

Modellizzazione ad agenti del binomio vettore fruitore nel trasporto pubblico torinese

Alessandro Greco

Dipartimento di fisica
Università di Torino

Emanuele Pepe

Dipartimento di fisica
Università di Torino

Mattia Furlan

Dipartimento di fisica
Università di Torino

ABSTRACT

Il seguente lavoro ha come fine ultimo l'indagine delle dinamiche di interazione tra i fruitori del servizio di trasporto pubblico e il vettore che lo fornisce. Come di consueto, in ambito di sistemi complessi, l'argomento è difficilmente indagabile mediante un approccio fisico standard; per questo motivo si è deciso di ricorrere ad una simulazione multiagenti. In particolare abbiamo realizzato un modello di mobilità urbana volto a riprodurre le caratteristiche generali dello spostamento su mezzo pubblico; interpretandolo come dinamica su network di fermate e collegamenti. È stato posto l'accento sulla decisione di acquisto del titolo di viaggio, le strategie di controllo ed il loro impatto sulle prime. Tutto ciò attraverso una serie di esperimenti mirati a indagare la stabilità del modello e la sua risposta a una repentina modifica nel numero di agenti controllori; sul breve e lungo termine.

Agenti

Gli agenti coinvolti nella simulazione sono di 6 specie, per ognuna di queste segue una rappresentazione schematica riguardante: variabili, funzioni e inizializzazione.

Building:

- id: parametro intero di identificazione,
- type: parametro intero associato al tipo di edificio.

Bus_stop:

- code: parametro intero di identificazione,
- popolazione: parametro intero contenente la popolazione della fermata nel turno corrente,
- popolazione_old: parametro intero contenente la popolazione nel turno precedente,
- out_edges: lista di Edge in uscita dalla fermata,
- vicini_geo: lista di Bus_stop in un raggio di 250 m,
- list_raggiungibili: lista contenente i code delle Bus_stop raggiungibili in un turno,
- stop_raggiungibili: lista di Bus_stop raggiungibili in un turno che sostituisce list_raggiungibili dopo l'inizializzazione,
- fattore_scelta: lista di 6 float contenente i fattori che rappresentano una sorta di potenziale attrattivo della fermata per una specifica classe di utenti in funzione del rapporto tra popolazione attesa e quella reale,
- appeal_matrix: matrice di interi 6 x 10 contenente i fattori che rappresentano una sorta di potenziale attrattivo della fermata per una specifica classe di utenti in funzione della fascia oraria e della loro posizione relativa rispetto agli edifici di interesse.

Facciamo ora una overview delle azioni:

- set_raggiungibili, questa azione ha lo scopo di inizializzare list_raggiungibili a partire dalla matrice globale raggiungibili costruita in fase di compilazione,
- trova_out_edges, questa azione ha lo scopo di inizializzare out_edges identificando gli Edge che partono dalla fermata chiamante,
- trova_vicini_geografici, questa azione ha lo scopo di inizializzare vicini_geo eseguendo una query su tutti i Bus_stop e selezionando quelli a una data distanza dalla fermata chiamante,
- set_appeal_matrix, questa azione ha lo scopo di inizializzare appeal_matrix a partire dalla matrice globale Pesi costruita in fase di compilazione.

Trattiamo ora i riflessi:

- update_fattore_scelta, questo riflesso viene eseguito ogni turno e ha lo scopo di aggiornare il fattore_scelta, specifico della fascia oraria, in funzione del numero di agenti passeggeri che popolano la fermata e del potenziale della stessa.

Concludiamo dunque con l'init, in questo caso è collocato all'interno di quello globale così da evitare criticità nell'ordine di inizializzazione dei gruppi di agenti dovute all'architettura di Gama e la sua funzione è quella di eseguire in successione le azioni di inizializzazione delle diverse variabili.

Edge:

- popolazione: parametro intero adibito a contenere il numero di viaggiatori circolanti sull'Edge nel turno corrente,
- popolazione_old: parametro intero contenente il numero di viaggiatori circolanti nel turno precedente

- arrivo: Bus_stop di arrivo dell'Edge,
- penalty: parametro float moltiplicativo utilizzato dagli agenti Smith in fase di costruzione del percorso da monitorare; ogni turno viene aggiornato a 1.0,
- avg_penalty: lista di 4 float rappresentanti le penalità medie per macrofascia oraria; vengono utilizzati da GTT in fase di gestione dei controllori,
- btw: parametro float adibito a contenere la betweenness centrality dell'Edge,
- fascia_btwn: parametro intero che individua una fascia di betweenness compresa tra 1 e 3,
- score: parametro float moltiplicativo utilizzato da GTT in fase di gestione dei controlli,
- avg_score: lista di 4 float rappresentanti gli score medi per macrofascia oraria; vengono utilizzati da GTT in fase di gestione dei controllori,
- peso: parametro intero proporzionale al numero di corse reali sulla tratta,
- saturation_index: parametro intero indicativo della pienezza dei mezzi che transitano sull'Edge in un dato istante; è compreso tra 0 e 4.

Facciamo ora una overview delle azioni:

- set_betweenness, questa azione inizializza btw a partire dalla matrice globale bwn.

Trattiamo ora i riflessi:

- update_variables, questo riflesso viene eseguito ogni turno con lo scopo di aggiornare alcune variabili dell'agente; rimandiamo l'analisi approfondita alla sezione "Logica del modello".

Concludiamo dunque con l'init, in questa sezione l'agente Edge definisce la variabile fascia_btwn come $\min(2, \text{int}(\frac{3 \cdot btw}{btw_max}))$; risulta dunque un valore intero compreso tra zero e 2.

Passenger:

- type: parametro intero identificativo della tipologia di passeggero,
- the_target: parametro Bus_stop che rappresenta la destinazione del passeggero,
- fermata: parametro Bus_stop che identifica il nodo della rete dove si trova il passeggero,
- percorso: path contenente la successione di Edge designati dall'utente per giungere a the_target partendo da fermata; viene azzerato ogni turno,
- avg_mobility: parametro float, indica la mobilità mediata sulle fasce orarie del singolo passeggero,
- flag_movimento: parametro booleano vero nel caso l'agente decida di muoversi; resettato a false ogni turno,
- titolo_viaggio: parametro intero utilizzato nella struttura di movimento e controllo; decrementato ogni turno,
- travel_mode: parametro intero identificativo della modalità scelta dall'utente per spostarsi; posto a 3 ogni turno,
- flag_abbonamento: parametro booleano vero nel caso l'agente possieda un abbonamento,
- risk: matrice di float 3 x 5 che esprime la percezione del rischio per un insieme di Edge, specificato tramite fascia_btwn e saturation_index,
- viaggi: matrice di float 3 x 5 che tiene conto del numero di viaggi per un insieme di Edge, specificato tramite fascia_btwn e saturation_index
- avg_risk: parametro float associato alla percezione del rischio del passeggero mediata su tutti i viaggi,
- num_viaggi: parametro float, contatore pesato dei viaggi effettuati dal passeggero,
- utility: lista di 3 float che rappresentano gli score attesi dall'utente in seguito alla decisione di una specifica tipologia di viaggio.

