Massively Parallel Simulations to Analyze Viral Infection Propagation Risk During Air Travel

Sirish Namilae
Embry Riddle Aeronautical University

Mathew Scotch, Anuj Mubayi and Rob Pahle
Arizona State University

Ashok Srinivasan
Florida State University

P Derjany, S. Chunduri, M Ghaffari, M Sadeghi
Motivation

- Infection transmission during air travel for many diseases
- There has been ban on flights from Ebola infected areas
  - Such measures early on can have large human and economic impact
  - Travelers with Ebola on passenger airplane in US
  - SARS transmitted during air travel
  - Evidence for many other disease transmission on airplanes
- Fine-tuned policy prescriptions for air-travel can be as effective
  - Reassures the public that action be being taken
  - Avoids negative human and economic impacts

Employ Molecular Dynamics and Monte Carlo based methods for pedestrian movement and infection transmission
Our Modeling Approach

- Use fine-scale model of human movement in planes to determine response to policies
- Link with phylogeographic model to examine global consequences
- Parameterize sources of uncertainty
  - Parameter sweep over this space to identify vulnerability
- Validate with similar diseases
Modeling Approach

- Pedestrian Location and Movement
- Pedestrian Dynamics
- Infectivity Profile
- Probabilistic Infection Profile
- Number of Contacts
- Susceptible-Infected Epidemic Model
Pedestrian Movement Models

- Different approaches to model pedestrian movement
  - Studies based on social psychology
  - Fluid flow based models for pedestrian movement
  - Geometric models to determine route through obstacles
  - Pedestrian movement based on cellular automaton
  - Social force model proposed by Helbing (2000)

- Social Force Model
  - Based on molecular dynamics and discrete element methods
  - Positions of people evolve in time based on interaction forces between other people and walls

\[
m_i \frac{dv_i}{dt} = m_i \frac{v_i^0(t) e_i^0(t) - v_i(t)}{\tau_i} + \sum_{j \neq i} f_{ij} + \sum_{W} f_{iw}
\]
Social Force Model

- Molecular dynamics is a mature simulation method in materials science and chemistry and serves as a framework for the model
  \[
  \vec{f}_i = \frac{m_i}{\tau} \left( \vec{v}_o^i(t) - \vec{v}^i(t) \right) + \sum_{j \neq i} \vec{f}_{ij}(t)
  \]

- We introduce location feedback
  - Location feed back on desired velocity
  - Effect of pedestrian density incorporated

\[
\vec{v}_o^i(t).\hat{e}_1 = (v_A + \gamma_i v_B) \left( 1 - \frac{\delta}{\vec{r}_i.\hat{e}_1 - \vec{r}_k.\hat{e}_1} \right)
\]

- Parallel computing for parameter reduction
Incorporating Line-Forming in Pedestrian Dynamics:

\[ \vec{f}^{\text{int}}_i = m_i \left( \frac{v_{0i}(t)-v_i(t)}{\tau} \right) \hat{e}_v \]

\[ \vec{v}_{0i}(t) = v_{0i}(t) \cdot \hat{e}_v = v_{0i}(t) \cdot \hat{e}_1 \]

\[ = (v_A + \gamma_i v_B)(1 - \frac{\delta}{||\vec{r}_i - \vec{r}_j||}) \cdot \hat{e}_1 \]

- \( v_A + \gamma_i v_B \): cumulative frequency distribution that varies between individuals.
- \( \gamma_i \) is a positive random variable less than unity.

- The repulsive force \( \vec{f}^{\text{ped}}_i \) ensuring impenetrability with other pedestrians and obstacles (Namilae, 2017):

\[ \vec{f}^{\text{ped}}_i = \sum_{i \neq 1} \vec{\vartheta} \phi(r_{il}) = \sum_{i \neq 1} \vec{\vartheta} \left[ \epsilon \left( \frac{\sigma}{r_{il}} \right)^{12} \right] \]
Crack propagation and dislocation emission during Liquid metal embrittlement in Al-Ga system – Namilae (2008)

Nanotube composite interface molecular dynamics model - Namilae (2006)
Parameter Estimation

