad esempio.

Betweenness Un'altra misura tenuta in considerazione per gli **studi di ricerca dei leader** è la **betweenness**; abbiamo valutato questa per i messaggi, in quanto un nodo centrale per i messaggi data la realtà direzionale indica che questo coordina nodi che di fatto non comunicano tra di loro. Abbiamo deciso ancora una volta di tenere in considerazione anche il commercio, in quanto pensiamo che un nodo centrale a livello di betweenness dati questi archi sia un player con più probabilità di essere centrale e di rilievo all'interno dell'Alleanza.

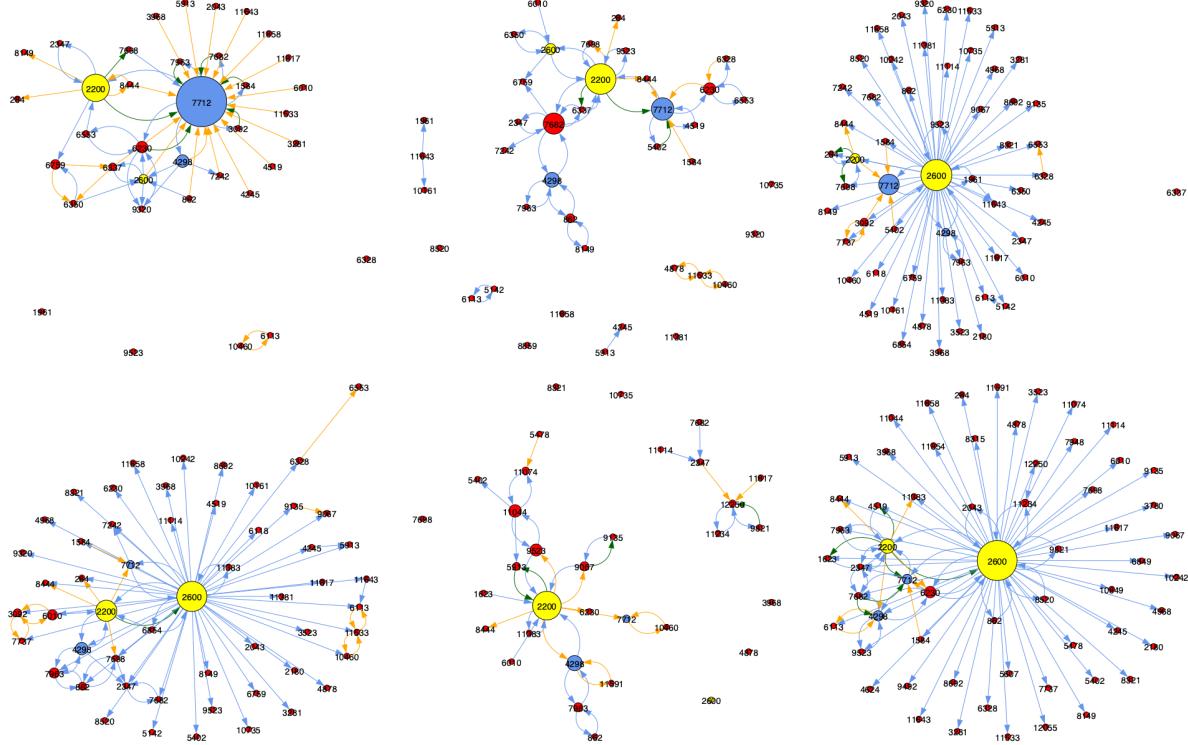


Figura 37: La community nei giorni 2, 7, 8, 13, 16 e 24

Possiamo osservare che i nodi già precedentemente identificati come centrali risultino esserlo nuovamente, inoltre i nodi identificati come diplomatici rivestano un ruolo importante, data questa misura, all'interno della community.

Closeness Questo indice di centralità fornisce informazioni riguardanti l'efficienza di un nodo a trasferire informazioni a tutti i nodi della rete; sostanzialmente ci dice quanto un determinato nodo risulti essere centrale fisicamente data la sua rete. Questa misurazione è stata **applicata ai messaggi** in quanto, date le analisi precedente, abbiamo un **trend unidirezionale dei nodi**, di conseguenza questo ci permetterà di visualizzare i nodi centrali data questa misura i quali sono poi coloro che effettuano questo ruolo di invio; abbiamo deciso di applicarla però anche agli archi di scambio commerciale in quanto abbiamo ritenuto, in maniera analoga a quanto detto prima, che considerare anche questo tipo di collegamento possa indirizzare in maniera più precisa e puntuale verso i **leader**.