Trattiamo ora i riflessi:

- decisione_abbonamento, in questo riflesso l'agente Passenger stima la spesa per i viaggi senza abbonamento nel mese successivo, e decide in base al confronto con il costo dell'abbonamento,
- movimento, questo riflesso è responsabile dello spostamento degli agenti Passengers nonché della loro decisione in merito a come muoversi. Il reflex viene eseguito qualora un'estrazione bernoulliana con probabilità:

$$1 + 0.4 \cdot \text{signum}(\text{titolo_viaggio})) \cdot \text{mobility}[\text{fascia_oraria}, \text{type}], \quad (1)$$

restituisca true. Soddisfatta la condizione si esegue una selezione random pesata su fattore_scelta, prima tra i Bus_stop in vicini_geo di fermata così da aggiornare fermata stessa, e a seguire sugli stop_raggiungibili della nuova fermata;

settando dunque `the_target`. Fatto ciò, viene presa una decisione su come raggiungere la destinazione; il che significa sostanzialmente utilizzare o meno il trasporto pubblico e nel caso acquistare o meno il biglietto. Il processo logico è basato sulla percezione di rischio dell'utente in funzione del path scelto nonché sulle tariffe GTT. In aggiunta si aggiornano tutte le variabili di sistema interessate come quelle di utility, risk, viaggi, popolazione di fermata e simili. Ovviamente il processo decisionale viene evitato qualora l'agente Passenger abbia un parametro `titolo_viaggio` diverso da zero; ergo un biglietto già validato.

GTT:

- `target_list`: lista di Edge contenente gli obiettivi per gli Smith istanziati,
- `utility`: parametro float adibito a contenere un valore numerico che rappresenta l'introito di GTT,
- `total_risk`: parametro float adibito a contenere la somma delle variabili `avg_risk` di ogni agente Passenger; viene aggiornato automaticamente ogni turno,

Facciamo ora una overview delle azioni:

- `crea_agente_Smith`: questa azione istanzia i controllori.
- `update_target_list`: questa azione aggiorna la lista di target in base all'assegnazione degli obiettivi ai controllori, penalizzando Edge già controllati,

Trattiamo ora i riflessi:

- `creation`, crea `n_controllori` Smith,
- `evaluation`, contiene pressoché tutte le azioni che l'agente GTT performa sul modello. In primis aggiorna il valore della utility sottraendo il compenso dei controllori; inoltre si aggiornano i target sui quali una porzione degli Smith sarà chiamata ad eseguire i controlli nei turni successivi. La logica che viene seguita prende in considerazione la prestazione dei singoli agenti Smith sulle specifiche tratte, puntando a modificare il comportamento dei peggiori. Inoltre l'agente GTT performa anche azioni "di sistema" come il reset a zero delle liste `score` e `avg_score` di Smith e Edge.

Smith:

- `score`: lista di interi costituita da 4 termini, adibita a contenere valori proporzionali alla prestazione del controllore nella specifica macrofascia oraria,
- `percorso`: lista di Edge che costituiscono il percorso seguito dall'agente Smith,
- `targets`: lista di Edge contenente i target dell'agente in funzione della macrofascia oraria.

Facciamo ora una overview delle azioni:

- `acquire_new_targets`, questa azione ha lo scopo di restituire in funzione della macrofascia oraria un nuovo target, selezionato dalla `target_list` di GTT. In seguito vengono incrementate le penalità per gli Edge entro 100m ed aggiornata `target_list` di GTT coerentemente,

Trattiamo ora i riflessi:

- `turn_shift`, questo riflesso ogni 5 ore colloca l'agente Smith nella location del target corrispondente alla corretta macrofascia oraria,
- `movimento`, questo riflesso costruisce il tragitto dell'agente Smith come successione di 7 Edge. Il percorso viene generato tramite un albero di Edge percorribili a partire dal target, e selezionando alla fine la sequenza di Edge contigui con uno score complessivo maggiore. Inoltre viene eseguito il codice responsabile dell'aggiornamento della penalità degli Edge controllati dall'agente Smith.
- `verification`: questo riflesso viene eseguito ogni ciclo e il suo scopo è quello di verificare i "titoli di viaggio" degli utenti Passenger che nel turno precedente hanno condiviso, anche solo parzialmente, il tragitto dell'agente Smith. I passeggeri controllati vengono estratti tra quelli in movimento, che condividono almeno un Edge con l'agente Smith, in modo random ed in funzione di popolazione e peso dell'Edge interessato. Tali agenti subiscono un incremento del parametro `risk` relativo alle specifiche dell'Edge: `saturation_index` e `fascia_btwn`. Nel caso di contravvenzione si incrementa anche il conteggio di tre Passenger estratti random e si aggiornano le variabili "di sistema" quali `utility`, `titolo_viaggio`, `score` e `travel_mode`.

Global:

- grafo: contenitore della rete di trasporto, costituita da Bus_stop (nodi) e Edge (link),
- fascia_oraria: ogni giornata è suddivisa in 10 fasce orarie da 2 ore, dalle 05:00 alle 01:00,
- macro_fascia_oraria: ogni giornata è suddivisa in 4 macrofasce orarie da 5 ore, dalle 05:00 alle 01:00,
- ticket: costo del biglietto in euro,
- abbonamento: costo dell'abbonamento in euro,
- multa: ammontare della multa in euro,
- n_controllori: numero di agenti Smith istanziati,
- n_passeggeri: numero di passeggeri istanziati,
- dimenticanza: base dell'esponenziale decrescente dei pesi sulle grandezze avg_risk e num_viaggi,
- dimenticanza2: base dell'esponenziale decrescente dei pesi sulle grandezze risk e viaggi,
- betw_max: valore massimo della betweenness centrality, usata nella normalizzazione delle fasce di betweenness per gli Edge,
- scala: scala computazionale tra la rete di trasporti reale e quella simulata,
- index: variabile di sistema indicativa della simulazione in corso,
- active_stops: Bus_stop dalle quali è possibile raggiungere altre Bus_stop,
- input_csv: booleano che indica l'inizializzazione del modello da file csv,
- salvataggio: stringa del percorso di salvataggio per la simulazione corrente.

Passiamo ora all'inizializzazione del modello:

- Assegnazione delle variabili globali index, n_controllori, multa, fascia_oraria, ticket, abbonamento e scala,
- import dei Building dagli shapefile opportuni,
- creazione dei passeggeri,
- creazione degli Edge e della struttura di network direzionato e pesato, calcolo della fascia_btwn,
- esecuzione della routine di inizializzazione degli agenti Bus_stop,
- eliminazione matrici globali di inizializzazione,
- creazione GTT,
- calcolo di active_stops.

La routine viene eseguita in maniera congrua al valore di input_csv, importando nel caso le variabili dai file di input.

Per una descrizione dettagliata della logica che sta alla base delle azioni e dei riflessi fondamentali per il modello si rimanda alla sezione specifica del paper.

Generic view

Costruzione network

La rete di trasporto pubblico torinese è stata modellata su un grafo diretto pesato dove i nodi sono le fermate e, qualora esista un mezzo che collega due vertici, si genera un link. In particolare, i pesi di questi ultimi sono proporzionali al numero di mezzi che transitano su una data tratta nel corso di una giornata. In questo modo, lo spostamento di un utente attraverso la rete di trasporto viene mappato in un cammino tra due nodi nel grafo del modello. Operativamente, il grafo è stato creato a partire dai dati GTFS disponibili online presso <http://www.gtfs-data-exchange.com/agency/gruppo-torinese-trasporti/>; nello specifico tra i file disponibili per il download sono stati utilizzati "stop_times.txt" e "stops.txt". Il primo è una collezione di viaggi eseguiti dal trasporto pubblico, nel quale vengono identificati univocamente i singoli viaggi effettuati dai mezzi (parametro "trip_id"), l'ordine di percorrenza delle fermate (parametro "stop_sequence") e il tempo intercorso tra due fermate successive. Il peso del collegamento tra due fermate viene incrementato di 1 ogni volta che una coppia di fermate orientata compare in un viaggio; l'output di questa elaborazione è una edge list riferita all'intera provincia di Torino. (<https://developers.google.com/transit/gtfs/reference>) Successivamente, per geolocalizzare i nodi e ridurre il settore di studio aderendo alle risorse computazionali disponibili, abbiamo scaricato dei file di natura territoriale nel formato shapefile riguardanti il comune di Torino all'indirizzo <https://it.wikipedia.org/wiki/Shapefile>. Il software usato per la manipolazione di questi files è Qgis (<http://www.qgis.org/it/site/>). I file usati contengono, oltre ai punti critici del trasporto pubblico, numerosi siti e informazioni legate a questo e allo spostamento dell'utenza. La discussione sull'uso di queste informazioni è rimandata alle sezioni successive.

