

Modello di Simulazione ad Agenti applicato ai comportamenti sociali determinanti la dinamica di pandemie (2 aprile 2020)

Autori dello studio:

Federico Cecconi - CSO QBT Sagl

Alessandro Barazzetti – Head of R&D, QBT sagl – Professor in Data analysis and advanced forecasting techniques, UCM Higher Education Institution Malta and Swiss Campus – Gen.Dir. Science Adventure Sagl Switzerland

Introduzione

Abbiamo realizzato un modello basato su agenti artificiali che descrive alcuni aspetti della dinamica di diffusione del COVID-19, in particolare gli aspetti legati al legame tra comportamenti sociali e dinamica di contagio. Ci sono due specificità nella struttura socio-economica dei territori del nord Italia colpiti inizialmente dalla pandemia: la diffusione di aziende medio-piccole, organizzate in cluster territoriali (sovrapponendo mappe della localizzazione delle aziende e mappe di diffusione del COVID 19 la relazione è evidente) e la presenza di una aggregazione dopolavoristica, tipica di quelle zone.

I modelli di simulazione ad Agenti

Un modello basato su agenti (ABM) appartiene a una classe di modelli computazionali per simulare le azioni e le interazioni di agenti autonomi (entità individuali o collettive come organizzazioni o gruppi) al fine di valutarne gli effetti sul sistema nel suo insieme. [1] Combina elementi di teoria dei giochi, sistemi complessi, emergenze, sociologia computazionale, sistemi multi-agente e programmazione evolutiva. I metodi Monte Carlo sono usati per introdurre casualità. Gli ABM sono utilizzati su domini scientifici non informatici tra cui biologia, ecologia e scienze sociali. I modelli basati su agenti sono una sorta di modello in microscala [2] che simula le operazioni e le interazioni simultanee di più agenti nel tentativo di ricreare e prevedere la comparsa di fenomeni complessi.

La situazione attuale

L'infezione da Coronavirus si è diffusa in Italia presentando evidenti anomalie di distribuzione di contagiati nel territorio. Si evidenzia una concentrazione di casi straordinaria nelle regioni del Nord in particolare in Lombardia dove ad oggi (1/4/2020) da fonti ufficiali abbiamo 43.208 casi di contagio. Studi svolti in questi giorni [3] confermano come in realtà il numero di contagiati sia molto più grande e pari a circa il 9,8% della popolazione residente ovvero 5.9 milioni di individui. Questo riporta il tasso di letalità della malattia a quell'1.14% [4] in linea con considerazioni circa tipologia di popolazione e andamento generale del virus.

Obiettivi da raggiungere

L'obiettivo che ci prefiggiamo di raggiungere è determinare i motivi per cui abbiamo assistito a una così rapida e estesa diffusione del virus in Lombardia rispetto ad altre regioni [5]. In particolare riteniamo che esistano essenzialmente ragioni di tipo comportamentali e di struttura sociale. Per fare questo utilizzeremo un modello ABM, di seguito descritto, che verrà applicato alle diverse realtà territoriali.

La validità del modello così realizzato potrà essere confermato dai dati reali di diffusione del virus e lo stesso modello potrà essere utilizzato per generare previsioni di diffusione anche per le successive ondate dell'infezione, permettendo agli operatori di operare selettivamente sulle misure di distanziamento sociale.

Descrizione del nostro modello

Abbiamo preso in considerazione una popolazione formata da 5.000 agenti rappresentativi della società lombarda per fascia di anzianità, sesso e appartenenza a tre gruppi sociali distinti: "gruppo sociale famiglia", "gruppo sociale lavoro" e "gruppo sociale tempo libero". La distribuzione è stata normalizzata attingendo ai dati ISTAT [6] I gruppi vengono popolati dinamicamente ad ogni simulazione secondo parametri di composizione.

Nel gruppo sociale famiglia vengono simulate le famiglie secondo fasce di età e numero di componenti: vengono naturalmente simulate anche famiglie composte da un solo agente, avendo quindi coperto le casistiche sociali delle persone anziane sole e dei single. Nel gruppo sociale lavoro troviamo i lavoratori di singole aziende secondo la specifica per cui aumentando il numero di agenti all'interno del gruppo sociale lavoro diminuisce il numero di addetti quindi si va verso un modello di piccola e media impresa. Analogamente nel gruppo sociale tempo libero all'aumentare del numero di gruppi tempo libero diminuisce il numero di partecipanti (al limite se il gruppo tempo libero aumenta considerevolmente siamo nell'ipotesi che ogni agente compie attività in autonomia ie tutti fanno jogging, al contrario se diminuisce il numero di gruppi tempo libero vuol dire che un numero ampio di agenti si reca in palestra).