First pass
Parameter estimation
On Bluewaters
Total simulations ~300,000

\[
\bar{v}_o^i(t) \dot{e}_i = (v_A + \gamma_i v_B) \left( 1 - \frac{\delta}{r_i^e - r_k^e} \right)
\]

\[
\bar{f}_i = \frac{m_i}{\tau} (\bar{v}_o^i(t) - \bar{v}_i^i(t)) + \sum_{j \neq i} \bar{f}_{ij}(t)
\]

\[
\bar{f}_{ij} = \frac{d}{dr_{ij}} \left( \sigma \left( \frac{\epsilon}{r_{ij}} \right)^{12} \right)
\]
Second pass
Parameter estimation
Total simulations ~300,000


Physical aspects like front to back deplaning considered for validating parameters
Model Validation

Different parameter combinations predict observed data for 5 airplanes

(Namilae et al., Physica A 2017)
Pedestrian Simulation of Embarkation & Deplaning:

Simulation snapshots of Airbus A320 deplaning at different time steps.
Pedestrian Simulation of Embarkation & Deplaning:

Simulation snapshots of an embarkation of an Airbus A320 from a departure lounge at different time steps.
A320 144 Seats Egress
A320 144 Seats INGRESS
Stochastic Infection Spread Model:

- The study of epidemics informs how a disease propagates and what are the suitable policies to suppress or inhibit its spread.

Assumptions:
- A population of size N consisting of I(t) infected and S(t) susceptibles: \( N = I(t) + S(t) \)
- A susceptible becomes infected when coming into direct contact with an infected. No reproduction is allowed within a short time range.
- “m” the total number of contacts per individual per time step.

- The probability that an infective at \( c \) days of infection meets other individuals is \( m/N \).
- Denote by \( P_c \) the probability that a contact between a susceptible and the infective, results in infection of the susceptible.
- \( P(\text{contact and infection}) = P(\text{infection/contact}) \cdot P(\text{contact}) = P_c \cdot \frac{m}{N} \)

Susceptibles  
\[ S \]  
Infected  
\[ I \]
Infection Transmission Model

i → each individual will consists of \([x, y, k]\) components at each time where,

\[(x, y) \rightarrow \text{position of the individual } i \text{ at time } t\]

\[k \rightarrow \text{represents the individual’s infection status } \varepsilon \{S, I\} \text{ at time } t\]

\[P_{ij} \rightarrow \text{probability of susceptible individual } i \text{ to receive infection successfully from infectious individual } j\]

\[P_{ij} = \begin{cases} f(r, y, \tau) & \text{if } i \text{ is in vicinity of } j \text{ whose } \text{"infectious radius" is } r, \\ & \text{"infectivity level" is } y, \\ & \text{"exposing for" } \tau \text{ time units} \\ 0 & \text{otherwise} \end{cases}\]
Infection Transmission Model

• Given the pedestrian trajectories from pedestrian model- we obtain contacts between people.

  \[ P(\text{contact and infection}) = P(\text{infection/contact}) \cdot P(\text{contact}) = P_c \cdot \frac{m}{N} \]

• Number of susceptible

  \[ S(t) = N - \sum_{c=1}^{d} i_c^0 = N - I(t) \]

• The number infected is binomially distributed (for demographic stochasticity) with parameters
  
  - \( n = S(t-1) \), the number of susceptibles at time \( t \), and \( p = P_c \cdot \frac{m}{N} \)

• Approximating the Binomial as Poisson. Number infected at time \( t \)

  \[ I(t) \sim \text{Poisson} \left( \sum_{c=1}^{d} \left( \sum_{i=1}^{i_c^0} \left( \frac{m_i(t-1)s_i(t-1)}{N} \right) \right) \right) \]

  • Location of Infected person is unknown and varied.
  • Parametric variations to quantify uncertainty and risk
Virus content in blood is used to create infectivity probability plot

Combined with dose response models

Difference between diseases like SARS and Ebola primarily dependent on contact definition (e.g. distance, time etc)
Since $R_0$ for Ebola is around 2, that means a typical infective individual will produce on an average two new secondary cases thus, replacing him or herself, producing additional case, and eventually leading to large outbreak in the population.
Results – Boarding Strategies

- Similar pattern for 144 seat A320 seating configuration & 182 seat Boeing 757

- There is clear difference between different boarding strategies.