Rischio

Il rischio su ogni singolo Edge viene concepito come rapporto tra incontri con gli agenti Smith e viaggi intrapresi; inserendo un effetto di dimenticanza sugli avvenimenti più remoti. La realizzazione di questa logica è affidata a una media pesata:

$$rischio_{jk} = \frac{\sum_{i=1}^{viaggi_{jk}} X_i \cdot dimenticanza^{viaggi_{jk}-i}}{\sum_{i=1}^{viaggi_{jk}} dimenticanza^{viaggi_{jk}-i}}; \quad (2)$$

dove con X_i si identifica un booleano pari a 1 qualora sia avvenuto un incontro con un controllore. I pesi variano in modo esponenziale al progredire del numero di viaggi intrapresi dallo specifico utente, infatti il numeratore di questo rapporto per $viaggi_{jk} + 1$ é:

$$dimenticanza \cdot \sum_{i=1}^{viaggi_{jk}} X_i \cdot dimenticanza^{viaggi_{jk}-i} + X(viaggi_{jk} + 1). \quad (3)$$

Poiché dimenticanza è un valore minore di 1 la legge esponenziale è decrescente, nello specifico la base si è scelta in modo tale che il peso di un incontro sia dimezzato dopo aver percorso 50 Edge appartenenti alla stessa classe di betweenness e saturation_index; indicati nella trattazione precedente dagli indici j e k. Il rischio assegnato ad un percorso, costituito da una successione di Edge contigui, è associato alla probabilità che il numero di controlli su tale tratta sia maggiore o uguale a uno. In termini probabilistici:

$$P(\#controlli \geq 1) = 1 - P(\#controlli = 0) = 1 - \prod_{Edge \in percorso} 1 - risk_{Edge}. \quad (4)$$

Avg_risk è, invece, una media temporale dei rischi percepiti sui singoli percorsi da parte di uno specifico utente. Questa variabile viene sottoposta alla medesima logica di dimenticanza; la base esponenziale è stata dedotta dalla precedente misurando che un viaggio dura in media 10 Edge.

Logica del modello

In questa sezione daremo una visione generale del funzionamento del modello concentrandoci sulla logica basilare con particolare attenzione in merito alle scelte fatte in fase di simulazione. Vista la finalità del modello, la chiave di volta è l'agente Passenger che rappresenta il fruitore dei servizi; colui che può effettivamente decidere se muoversi e con che modalità. Come già accennato in fase di descrizione dell'agente i riflessi fondamentali che governano questa dinamica sono decisione_abbonamento e movimento. Il primo è stato strutturato, nel processo di sviluppo del modello, in diverse versioni di complessità crescente. La scelta della filosofia definitiva che si è adottata è stata presa sulla scia dell'argomentazione di teoria dei giochi presentata in modo esaustivo in appendice. L'unica differenza riguarda l'iterazione del gioco per tutti i turni decisionali che l'agente passeggero valuta di dover affrontare. Per quanto riguarda la seconda decisione, come già accennato in fase di descrizione del riflesso, viene presa solamente dai passeggeri che in un dato turno si trovino a muoversi privi di un abbonamento. Questo avviene qual'ora la seguente relazione restituisca true:

$$flip((1 + 0.4 \cdot sign(titolo_viaggio)) \cdot mobility[fascia_oraria, type]) \quad (5)$$

Abbiamo deciso di favorire gli agenti già in possesso di un titolo di viaggio di un fattore arbitrario 0.4 che rappresenta un incremento del 40% della probabilità di movimento. Una volta nel riflesso, il passeggero modifica la sua posizione riassegnando la variabile fermata in modo random, tra tutte le Bus_stop geograficamente prossime a quella sulla quale si trova; quest'ultima compresa. La motivazione di questo passaggio sta nell'incrementare la flessibilità del movimento così che gli agenti possano visitare più agevolmente una porzione maggiore delle destinazioni fruibili. Il passo successivo è molto simile e consiste in una scelta analoga alla precedente ma fatta sui raggiungibili della nuova fermata; mediante la funzione di GAMA path_between si costituisce dunque quello che sarà il percorso del passeggero. Deciso il tragitto è necessario capire come coprirlo, per prima cosa si aggiornano le variabili di rischio in funzione degli Edge che costituiscono il path designato.

$$risk_max \leftarrow 1 - \prod_{edge \in path} \left(1 - \frac{risk[saturation_index, fascia_btwn]}{viaggi[index, fascia_btwn]} \right) \quad (6)$$

$$avg_risk \leftarrow dimenticanza \cdot avg_risk + risk_max \quad (7)$$

Il valore medio di rischio viene moltiplicato per il parametro dimenticanza così che le decisioni pregresse assumano una risonanza minore sulla decisione del passeggero; secondo la legge esponenziale sopra descritta. Valutato il rischio si compila la lista di utility per designare la metodologia di spostamento, secondo le seguenti espressioni:

$$e^{-length(percorso.edges)} \quad (8)$$

$$1 - e^{\frac{-length(percorso.edges)}{(risk_max \cdot multa)}} \quad (9)$$

$$1 - e^{\frac{-length(percorso.edges)}{ticket}} \quad (10)$$

L'espressione 8 corrisponde alla decisione di compiere il tragitto senza usufruire dei mezzi pubblici e come si può vedere risulta inversamente proporzionale alla lunghezza di quest'ultimo secondo una funzione esponenziale. L'espressione 9 corrisponde alla decisione di prendere i mezzi ma non pagare il biglietto, la utility in questo caso viene favorita dalla lunghezza del percorso ma risulta inversamente proporzionale al prodotto rischio \cdot multa. L'ultima possibilità, viaggiare acquistando il biglietto, viene codificata in 10 ed è analoga alla precedente ma il prodotto a denominatore viene sostituito dal costo del titolo di viaggio. Nuovamente si incappa in un grado di libertà del modello fissato a nostra discrezione. La scelta sembrerebbe sfavorire fortemente la modalità di viaggio 1, ma questo in parte è giustificato dal fatto che il modello ha come finalità lo studio del sistema di trasporto pubblico; e dunque del comportamento dell'utenza interessata a usufruire dei mezzi. Osservando le altre due equazioni si può ridurre il confronto ai denominatori ed essendo i parametri multa e ticket tra i pochi definibili a partire da una situazione reale il confronto delle due utility ruota intorno alla percezione di rischio. Tutta questa struttura, come già detto assolutamente discrezionale, è stata ideata in modo tale che i valori di utility siano confrontabili e compresi tra zero e l'unità; con un andamento sensato in funzione dei parametri del modello. Variare queste funzioni, magari in seguito a uno studio in ambito sociale significativo, potrebbe rappresentare un buon punto di partenza per uno sviluppo futuro. Al termine di questo processo decisionale avvengono tutti gli aggiornamenti delle variabili interessate come: il titolo di viaggio, la utility di GTT qualora si sia acquistato un biglietto o un abbonamento, la posizione dell'agente passeggero sulla rete di fermate, la popolazione delle fermate di arrivo e partenza e simili.