La simulazione avviene nell'arco delle 24 ore giornaliere in cui vengono simulate le interazioni sociali degli agenti e dei gruppi di agenti per un totale di 1200 ore pari a 50 giorni.

Abbiamo quindi descritto le modalità di diffusione della malattia attraverso interazioni sociali utilizzando i dati ad oggi disponibili [5] attribuendo per ogni agente in base allo stato in cui si trova (infetto, ammalato, deceduto o guarito) al gruppo di appartenenza (o nessun gruppo), alla fascia di età e ad un orario ben preciso, con che probabilità cambi di stato secondo la sequenza predeterminata utilizzata nel modello SIR [7].

La sequenza possibile è una matrice di probabilità così descritta:

Per ogni fascia di età cui appartengo	Allora			
Se sono	Mi ammalo	Contagio, per gruppo di appartenenza, per ogni fascia di età, in base all'orario	Guarisco	Decedo
Contagiato	P1	P2	P3	P4
Ammalato		P5	P6	P7
Guarito	ND	ND	ND	ND
Deceduto	ND	ND	ND	ND

Il modello è stato sviluppato utilizzando R: i file di input dei dati sono in formato csv mentre i parametri di composizione della popolazione sono inseriti nel codice.

Risultati dell'elaborazione

Abbiamo effettuato simulazioni utilizzando il nostro modello impostando diversi parametri che rappresentano i territori della Lombardia assumendo che vi siano alte concentrazioni di agenti anche nelle fasce di età alte in contesti lavorativi e in contesti di tempo libero con forte integrazione tra socialità professionale e tempo libero. I casi tipici sono le attività consulenziali post pensione, il fenomeno dei nonni che accudiscono i nipoti, l'altro numero di anziani negli ospedali per esami e in generale le frequentazioni di gruppo anche per meno anziani come palestre e svago ricreativo post lavoro.

Abbiamo quindi impostato due parametrizzazioni numeriche così descritte:

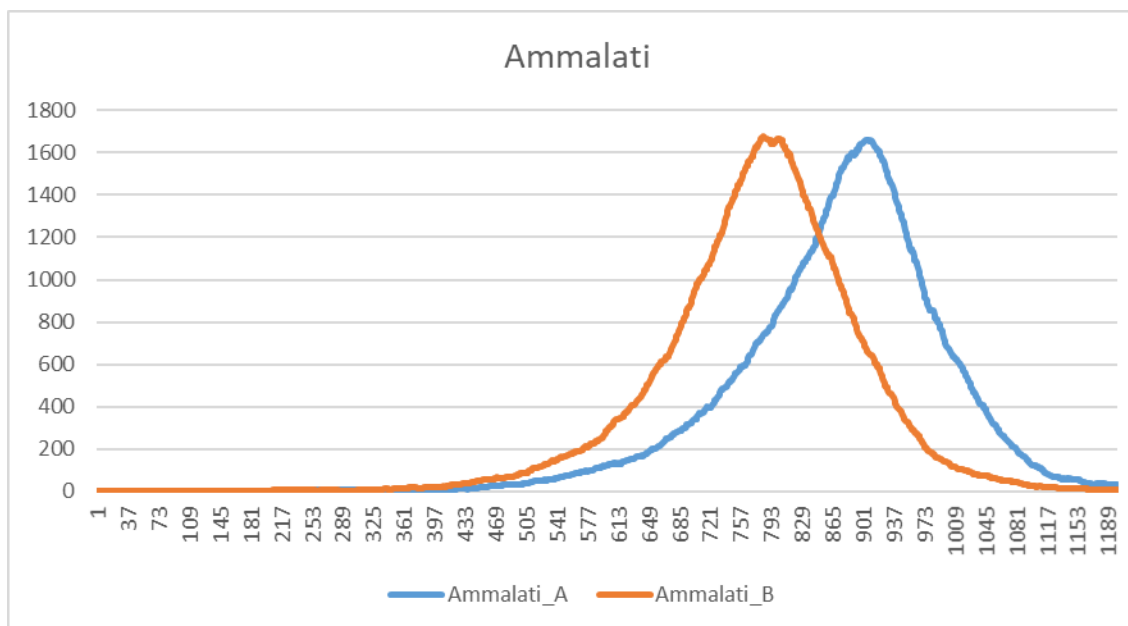
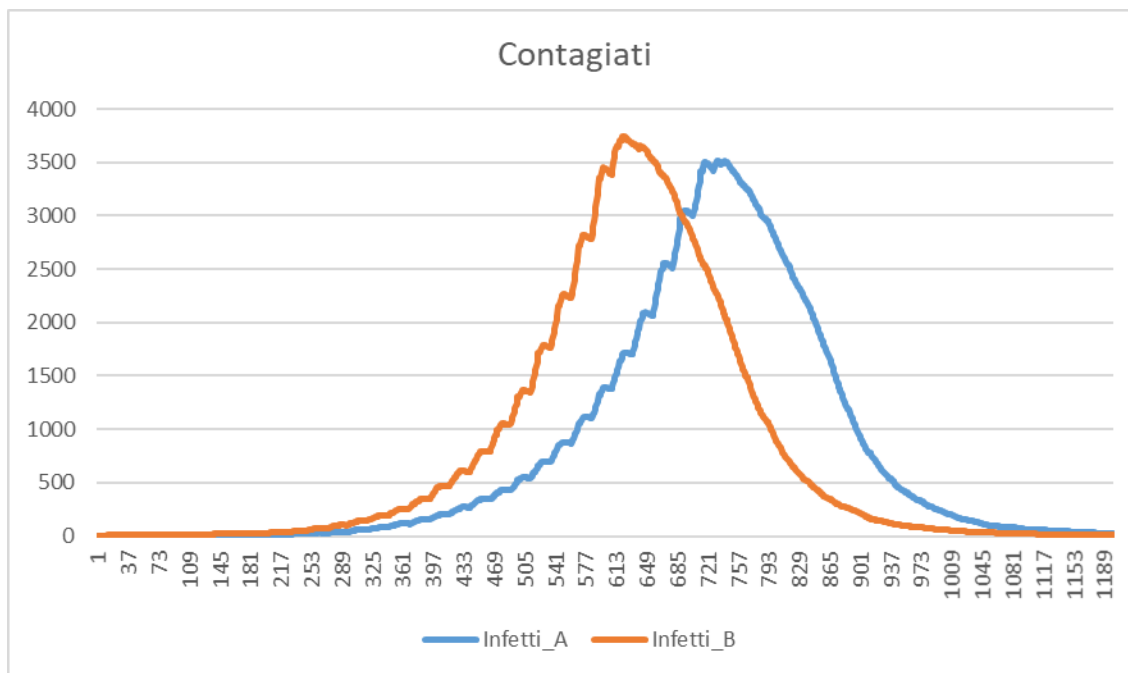
Soluzione A