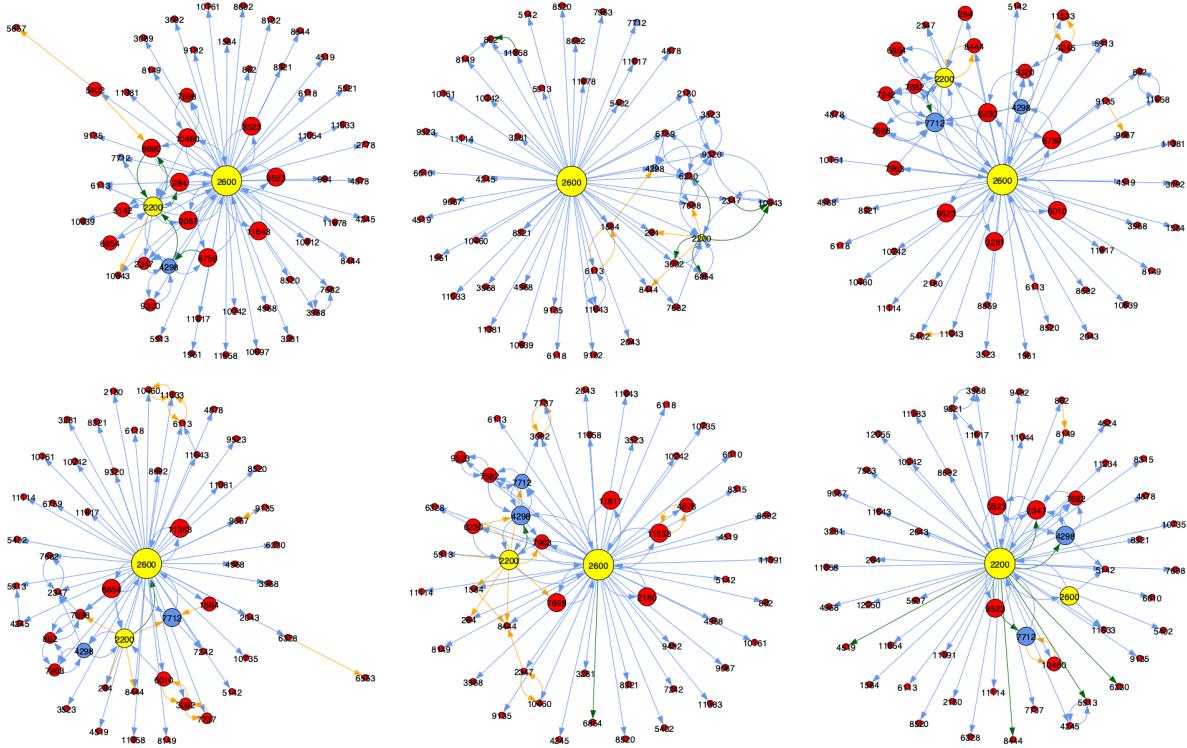


Figura 38: La community nei giorni 9, 10, 11, 13, 26 e 30

Ancora una volta i nodi precedentemente centrali nelle precedenti misure risultano tra i primi; notiamo che anche i nodi identificati come diplomatici, data questa misura, sono protagonisti. Notiamo anche la presenza di altri nodi centrali data questa misura, questi non hanno un ruolo definito; potrebbero comunque ricoprirlo. Sarebbe interessante fare ulteriori analisi.

PageRank Ultima misura di centralità riguarda il **PageRank**, questa è stata applicata solo ai **messaggi** in quanto riteniamo sia, data la misura di centralità in questione, l'unica che ha senso valutare. Il fatto che un utente nella rete scriva a un utente centrale non fa sì che anche il primo utente debba essere considerato importante o centrale, invece se il nodo centrale risponde o scrive ad un nodo allora ha senso valutare la distribuzione della propria importanza. Ovviamente, dato che abbiamo notato la tendenza unidirezionale dei messaggi, viene valutata la **numerosità dei messaggi** nella valutazione di questa misura (si definiscono quindi dei pesi).