Fino ad ora abbiamo descritto un verso del legame tra agenti passeggeri e GTT, il viceversa è invece mediato dall'operato di controllo gestito dall'agente GTT tramite l'istanziamento degli agenti Smith e l'assegnazione dei target sui quali eseguire i controlli. Il numero di controllori che vengono istanziati è un parametro fisso del modello, così come il costo del biglietto e dell'abbonamento, eventualmente modificabile mediante esperimenti batch; il secondo punto di controllo viene codificato nel riflesso evaluation. In primis viene decrementata la utility di GTT, sottraendo il costo settimanale dei controllori che noi abbiamo stimato ammontare a 625 euro. Fatto questo GTT riassegna target_list iniziando ogni elemento lista, corrispondente a una specifica macrofascia oraria, con gli agenti Edge in ordine decrescente di $avg_score \cdot avg_penalty$.

$$target_list[macro_fascia] \leftarrow n_Edges \text{ sort_by } (-avg_score[macro_fascia] \cdot avg_penalty[macro_fascia]). \quad (11)$$

A seguire si ripristina la penalty in modo tale che a ogni turno la penalità associata ad ogni Edge diminuisca, analiticamente ciò si traduce nell'incremento del termine moltiplicativo avg_penalty di un valore 0.1 tra un minimo di 0.0 e un massimo di 1.0. Si giunge ora al core del riflesso: per ogni data macrofascia oraria tutti gli agenti Smith superiori alla media, in termini di score, mantengono i target sui quali operare e si limitano a diminuire avg_penalty degli Edge entro i 100 metri di un fattore 0.5. Per quanto riguarda i controllori sotto la media, questi penalizzano i target rivelatisi infruttuosi, riducendo avg_penalty di 0.5, e in seguito eseguono la funzione acquire_new_targets. In questo frangente, tramite un'estrazione bernoulliana, decidono se acquisire un nuovo target in maniera random, tra quelli non penalizzati, o se selezionarlo tramite la target_list di GTT.

$$targets \leftarrow GTT.target_lists \text{ sort_by } (-avg_score \cdot avg_penalty)). \quad (12)$$

In questo modo gli Edge per i quali il coefficiente moltiplicativo è unitario, e dunque massimo, vengono favoriti; ed ecco palesato il ruolo di `avg_penalty`. Il passo successivo è aggiornare proprio questo valore riducendolo di 0.9 per gli Edge selezionati e di 0.5 per quelli geograficamente prossimi; come prima entro 100 metri. L'intera struttura di selezione ha la finalità di favorire il controllo degli Edge sui quali i controllori trovano il numero maggiore di morosi e, al contempo, di sfavorire l'agglomerarsi degli agenti Smith in una sola porzione della rete. Questa logica permette di esplorare la rete, in modo tale da evitare l'istaurarsi di zone di illegalità poiché sempre esenti dal controllo, e al contempo di monitorare gli Edge che si dimostrano essere più importanti al fine della utility di GTT. Gli ultimi riflessi fondamentali da esplicitare sono il movimento degli agenti Smith, codificato in movimento, e l'azione di controllo dei titoli di viaggio; `verification`. Partendo dal primo vengono definite tre variabili di "sistema", `path_score`, `path_score_max` e `available`, rispettivamente due float necessari per il selezionamento del percorso e una variabile ausiliaria il cui unico scopo è quello di contenere gli Edge disponibili per la costruzione del percorso. La logica è iterativa, partendo dal target si considerano tutti gli Edge in uscita da quest'ultimo e dai suoi vicini geografici; questo per incrementare la flessibilità del controllore. Segue una serie di 7 cicli annidati che, iterando su tutti gli Edge, esplorano i percorsi possibili e computano per ognuno il valore di `path_score`; ovvero la somma di tutti gli score dei singoli Edge costituenti il percorso moltiplicati per `penalty`. Si riassegna dunque il valore di `penalty` lungo tutto il percorso dimezzandola. Questa struttura logica permette di assegnare a ogni controllore un target univoco e fa in modo che vengano fortemente sfavorite le intersezioni lungo il tragitto. Affrontiamo ora il riflesso `verification`, per ogni Edge costituente il percorso del dato agente Smith questo richiede a $48 \cdot \frac{\text{edge.popolazione}}{\text{edge.peso}}$ Passenger, che abbiano ivi transitato, di incrementare il proprio rischio di un'unità; come di consueto in base alla fascia di rischio e di saturazione del tratto. Si compila dunque la lista morosi che va a contenere tutti gli agenti Passenger privi di biglietto il cui tragitto comprenda l'Edge in questione. Si incrementa quindi lo score dell'agente Smith, e la utility di GTT, di un valore pari a multa per ogni moroso acciuffato; al fine di tenere in considerazione i limiti operativi di un individuo reale si è deciso di imporre un numero massimo di controlli pari a 4. Fatto ciò si selezionano i morosi multati in modo random, il loro titolo_viaggio viene portato a 3 mentre il `travel_mode` lo si pone a 2. Inoltre questi domandano a una Terna di passeggeri selezionati a caso di modificare la percezione di rischio e il numero di viaggi nella classe di appartenenza dell'Edge sul quale i morosi in questione sono stati multati,

$$\text{risk}[\text{saturation_index}, \text{fascia_btwn}] \leftarrow \text{risk}[\text{saturation_index}, \text{fascia_btwn}] + 0.5, \quad (13)$$

$$\text{viaggi}[\text{saturation_index}, \text{fascia_btwn}] \leftarrow \text{viaggi}[\text{saturation_index}, \text{fascia_btwn}] + 1. \quad (14)$$

La motivazione alla base di questa implementazione logica è quella di tenere in considerazione la relazione tra agenti che, nella realtà dei fatti, condividono esperienze e si influenzano vicendevolmente. Mediante la comunicazione, l'esperienza di un passeggero che non è di fatto stato multato su una data tratta viene equiparata a quella del passeggero colto in flagrante.

Riscaldamento

Una delle problematiche che si sono dovute fronteggiare in fase di simulazione è stata, come di consueto quando si lavora in questo ambito, il limite computazionale delle macchine a nostra disposizione. In linea di massima, per simulare un sistema come quello di trasporto pubblico, la cosa migliore sarebbe istanziare un numero di utenti consistente con l'ordine di grandezza reale e farli muovere sulla rete reale. Questo è particolarmente vero in fase di verifica, il termine validazione sembra eccessivo nel contesto del modello stesso. Nello specifico il numero di utenti che usufruiscono giornalmente del trasporto urbano ammonta a 640000 unità, questi si muovono su un network costituito da 3152 fermate, 4562 edge e vengono controllati da 180 dipendenti del Gruppo Torinese Trasporti. Alla luce di come si è deciso di strutturare gli agenti questi sono, per noi, valori inarrivabili; si è dunque reso necessario riscaldare il sistema. In primis si è deciso di lavorare su una porzione ridotta della rete locale tale per cui il numero di fermate totale è 5.69 volte il numero di quelle considerate. Riscaldando a sua volta i passeggeri questi risultano 112578 e ciò implica che, considerandoli uniformemente distribuiti tra i 40 step giornalieri sui quali si è strutturata la dinamica del modello, ogni turno debbano muoversi circa 2811 agenti. Considerando che solo il 10% dei passeggeri istanziati si muove ad ogni turno, il valore totale è di 28119. Facciamo dunque un ragionamento analogo per i controllori, ogni giorno solamente $\frac{6}{21}$ dei controllori assunti sono in servizio simultaneamente, per questioni di orario e giorni lavorativi, e questo corrisponde a circa 51.42 agenti. Riscaldando secondo la proporzionalità dataci dalla riduzione del numero di fermate si passa a 9.05 controllori e dunque 4.5 coppie. Detto questo, per mantenere le proporzioni reali del sistema in esame, è necessario ridurre linearmente i valori dei passeggeri e dei controllori; l'unico modo per poter procedere alla simulazione con i nostri mezzi è utilizzare un solo Smith e di conseguenza 15000 passeggeri. L'implementazione del modello è però orientata alla gestione di più controllori, basandosi le valutazioni e le strategie su grandezze medie. Quindi, al fine di mantenere variabilità nelle strategie, abbiamo deciso di deviare dalla scala analitica istanziando 4 agenti controllori anziché uno; consci dell'impossibilità, in questo framework, di fare valutazioni quantitativamente significative.