- Strategies that lead to arbitrary movement along the cabin preventing clustering reduces infection transmission.


Boeing 757-200 for different boarding patterns for a single imported infective at the $9^{th}$ day of infection (probability of infection = 0.06), critical radius of infection 1.2 m.
Boeing 757-200 for different exiting patterns for a single imported infective at the 9th day of infection (probability of infection = 0.06), critical non-successive contact number of 3 and critical radius of infection 48 feet. The contacts of egressing passengers outside the airplane are not taken into account.
Boeing 757-200 combined baseline egress and different boarding patterns for a single imported infective at the 9th day of infection (probability of infection = 0.06), critical radius of infection 48 in.
Parameter variation – Infection radius

- Critical model parameter - contact radius defined as minimum distance to define a contact
- Particles travel depending on (1) size and (2) fluid mechanics in expiratory events (e.g. coughing)
- Particle size 0.1 to 10 micrometers. Distance travelled up to 2m
- Mechanism of infection –
  - long distance - transmitted by small particles like aerosols (SARS, H1N1)
  - Short distance – transmitted by coarse droplets e.g. Ebola.

Boeing 757-200 combined for a single imported infective at the 9th day of infection (probability of infection = 0.06), different critical radii of infection.
Long vs short contact radius SARS vs Ebola

- Infection radius for Ebola 1.2 m and for SARS 2.1 m
- SARS more contacts and infection. SARS was transmitted on airplanes *
- Model includes airport gate.

Parameter Variation and Uncertainty Quantification

Long vs short contact radius, SARS & H1N1 vs Ebola during random ingress

- At short transmission distance, large airplanes, above a certain seating threshold produces less infections to slightly smaller ones since passengers are dispersed over a wider range.
- At peak days of infectivity, Ebola has a slight shift due to its low infectivity ($\lesssim 10\%$)
Parameter Variation and Uncertainty Quantification

- At large transmission distance, large airplanes produce more infections since a greater number of passengers are within the critical radius of infection.
- At peak days of infectivity, SARS still has a slight shift due to its low infectivity ($\lesssim 10\%$)
Airplane Size

Number of contacts for transporting 1000 passengers in different airplanes boarding and deplaning by default methods.

- Smaller airplanes result in lower number of
- The smaller size of the susceptible population, lower number of susceptibles within a given contact radius and the reduced time of in-plane movement are some of the factors that benefit smaller airplanes.

Infection distribution profile for random boarding strategy varying the airplane size.
Policy decision makers meet in a decision meeting to explore policies for containing the disease spread.

**Real-time response** to policy queries is required.

SPED has limitations for real-time response:
- Even a parallel sweep is limited by the speed of slowest simulation (~20 minutes).
- The computational bottleneck becomes severe with increasing number of parameters.

**Solution:** separate the computation of trajectories (SPED simulation) and the analysis of contacts (based on these trajectories):
- **Pre-compute the trajectories**
- Analysis phase can be parallelized **for real-time response**.

In this paper, we discuss the optimizations we made in parallelizing the Analysis phase.
Types of Parameter sweep

- **2D Lattice**
  - Parameter space coverage: **inefficient**
  - Convergence check: **inefficient**
  - Factor 2\(^d\) gap between convergence checks

- **2D Random**
  - Parameter space coverage: **inefficient**
  - Convergence check: **efficient**
  - Factor 2 gap between convergence checks

- **2D LDS**
  - Parameter space coverage: **efficient**
  - Convergence check: **efficient**
  - Factor 2 gap between convergence checks

SPED model in this study uses 5 parameters
- 5-D Lattice based and 5-D Scrambled Halton Low Discrepancy Sequence (LDS) based parameter sweeps used for infection spread analysis
Parameter sweep Using LDS

Parameter space coverage:

- **inefficient**
  \[ \text{Discrepancy} \propto \frac{\log \log N}{N^{1/2}} \]

- **efficient**
  \[ \text{Discrepancy} \propto \frac{\log^d N}{N} \]

SPED model in this study uses 5 parameters

- 5-D Lattice based and 5-D Scrambled Halton Low Discrepancy Sequence (LDS) based parameter sweeps used for infection spread analysis
Convergence for lattice sweep