Abbiamo deciso di applicare tale misura solo al **messaggi** in quanto riteniamo che in questo caso avremmo ottenuto informazioni di poco o nullo valore. Forniamo degli screen riguardanti i giorni più interessanti per possibili valutazioni:

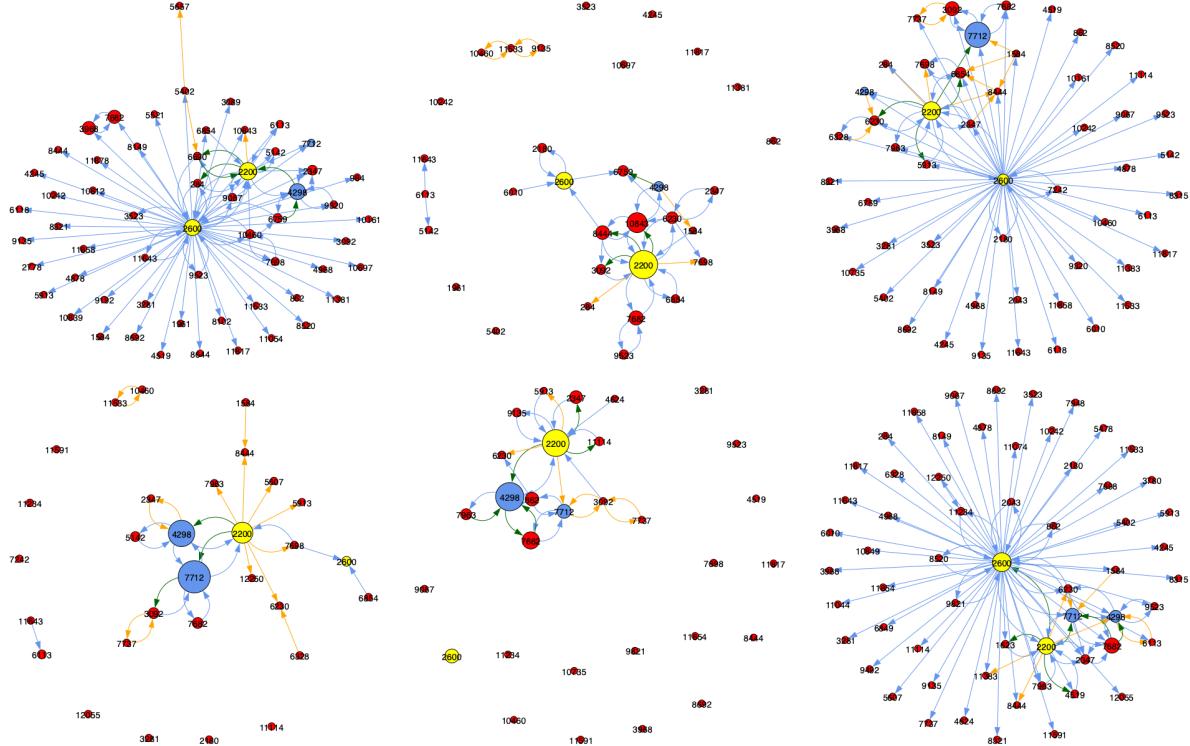


Figura 39: La community nei giorni 2, 6, 15, 17, 21 e 30

La tendenza di centralità segue le orme di quanto già visto precedentemente, tuttavia notiamo la presenza di nodi centrali per tale misura (risultavano centrali anche per misure precedenti, come ad esempio la Betweeness). Sicuramente questa alleanza ha altre figure “istituzionali”, date le informazioni in nostro possesso non possiamo fare analisi più approfondite.

5.5.6 Relazioni con i diplomatici

Date le due figure identificate all’interno della community, e nel tentativo di supportare ulteriormente le nostre tesi su queste, abbiamo pensato di **valutare le relazioni sull’intero periodo a livello di messaggi** tra coloro che abbiamo identificato come leader e quelli che abbiamo indentificato come diplomatici; in parallelo valutare le relazioni con gli utenti “normali”. **Ci aspettiamo di avere un valore maggiore di messaggi scambiati tra diplomatici e leader**; ovviamente il senso di analisi del collegamento vede come mittente il diplomatico o utente normale e come destinatario il leader. I risultati ottenuti sono i seguenti:

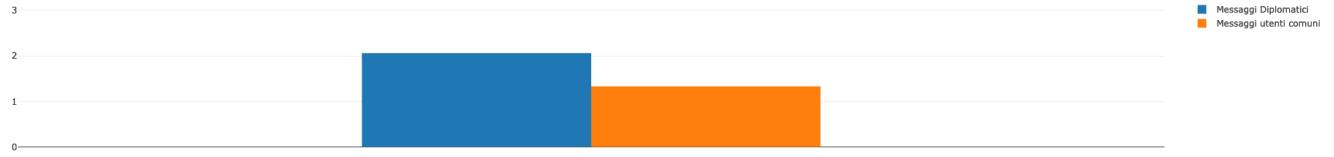
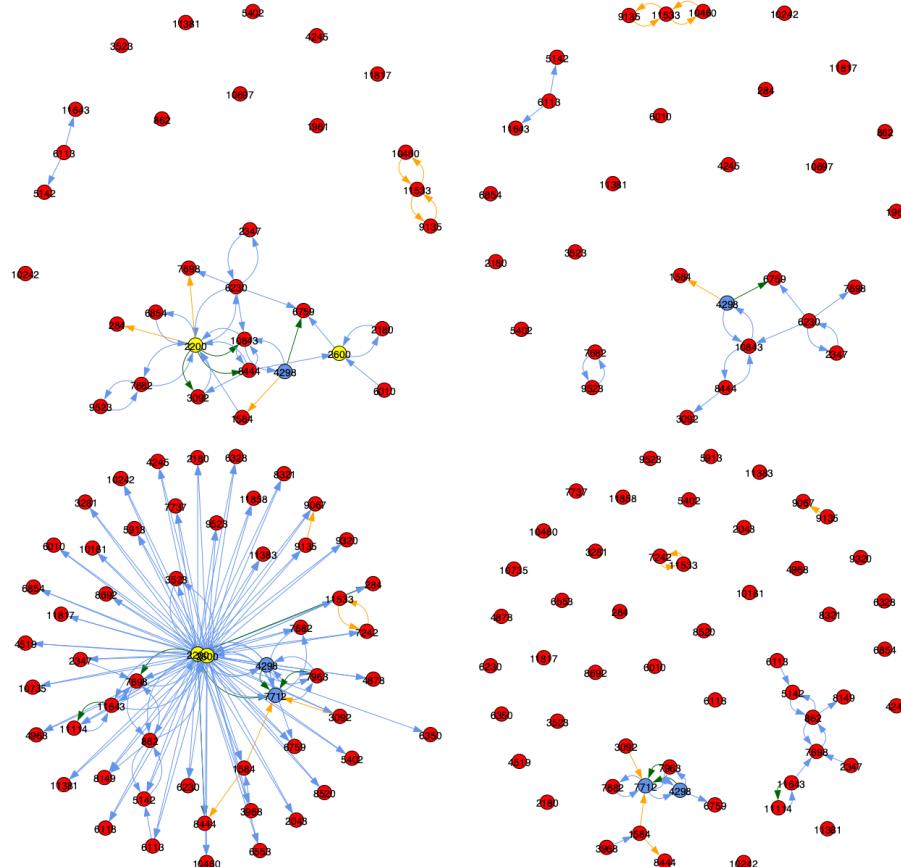


Figura 40: Numero di messaggi inviati dai diplomatici ai leader e dagli utenti normali ai leader

Un diplomatico, al giorno, invia 2 messaggi ai leader e invece un normale utente 1 (1.3). Questo dato rappresenta un’ulteriore prova del fatto che i diplomatici identificati sono, probabilmente, veri.

5.5.7 Eliminazione dei leader

In quest’ultima analisi, per lo più di natura ‘visiva’, abbiamo voluto provare ad **eliminare i nodi identificati come leader** al fine di paragonare la situazione pre e post tale eliminazione. L’idea iniziale era osservare cosa succedeva alla comunità, mediante anche uno studio di community detection, tuttavia il risultato ci è sembrato molto esplicito e abbiamo ritenuto sufficiente un approccio **prettamente visivo**.