Esperimenti

Considerazioni preliminari

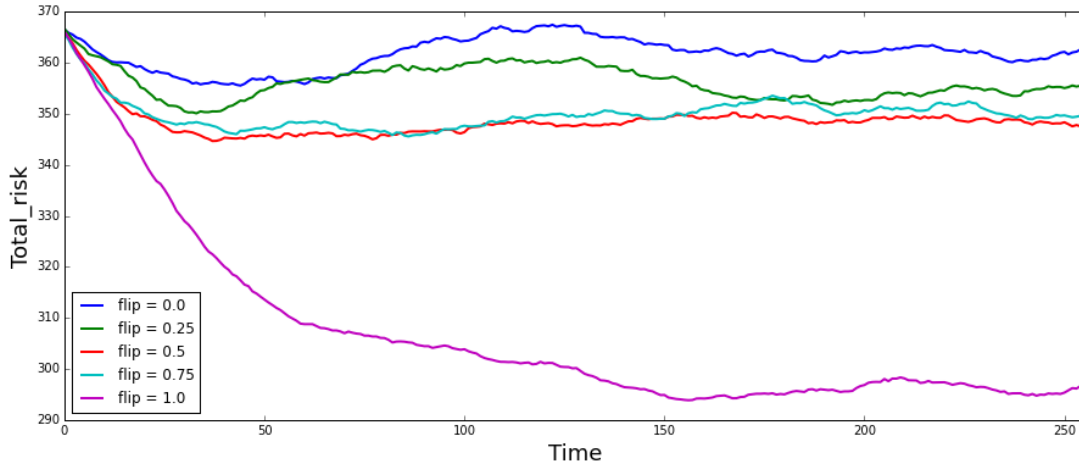


Figura 1. Batch sul flip del controllore

Il primo esperimento che si è performed si propone di dare una valutazione orientativa in merito all'efficacia delle specifiche strategie attuabili dai controllori nella selezione degli Edge sui quali operare. Nella loro riallocazione i meno efficienti in termini di score decidono con una determinata probabilità, la quale è stata variata tra 0.0 e 1.0 a step di 0.25, di selezionare in maniera random un Edge come target. Per i nostri grossi limiti computazionali le prove sono state effettuate singolarmente. Consci del fatto che i risultati così ottenuti non abbiano valenza statistica, eseguire una simulazione di ensemble che stimi gli errori e fornisca un parametro più preciso sarebbe certamente un passo da compiere per ulteriori sviluppi del modello. Osservando i grafici ottenuti a partire dai dati raccolti () si può notare come sussista una netta sovrapposibilità tra tutte le curve, con unica eccezione per quelle ottenute imposto flip a 1.0, corrispondente ad una scelta completamente casuale dei target. Questo ci porta a osservare come la strategia di scelta deterministica sia probabilmente più efficiente di una scelta completamente random. Il presunto legame tra la dimensione della rete e le strategie di controllo sarebbe opportunamente indagabile cambiando la dimensionalità e la topologia del grafo sul quale si svolge la dinamica del sistema; anche questo rappresenta un naturale sviluppo dell'indagine che con questo lavoro si è voluta iniziare. Alla luce dei limiti appena sottolineati, a prescindere dalla somiglianza delle curve, il fatto che si sia deciso di eseguire tutte le simulazioni che seguono con flip pari a 0.5 è da considerarsi a tutti gli effetti un grado di libertà del nostro modello.

Batch per la stabilità

Il primo esperimento vero e proprio riguarda lo studio della stabilità del modello, questo è stato fatto monitorando la variabile rischio. Si è raccolto un campione di 20 simulazioni analoghe ottenute cambiando il seed per l'estrazione dei numeri random, scegliendo tra due possibili valori per questioni tecniche alla base del batch, e le condizioni iniziali sulla posizione degli agenti. Nello specifico, per ognuno dei due blocchi di simulazioni caratterizzati dal seed scelto, la prima è stata inizializzata in maniera random mentre le successive hanno preso in input le posizioni degli agenti al termine della precedente. Il tempo di ognuna delle run è di 146000 cicli, corrispondenti a poco più di 10 anni, e completato il campione, si è calcolata la curva media e la sua deviazione standard. Fatto ciò si sono fittati i valori con un andamento esponenziale:

$$rischio = A \cdot \left(1 - \exp^{-\frac{t}{\tau}}\right), \quad (15)$$

il quale ci restituisce i parametri:

$$A = 346.6 \pm 0.4, \quad \tau = 6772 \pm 9. \quad (16)$$

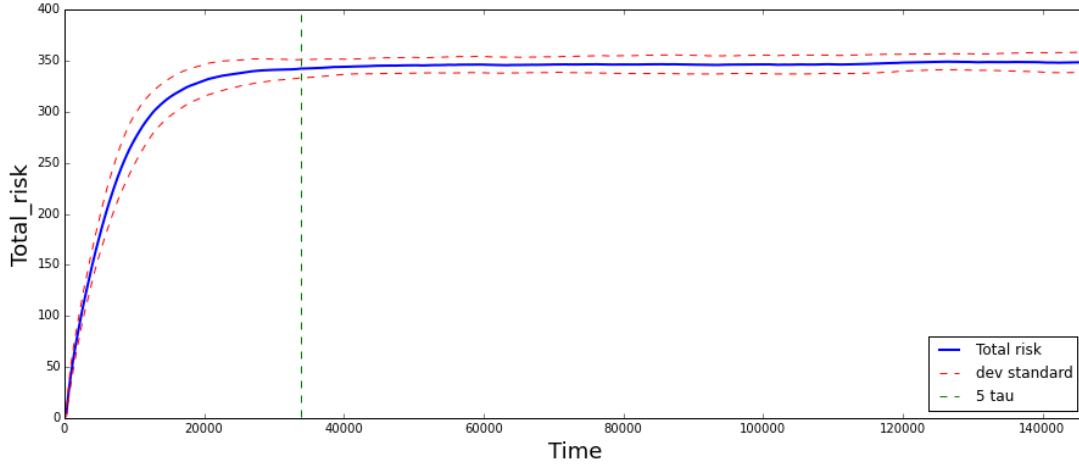


Figura 2. Andamento del rischio medio percepito

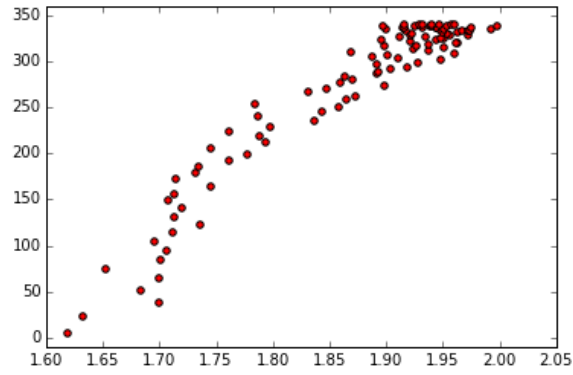


Figura 3. Nuvola di punti generate da rischio totale e abbonati

Questi dati ci hanno permesso di stimare, come tempo di reazione del nostro sistema, un valore di 5τ . É doveroso sottolineare come la variabile migliore al fine di studiare il comportamento degli agenti passeggeri, e del modello in genere di conseguenza, sia il numero di abbonati. Questo poiché identificativo delle modalità di spostamento, esattamente come il numero di biglietti venduti, ma al contempo meno rumoroso e oneroso in termini computazionali. Per valutazioni di natura analoga si è deciso di passare, dalle prime considerazioni in termini di abbonamenti, ad un'analisi incentrata sulla variabile di rischio. Prima di scegliere questo parametro si sono fatte, però, alcune considerazioni in merito alla correlazione tra le grandezze in gioco (Figura:3); così da legittimare la nostra scelta. In primis si sono calcolati, sulle misure raccolte per l'esperimento appena descritto, alcuni coefficienti standard di correlazione:

$$\rho_{Spearman} = 0.31, \quad r_{Pearson} = 0.86. \quad (17)$$

La differenza tra questi due parametri ci ha spinti ad indagare ulteriormente. La fase iniziale della simulazione, compresa entro i 5τ , mostra una forte correlazione, rispettivamente ranked e lineare:

$$\rho_{Spearman} = 0.75, \quad r_{Pearson} = 0.95. \quad (18)$$

Questo lo si può motivare supponendo che la stabilizzazione del modello riduca la risonanza del rumore su entrambe le grandezze osservate. Giunti a regime la compagine stocastica intrinseca al modello risulta più significativa, inoltre le diverse interazioni non lineari tra questo fattore e le grandezze osservate, conducono a una riduzione della correlazione osservabile

mediante i coefficienti considerati:

$$\rho_{Spearman} = 0.29, \quad r_{Pearson} = 0.27. \quad (19)$$

Nello specifico lo stimatore di Spearman risulta basso se valutato sull'intero dominio poiché, valutando il ranking di tutte le singole coppie di valori, viene influenzato dalla grossa porzione di dominio che definiamo stabile dove i dati non seguono un andamento crescente o decrescente ben definito a causa del rumore. Per concludere, la correlazione forte osservata nella fase iniziale della simulazione e le giustificazioni fornite per motivare la correlazione debole successivamente osservata, crediamo siano sufficienti a legittimare la nostra scelta in termini di analisi.

Utility

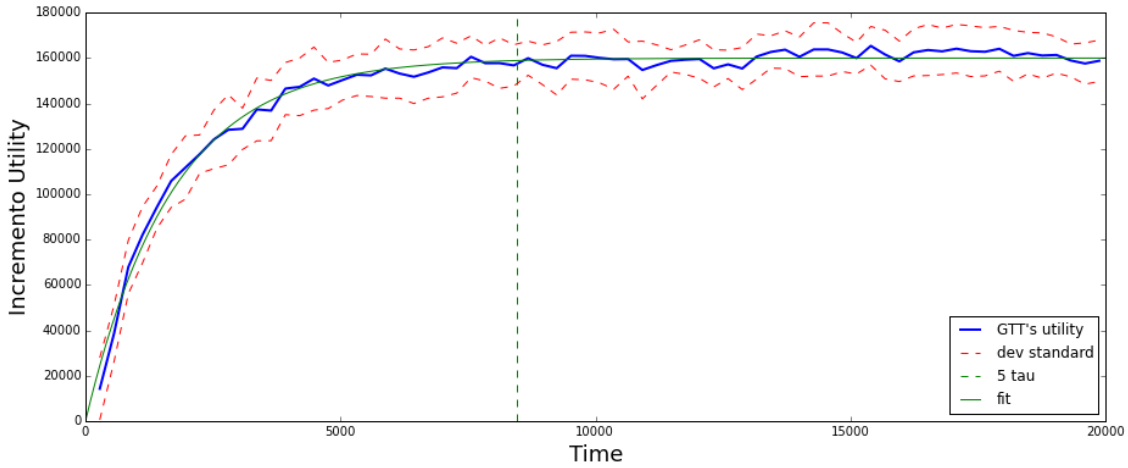


Figura 4. Utility media di GTT con 4 controllori su 10 simulazioni

L'ultimo esperimento che si è deciso di eseguire riguarda lo studio dell'andamento della utility di GTT in funzione del numero di controllori istanziati. Si è presupposto un andamento consistente con la curva di Laffer che, in ambito economico, prevede sussista un livello di prelievo fiscale oltre il quale l'attività economica non risulti più conveniente; e di conseguenza l'introito dello stato ne risulti sfavorito. Nel nostro modello, invece di considerare il rate di tassazione e la base, viene modulato il numero di controllori; con conseguente incremento del gettito e delle spese. È ovvio che, essendo la massima cifra guadagnabile da parte di GTT finita, la curva che descrive l'andamento della utility tornerà a valere zero per un numero specifico di agenti Smith; ciò che non è scontato è la forma del sostegno il quale può intersecare più volte l'asse delle ascisse. Essendo l'indagine di questa zona dello spazio delle fasi troppo onerosa in termini di calcolo, a causa del peso computazionale associato ad ogni controllore, ci siamo concentrati sulla ricerca del massimo della curva così da capire quale sia il numero ottimale di controllori da istanziare.

In primis, osservando le curve di utility in funzione del tempo derivante dalle misure precedenti, si è visto come il tempo di convergenza ad una condizione di stabilità per questa variabile sia differente dal τ precedentemente calcolato. Questa è un fenomeno che emerge sovente in ambito di sistemi complessi i quali mostrano, in genere, competizione tra scale. Lo si è dunque dovuto stimare prima di fare considerazioni più accurate per i singoli valori. Per far ciò abbiamo operato come in precedenza, su 10 esperimenti con 4 controllori, ottenendo i parametri:

$$A = 159999 \pm 1292, \quad \tau = 1691 \pm 106. \quad (20)$$

Alla luce di queste considerazioni abbiamo deciso di utilizzare, come condizioni iniziali, dei dati compatibili con il maggiore delle due scale temporali in esame; ovvero quella relativa alla variabile di rischio. Dunque abbiamo ispezionato lo spazio delle fasi tramite un esperimento batch che ci permettesse di studiare l'andamento della utility in funzione del numero di controllori. Nello specifico, abbiamo variato il loro numero da un minimo di 1 a un massimo di 19 con step unitari;

procedendo poi per 15000 step. Proseguendo con l'esperimento si è selezionata la porzione di dominio più plausibile in funzione del valore assunto dalle utility all'ultimo step temporale; eseguendo dunque una valutazione di ensemble. Siamo consci del fatto che essendo questa considerazione fondata su singole realizzazioni la magnitudine delle utility in esame non è un valore identificativo di un trend medio; ciò nonostante si tratta di un'ipotesi di partenza sensata. Si sono dunque eseguite 10 simulazioni per ogni numero di controllori tra quelli selezionati, queste sono state impostate analogamente alle precedenti tranne per il tempo di conclusione che è stato fissato a 20000 step. Il primo tentativo di valutazione del massimo incremento della utility lo si è fatto fittando i dati raccolti, assunti lineari, e valutando i coefficienti angolari. Questo approccio si è rivelato inconcludente poiché il rumore sui dati inficiava troppo pesantemente le conclusioni inferibili rispetto al campione esiguo di dati. Si è dunque cercato un approccio meno sensibile alle fluttuazioni stocastiche valutando direttamente la utility; ovvero la grandezza integrale associata al coefficiente angolare. Nello specifico si sono considerati due lassi temporali: i primi 6 mesi di simulazione e l'ultimo anno. Con i dati così raccolti si sono ottenute le due curve (*Figura:5, Figura:6*) che rappresentano