Histogram for subgrid of size $5^5$

Histogram for subgrid of size $9^5$

Histogram for $17^5$ grid
Convergence for LDS sweep

Histogram for $17^5$ grid

Histogram for subgrid of size $9^5$

Histogram for subgrid of size $5^5$

Histogram for $17^5$ points

Histogram for $2^{18}$ points

Histogram for $32768$ ($2^{15}$) points
Convergence algorithm

Algorithm 1: Histogram Convergence Scheme

procedure
  Initialization
  for each process do
    Analyze $N_0$ files
    Update local moments
    Compute global statistics
  end for
  $i = 0$
  while not converged and files remain unanalyzed do
    $i++$
    Analyze $(2^i - 2^{i-1})N_0$ files
    Update local moments
    Compute global statistics
    Check for convergence
  end while
end procedure

LDS: Relative difference in statistics used in convergence check
Load imbalance in lattice vs. LDS sweeps

Load imbalance across processes is defined as

\[ \frac{\text{MaximumLoad} - \text{AverageLoad}}{\text{AverageLoad}} \]

0 when load is perfectly balanced

- Lattice sweep is well balanced
- LDS has a poor balance with 1000 and 1024 processes
- LDS performs better than Lattice for 1003 processes
  - 1003 is divisible by 17 (parameter values)

Load imbalance for Lattice and LDS sweep of the entire data set $17^5$ (without convergence checks) using cyclic distribution

1000 and 1024 are products of primes used in the LDS

*IEEE CCgrid (2018)*
Load balancing LDS

- Cyclic: Load is not well balanced in the initial stages even with 1003 processes
- Block: Does not work well for small number of samples
- Dynamic: Master-worker based dynamic load balancing works best overall
Scaling of dynamic load balancing of LDS

- Master-worker approach has the overhead of communication with master
- Good weak scaling efficiency
- Strong scaling results show good efficiency until around 4000 processes, with a sharp drop-off after that

Weak scaling of LDS with dynamic load balancing

Strong scaling of LDS with dynamic load balancing
Phase Maps for policy analysis

Day 1 of infection

- At day 1, Ebola with a lower critical radius compensated by a higher infectivity, produces the same number of infected compared to H1N1.

Peak day of infection

- SARS combining the higher infectivity and radius of infection produces greater number of infected travelers.
- At peak day of infection H1N1 disperses more among the passengers especially on stretched-aircraft configuration.
Summary

- Pedestrian movement model combined with stochastic Susceptible–Infected model for Ebola and SARS

- Airplane movement patterns evaluated for disease propagation
  - Boarding in 2 sections with randomized within sections is the most effective strategy to reduce infections
  - Boarding has higher effect than deplaning.
  - Smaller airplanes are better

- Parameter sweep with LDS is more efficient than with Lattice

- Better coverage of parameter sweep and faster convergence
  - It made feasible analysis that was not feasible earlier

- Load imbalance is a potential problem with LDS and is related to its number-theoretic properties
  - Identified techniques, that can lead to good load balancing under different applications scenarios

- Acknowledgements: NSF PRAC and Rapid grants
Extensive News coverage of our Research

• Covered in over 75 news outlets in four continents

Publications

- S. Namilae, P Derjany, A Mubayi, M Scotch and A Srinivasan, Multiscale Model For Infection Dynamics During Air Travel, Physical review E, 002300 (2017)
- P Derjany, S Namilae, A Mubayi and A Srinivasan, Computational Model for Pedestrian Movement and Infectious Diseases Spread During Air Travel, AIAA Scitech (2018)
Results
Boarding and Deplaning strategies

Number of contacts for different deplaning strategies in 144 seat Airbus A320 and 182 seat Boeing 757-200 seating configurations.

Number of contacts for different boarding strategies in 144 seat Airbus A320 and 182 seat Boeing 757-200 seating configurations. The bars represent standard deviation.
Number of human-human contacts during boarding for the five airplanes during deplaning for contact threshold of 18 inches and 30 inches. The bars represent standard deviation.

Percentage of first class seats vs contacts during boarding and deplaning for contact threshold of 18 inches.