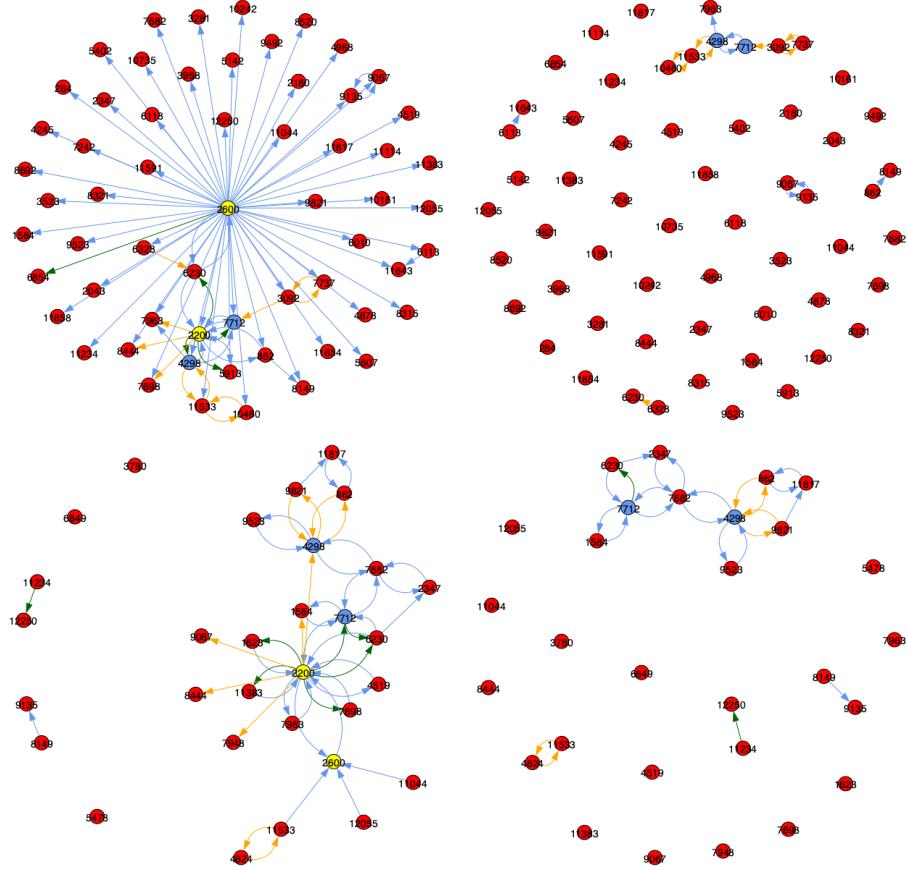


Figura 41: Confronto pre e post eliminazione giorni 6, 12, 18 e 29

Possiamo quindi visivamente notare come la tendenza ai messaggi unidirezionali, nei giorni in cui questi occorrono, **pare del tutto cancellata dall'eliminazione dei nodi identificati come leader**. Possiamo però notare come gli utenti identificati come **diplomatici continuino a ricoprire un ruolo attivo** all'interno della community, possiamo infatti notare come questi appartengano sempre alla componente attiva della community.

6 Possibili sviluppi futuri

Un'analisi di particolare interesse che si potrebbe realizzare è relativa al lato bellico ed alle azioni offensive intraprese tra varie community ed utenti, in particolare sarebbe interessante poter **analizzare come le interazioni offensive tra community possano essere in grado di modificare le gerarchie delle comunità coinvolte o addirittura farle dissolvere a seguito di sconfitte.**

Sempre relativamente alle azioni offensive, avendo a disposizione più informazioni relative ai singoli messaggi, si potrebbe cercare di **capiere in che modo gruppi di utenti appartenenti ad una comunità organizzano le loro offensive comuni** in modo da essere più efficaci. Inoltre sarebbe interessante poter indagare su alcune possibili relazioni di causa/effetto cercando correlazioni tra i vari tipi di azioni come ad esempio un possibile **legame tra messaggi e commercio o attacco.**

Il nostro lavoro è stato in parte limitato dalle **poche informazioni disponibili sugli utenti e sugli archi che li connettono**, avendo a disposizione maggiori informazioni come ad esempio gli esiti degli attacchi, il contenuto dei messaggi o ancora le quantità di materiale scambiato nei commerci si potrebbero analizzare in modo ancor più approfondito le dinamiche sociali (e non) all'interno del gioco ed affinare ulteriormente lo strumento di identificazione dei giocatori scorretti il quale non avendo a disposizione informazioni sulla quantità di materiale scambiato può essere facilmente **aggirato** inviando messaggi vuoti o scambiando quantità minime di materiali.

Bibliografia

- [1] Kiran Lakkaraju Hamidreza Alvari Rolf T. Wigand Nitin Agarwal Alireza Hajibagheri, Gita Reese Sukthankar. Using massively multiplayer online game data to analyze the dynamics of social interactions. URL <https://www.semanticscholar.org/paper/Using-Massively-Multiplayer-Online-Game-Data-to-the-Hajibagheri-Sukthankar/7b92bb07374cd17d9f9220a222ee4f8b4e01b26>.
- [2] Wikipedia. Travian. URL <https://it.wikipedia.org/wiki/Travian>.
- [3] Travian. Travian rules. URL <https://www.travian.com/it/gamerules>.