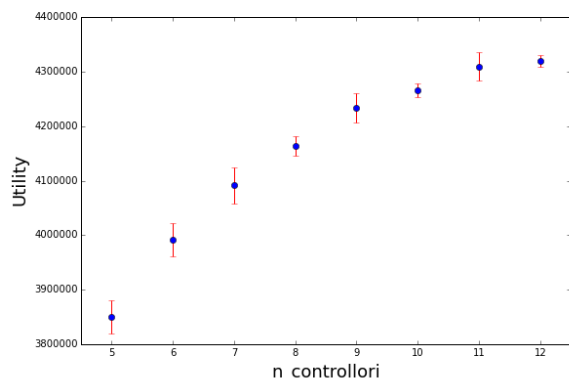


Figura 5. Utility entro i primi sei mesi di simulazione

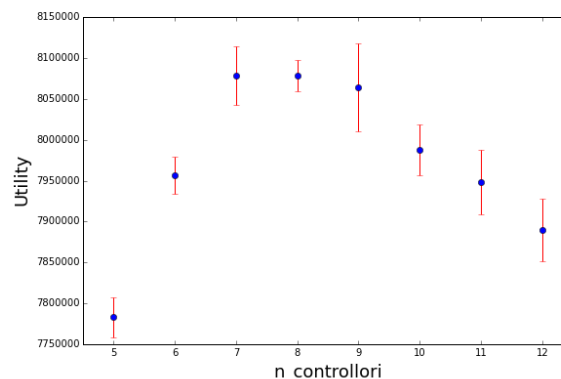


Figura 6. Utility entro gli ultimi 12 mesi di simulazione

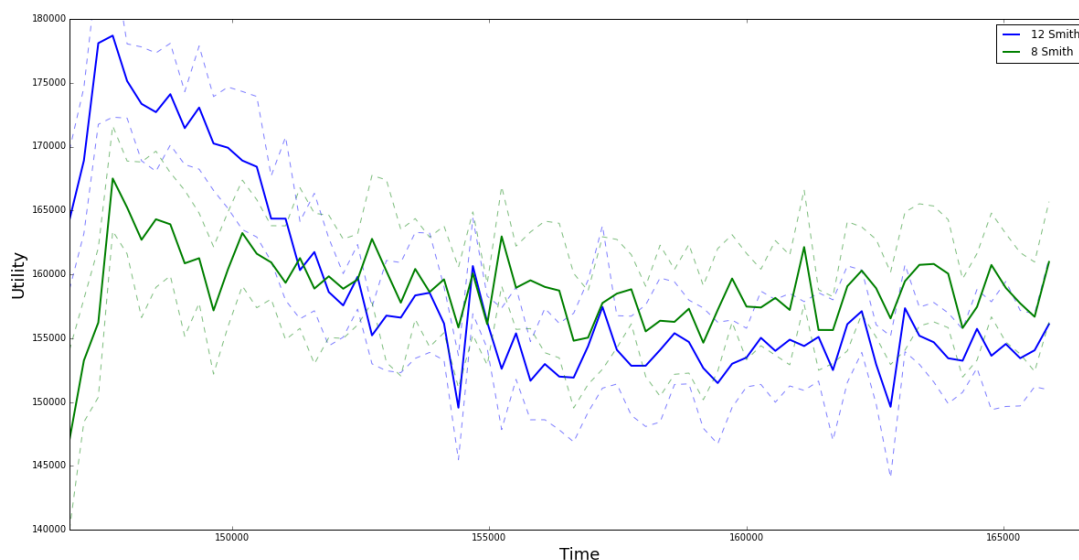


Figura 7. Curve rappresentative della utility al variare del numero di controllori

l'esito delle politiche di maggiore e minore controllo. Si osserva come, a breve termine, l'introito di GTT si proporzionale

al numero di controllori istanziati; questo è dovuto all'impatto di una politica di controllo più stringente partendo da una condizione ad elevata percentuale di morosi. Se si considera invece il risultato sul lungo termine la dinamica non è monotona e si osserva un massimo nel range 8 - 10 controllori; ciò significa che il comportamento degli utenti è cambiato e gli introiti diminuiscono a causa del costo degli agenti Smith. È interessante notare, osservando le curve rappresentative della utility 7, come le funzioni integrali catturino il picco che viene a generarsi nella fase iniziale della nuova politica di controllo. L'ampiezza di questa zona si mostra direttamente proporzionale al numero di controllori e questo è il motivo numerico alla base del fenomeno appena descritto. Si può fare una considerazione simile notando come, superata questa zona, le curve tendano ad invertirsi con valorizzazione conseguente delle politiche di controllo con numero di agenti Smith ridotto.

# Smith	UtilityShortTerm(10^6)	UtilityLongTerm(10^6)
5	3.85 ± 0.03	7.78 ± 0.04
6	3.99 ± 0.03	7.96 ± 0.02
7	4.09 ± 0.04	8.08 ± 0.05
8	4.16 ± 0.02	8.08 ± 0.02
9	4.23 ± 0.03	8.06 ± 0.05
10	4.27 ± 0.02	7.99 ± 0.03
11	4.31 ± 0.03	7.95 ± 0.05
12	4.329 ± 0.012	7.89 ± 0.03

Tabella 1. Risultati dell'analisi su utility

Tornando allo short term, studiarne l'estensione temporale del picco potrebbe essere uno sviluppo interessante così da avere un'informazione aggiuntiva in merito al tempo di risposta del nostro modello.

Conclusioni e prospettive

Procediamo, in conclusione, ad un'analisi cumulativa del lavoro fatto.

Pregi del modello

Uno dei pregi più significativi del lavoro svolto, a nostro avviso, è il fatto che si basi su un framework network. Questo lo rende facilmente implementabile tramite gli strumenti di teoria reti e, inoltre, permette agevolmente di cambiare la topologia in esame. Un secondo punto forte dell'analisi riguarda la risposta del modello, in primo luogo questo mostra convergenza ad una condizione di stabilità entro un tempo accettabile. Inoltre, in seguito ad una variazione del numero di controllori, si è osservata una reazione del modello coerente con le attese; ad esempio un controllo più serrato comporta una crescita della percezione del rischio proporzionale all'incremento dei controllori.

Difetti del modello

Per quanto riguarda le problematiche del nostro modello la maggiore tra queste riguarda l'arbitrarietà dei meccanismi di movimento e controllo; oltre che dei parametri arbitrariamente introdotti per simulare il comportamento reale dell'utenza. Non avendo potuto lavorare su dati reali, provenienti da indagini sociali o direttamente dal GruppoTorineseTrasporti, tutte le meccaniche da noi ideate sono state dettate dalla logica personale. Sicuramente questo al momento rappresenta un fertile terreno di dibattito; di positivo si può annoverare come gli spostamenti degli agenti siano consistenti con quanto atteso in funzione dei punti di interesse e delle fasce orarie. Gli altri problemi del nostro modello sono stati principalmente dettati dalla limitatezza delle risorse di calcolo. In primis si ha poca significanza statistica, questione risolvibile eseguendo validazioni di ensemble più consistenti e per ogni parametro, e inoltre si è reso necessario un riscaldamento del sistema reale. Non essendo questa una procedura scontata, nuovamente discutibile inquanto fondata su considerazioni personali, sarebbe preferibile riuscire a evitarla; incrementando la potenza di calcolo sfruttata.

Risultati e sviluppi

Tirando le somme il risultato più significativo ottenuto ad ora riguarda le strategie di controllo. Abbiamo osservato, tramite il nostro modello, come esista una configurazione capace di massimizzare l'introito di GTT qualora si debba confrontare con un'utenza razionale. A questo si aggiunge una considerazione collaterale: è emerso dalle nostre simulazioni come il guadagno del vettore dei servizi non sia massimo per una condizione di completa legalità. Parafrasando, al Gruppo Torinese Trasporti conviene che esista una compagine di agenti truffaldini persuasi del fatto che viaggiare senza titolo di viaggio convenga; così da poterli multare e speculare su questo. Per un eventuale sviluppo sarebbe opportuno migliorare il network alla base del modello simulando le linee reali, eventualmente tramite una struttura multigrafo. Sarebbe inoltre auspicabile l'eliminazione delle molteplici approssimazioni di campo medio, come quella con la quale si sono stimate le corse su un dato Edge, così da essere più aderenti alla realtà. In ultima istanza, i dati raccolti con le molteplici simulazioni, ci portano a credere che l'operato degli agenti puramente razionali non sia completamente rappresentativo della logica reale adottata dagli utenti del servizio di trasporto pubblico. Per questo motivo uno sviluppo interessante sarebbe introdurre una logica più complessa, per la scelta della modalità di viaggio, capace di simulare l'impatto di politiche sociali e di comparare la loro efficacia in funzione del costo; eventualmente confrontando il tutto con delle analoghe politiche di controllo.