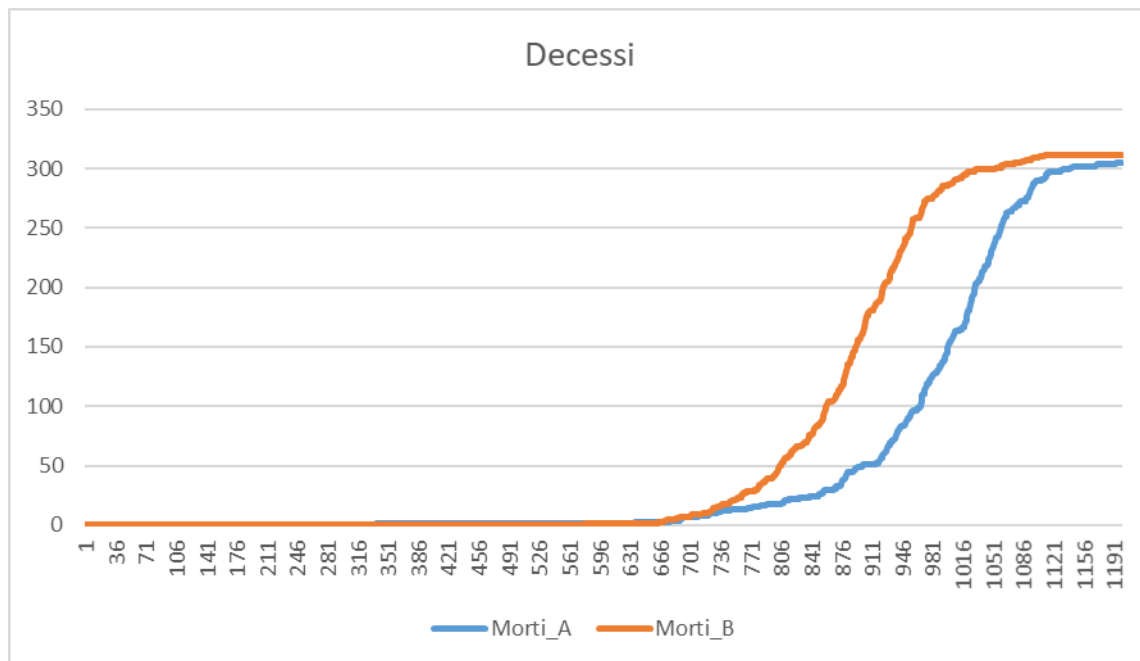
Poche aziende con alto numero di agenti per azienda e molti luoghi di aggregazione con un minor numero di agenti per luogo.

Soluzione B

Molte aziende con basso numero di addetti per azienda e pochi luoghi di aggregazione ma con maggior numero di agenti per luogo di aggregazione.

Le due simulazioni danno il seguente risultato:





Notiamo come la *Soluzione B* anticipa di molto la *Soluzione A* con un maggior numero di contagiati nel picco.

Conclusioni

Le due simulazioni evidenziano la rapidità di diffusione nell'arco di soli 50 giorni e quanto siano influenti i parametri della socialità sia lavorativa che di tempo libero sulla diffusione. Il modello sembrerebbe dimostrare che non sono necessarie ipotesi circa la natura del virus o le caratteristiche dell'ambiente per avere ripide curve e altissimi picchi di contagio. Semplicemente le caratteristiche socioeconomiche di una regione determinano comportamenti che favoriscono la diffusione della pandemia.

A maggior ragione quindi ci aspettiamo che le misure di distanziamento sociale, in assenza di vaccini preventivi o farmaci curativi, siano necessarie e sufficienti per fermare la diffusione del virus.

Prossimi approfondimenti

Procederemo ora con l'estensione del modello secondo il seguente schema:

- Estensione ad altre regioni simulando diversi comportamenti sociali
- Modellazione per fascia d'età per dettagliare i comportamenti sociali sulle popolazioni di anziani maggiormente colpite dalla pandemia
- Ampliare matrice delle probabilità considerando nuove casistiche, questo possibile solo quando saranno disponibili nuovi dati per poter determinare al meglio le probabilità.
- Simulazione distanziamento sociale, applicare al modello soluzioni di distanziamento sociale per verificare i cambiamenti nelle curve.

Bibliografia

- [1]. Niazi, Muaz; Hussain, Amir (2011). "Agent-based Computing from Multi-agent Systems to Agent-Based Models: A Visual Survey" (PDF). *Scientometrics*. 89 (2): 479–499. arXiv:1708.05872. doi:10.1007/s11192-011-0468-9. Archived from the original (PDF) on October 12, 2013.
- [2] Gustafsson, Leif; Sternad, Mikael (2010). "Consistent micro, macro, and state-based population modelling". *Mathematical Biosciences*. 225 (2): 94–107. doi:10.1016/j.mbs.2010.02.003. PMID 20171974.
- [3] 30 March 2020 Imperial College COVID-19 Response Team Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries.
- [4] ISPI ANALYSIS 27/03/2020 CORONAVIRUS: LA LETALITÀ IN ITALIA, TRA APPARENZA E REALTÀ Matteo Villa (ISPI Research Fellow)
- [5] Ministero della Salute www.salute.gov.it; Worldmeters.info <https://worldmeters.info>
- [6] ISTAT istituto Nazionale di Statistica www.istat.it
- [7] Hethcote H (2000). "The Mathematics of Infectious Diseases". *SIAM Review*. 42 (4): 599–653. Bibcode:2000 SIAMR..42..599H. doi:10.1137/s0036144500371907