Funzioni arbitrarie del modello

Matrici

Ci siamo trovati davanti al problema di capire quali fossero le ragioni che spingono gli utenti a muoversi verso una specifica zona. Ci è dunque sembrato opportuno, data la mancanza di dati empirici, affidarci al nostro buon senso nel predire gli interessi degli utenti.

Per effettuare questa operazione abbiamo preferito scomporre quanto più possibile il problema e scrivere una serie di matrici di nostro pugno, basate sull'idea che un utente, a seconda dell'orario, si muova con maggiore probabilità verso un centro d'interesse (scuole, ASL,...).

Riportiamo dunque di seguito le matrici utilizzate:

- *orario*: cerca di stimare quanto alta sia l'attività di un centro d'interesse a seconda dell'orario (ad esempio le università dopo le 8 sono chiuse mentre gli ospedali rimangono sempre aperti).
- *affinity*: in questa matrice abbiamo cercato di stimare quanto ogni tipo di utente provi interesse per un dato tipo di edificio (uno studente universitario sarà attratto dall'università mentre una persona anziana tenderà ad andare a teatro la sera).
- *pesi*: viene generata a partire dalle informazioni contenute nelle strutture appena descritte, il suo scopo è di attribuire ad ogni *Bus_stop* una matrice che determini il suo peso in funzione degli edifici d'interesse entro 500 m di distanza.

$$\sum_{\text{Edifici entro 500 m}} \text{orario}[fascia, edificio.type] \cdot \text{affinity}[utente, edificio.type] \cdot \left(1 - \frac{\text{distanza}}{\text{distanza_critica}}\right). \quad (21)$$

- *mobility*: attribuisce ad ogni tipo di utente la probabilità di muoversi in una data fascia oraria.

Variabili di percezione

In questa sezione descriveremo due variabili cruciali al fine del corretto funzionamento del modello che rientrano, per come definite, nella compagine delle funzioni da noi arbitrariamente introdotte. La prima è *saturation_index*, valore numerico intero compreso tra 0 e 4. Viene aggiornato ogni turno per ogni elemento Edge secondo la seguente logica:

$$\text{saturation_index} \leftarrow \text{int} \left(\text{floor} \left(4 \cdot \frac{1}{\frac{100}{\text{scala}}} \cdot \frac{\text{popolazione_old}}{\frac{\text{peso}}{48}} \right) \right) \quad (22)$$

Il fattore 4 è necessario per normalizzare a 4 il valore di saturazione, il termine *popolazione_old* è, invece, indicativo del numero di utenti che popolano il dato Edge in un turno e viene comparato tramite la frazione con il numero atteso. Questo valore lo si ottiene a partire da *peso*, che rappresenta una stima reale del numero di utenti transiti su un dato collegamento nell'arco di una giornata, diviso il numero di mezzogiornate; questo in accordo con la durata di un turno di spostamento che nel nostro caso vale 30 minuti. L'ultimo termine doveroso da commentare è il rapporto $\frac{100}{\text{scala}}$, il numeratore rappresenta la dimensione dei mezzi che viene dunque riscalata. La seconda variabile in questione è *fascia_btwn*, anche in

questo caso si tratta di un intero compreso tra 0 e 2 ma non necessita di aggiornamento in quanto proprietà del singolo Edge. Viene definita in fase di inizializzazione di tali agenti come:

$$fascia_btwn \leftarrow \min(2, \frac{3 \cdot \text{int}(btw)}{betw_max}), \quad (23)$$

dove il termine numerico ha lo scopo di fare in modo che più di un Edge abbia betweenness massima; per questo motivo non si è seguita la logica precedente moltiplicando per 2. Lo scopo di queste variabili, come ampiamente emerso nelle sezioni precedenti di questo lavoro, è principalmente quello di classificare gli Edge che costituiscono la rete. Inoltre l'indice di saturazione del mezzo entra in gioco anche nella valutazione dello score come evidenziato nella sezione specifica. La ragione per la quale si è deciso di discretizzare questa coppia di variabili è stata di restituire una caratteristica intrinseca degli agenti reali; ovvero la percezione imperfetta della realtà.

Score degli Edge

La funzione che abbiamo deciso di implementare per la valutazione dello score degli Edge è la seguente:

$$score \leftarrow \frac{\left(\frac{btw}{betw_max}\right)^{\frac{1}{3}} \cdot saturation_index}{\exp^{2.131 \cdot saturation_index} + 1}. \quad (24)$$

Questo funzionale è completamente arbitrario, si è deciso di ricorrere a questa forma analitica in modo da penalizzare la verifica dei target eccessivamente popolosi o sgombri. Questo serve a rispecchiare una dinamica reale che si osserva sui mezzi pubblici dove un eccesso di passeggeri rende difficoltoso il transito dei controllori.

Decisioni in ottica di Game Theory

La scelta di un agente di fare o meno l'abbonamento si basa sulla previsione di ciò che avverrà nel mese successivo al momento della decisione. Non essendo possibile un calcolo esplicito dei viaggi futuri e del loro costo, si rivela necessario eseguire una previsione quantitativa. Il costo dell'abbonamento viene confrontato con il costo della mobilità urbana nel mese successivo se l'agente dovesse eseguire decisioni puntuali ad ogni viaggio. Al fine di eseguire questa stima è possibile interpretare le scelte in questione come round di un gioco tra GTT e passeggero. Le possibili strategie per l'utente sono tre: fare il biglietto (T), non farlo (NT) o viaggiare in maniera alternativa; effettivamente quest'ultima non viene considerata poiché da un'analisi a posteriori emerge che non viene mai giocata da alcun utente. Per quanto riguarda l'agente Smith le possibili strategie sono controllare il passeggero (C) o non farlo (NC); segue la matrice dei payoff. Viene rappresentato

	C	NC
T	-ticket	-ticket
NT	-multa	0.0

Tabella 2. Matrice dei payoff

solo il payoff dell'utente, opposto a quello di GTT trattandosi di un gioco a somma nulla. La strategia di GTT, secondo la percezione dell'agente, è riducibile a una variabile aleatoria distribuita in modo binomiale e associata al numero di controlli che subisce da parte degli agenti Smith. La miglior stima di questa probabilità di controllo è proprio il rischio medio calcolato; in base a questo valore il passeggero decide quale strategia adottare secondo la seguente logica. Assumendo che la strategia sia la probabilità di fare il biglietto ($p(T)$), ogni round l'agente lancerà una moneta truccata e in base all'esito acquisterà o meno il ticket; inseriamo $p(T)$ nell'equazione del payoff atteso e massimizziamo così da trovare la probabilità ottimale.

$$Payoff_utente = -p(T) \cdot p(C) \cdot ticket - p(T) \cdot p(NC) \cdot ticket - p(NT) \cdot p(C) \cdot multa + 0.0 \cdot p(NT) p(NC), \quad (25)$$

Considerando che $p(NC) = 1 - p(C)$ e $p(NT) = 1 - p(T)$, semplificando si ottiene:

$$Payoff_{utente} = -p(T) \cdot ticket - (1 - p(T)) \cdot p(C) \cdot multa. \quad (26)$$

L'equazione è una retta la quale, agli estremi del dominio chiuso e limitato, presenta estremi assoluti. La strategia dell'agente dunque è binaria: fare sempre il biglietto o non farlo mai. Il punto critico, in questo fenomeno di transizione di fase, coincide con l'uguaglianza tra i payoff associati alle due diverse strategie; questo avviene qualora $p(C) = \frac{ticket}{multa}$. Al di sopra di questo valore l'agente riterrà opportuno fare il biglietto.