Dynamic Random Graph Modelling and Applications in the UK 2001 Foot-and-Mouth Epidemic

Christopher G. Small

Joint work with

Yasaman Hosseinkashi, Shoja Chenouri

(Thanks to Rob Deardon for supplying the data.)
Outline

- Foot-and-mouth disease outbreak in the UK, 2001 and dynamic random graphs (DRG)
- Introducing a Markov model for DRGs
- Inferring the missing edges and some epidemiological factors
- Application to the UK 2001 foot-and-mouth (FMD) epidemic
- Simulation results and detecting the high risk farms
Foot-and-mouth outbreak, UK 2001

At day 4

All farms involved in the epidemic

Farm known to be infectious at the current day

Edges show possible disease pathways

At day 5

New infection discovered
Foot-and-mouth outbreak, UK 2001

At day 44

At day 184
Foot-and-mouth outbreak, UK 2001

Cumulative number of infectious farms over time

Changes in the diameter of the infectious network over time
Infectious disease outbreak as a dynamic random graph

The development of an infectious disease outbreak can be modeled as a sequence of graphs with the following vertex and edge sets:

- Vertices: individuals (patients, susceptible).
- Edges: disease pathways (directed).

\[
G = \{G(t) = (V(t), E(t)) : t \in T\}
\]
A Markov model for dynamic random graphs

Model assumptions:

• Every infectious farm $i$ may infect a susceptible farm $j$ with exponential waiting time $Y_{ij} \sim \exp(h_{ij})$.

• Simultaneous transmissions occur with very small probability.

• Transmission hazard $h_{ij}$ can be parameterized as a function of local characteristics of both farms and their Euclidean distance.

\[ G = \{ G(t) = (V(t), E(t)) : t \in T \} \]

is a sequence of random directed graphs over a continues time domain $T$ where

\[ V = V_{inf}(t) \cup V_{sus}(t) \cup V_{rem}(t) \quad \text{for all } t \in T \]
A Markov model for dynamic random graphs

Embedded in discrete time domain:

\[ \{(T_k, G_k) \mid k : 1, 2, \cdots, K\} \]

- Waiting time to get into k’th state
- All farms with their active transmission pathways

- The composition of infectious/susceptible/removed farms changes over time by new infections and culling the previously infected farms.

- The k’th transition occurs when the k’th farm is infected.

- Culling modeled as deterministic process.
A Markov model for dynamic random graphs

At state 3:

\[ V_3^{\text{inf}} = \{4, 2, 3, 1\} \]
\[ V_3^{\text{sus}} = \{5, 6\} \]
\[ V_3^{\text{rem}} = \{7\} \]

At state 4:

\[ V_4^{\text{inf}} = \{4, 2, 3, 1, 5\} \]
\[ V_4^{\text{sus}} = \{6\} \]
\[ V_3^{\text{rem}} = \{7\} \]

Edges are usually missing in the actual epidemic data!
A Markov model for dynamic random graphs

Transition probabilities and likelihood:

\[
f ( T_{k+1}, G_{k+1} \mid T_k, G_k ) = P(G_{k+1} \mid G_k) \times f(T_{k+1} \mid G_k)
\]

\[
P ( G_{k+1} \mid G_k ) = \frac{\sum_{i \in V_{k+1}^{\text{inf}}} h_{ij_{k+1}}}{\sum_{j \in V_{k}^{\text{sus}}} \sum_{i \in V_{k}^{\text{inf}}} h_{ij}}
\]

\[
V_{k+1}^{\text{inf}} = V_k^{\text{inf}} \cup \{ j_k \}
\]

\[
f (T_{k+1} \mid G_k) = \left( \sum_{j \in V_{k}^{\text{sus}}} \sum_{i \in V_{k}^{\text{inf}}} h_{ij} \right) \exp \left\{ -T_{k+1} \times \sum_{j \in V_{k}^{\text{sus}}} \sum_{i \in V_{k}^{\text{inf}}} h_{ij} \right\}
\]

\[
L(\theta) = \prod_{k=1}^{K-1} P(T_{k+1}, G_{k+1} \mid T_k, G_k)
\]
A Markov model for dynamic random graphs

Transition probabilities and likelihood:

\[ \ell(\theta) = \sum_{k=1}^{K-1} \left[ \log \left( \sum_{i \in V_k^{\text{inf}}} h_{ij_k} \right) - T_{k+1} \sum_{j \in V_k^{\text{sus}}} \sum_{i \in V_k^{\text{inf}}} h_{ij} \right] \]

Model parameters:

\[ \theta = (\alpha, \beta, \gamma, \delta) \]

Density of livestock

Euclidean distance

\[ h_{ij} = n_i^\gamma \exp \left\{-\alpha d_{ij}^\beta \right\} n_j^\delta \]
Inference on missing edges and epidemiological factors

Probability distribution for the $k$'th transmission pathway:

$$p_k(i) = P \left( e_{ij}(k) = 1 \mid V_k^{\inf}, V_k^{\sus}, V_{k+1}^{\inf} = V_k^{\inf} \cup \{ j_k \} \right)$$

$$= \begin{cases} \frac{h_{ij_k}}{\sum_{l \in V_k^{\inf}} h_{lj_k}} & i \in V_k^{\inf} \\ 0, & i \notin V_k^{\inf} \end{cases}$$

$V_k^{\sus}, V_k^{\inf}$: Infectious and susceptible farms at state $k$

$j_k$: The farm which is infected at $k$'th transition
The basic reproduction number ($R_0$) is defined as the expected number of secondary cases produced by a typical infected individual during its entire infectious period, in a population consisting of susceptibles only (Heesterbeek & Dietz, 1996):

\[ R_0 = S_0 \int_0^\infty A(\tau) d\tau \]

The density of population at the start of the epidemic when every individual is susceptible.

Expected infectivity of an individual with infection age $\tau$. 
Inference on missing edges and epidemiological factors

\[ R_0 = S_0 \int_0^\infty A(\tau) d\tau \quad \Rightarrow \quad \text{Assuming homogenous infectivity over the population!} \]

- The basic reproduction number can be estimated by the expected out degree of a vertex, summed over all transitions (which is equivalent to sum over its infectious period).

\[ \hat{r}_i = \sum_j \hat{E}(e_{ij}) = \sum_k \hat{p}_k(i) \quad \text{Expected outdegree of the vertex } i \text{ during the epidemic, which can be considered as the specific basic reproductive number for vertex } i. \]

\[ \hat{R}_0 = \frac{1}{|V|} \sum_i \hat{r}_i \]
Inference on missing edges and epidemiological factors

Example:

Transition one

\[ p_1(4) \]

Transition two

\[ p_2(4) \]
\[ p_2(2) \]

Transition three

\[ p_3(4) \]
\[ p_3(2) \]
\[ p_3(3) \]

Transition four

\[ p_4(4) \]
\[ p_4(2) \]
\[ p_4(1) \]
\[ p_4(3) \]

\[ r_4 = p_1(4) + p_2(4) + p_3(4) + e_4(4) \]
\[ r_2 = p_2(2) + p_3(2) + p_4(2) \]
Inference on missing edges and epidemiological factors

Considering all potential disease pathways at each transition:

- The cumulative risk that a farm has tolerated before becoming infected is summarized by the expected potential indegree over the time interval that it belongs to the susceptible subgraph.

- The cumulative threat that a farm has caused is computed by its expected potential outdegree over the time interval that it belongs to the infectious subgraph.

\[
\begin{align*}
V_3^{\text{inf}} &= \{4, 2, 3, 1\} \\
V_3^{\text{sus}} &= \{5, 6\} \\
V_3^{\text{rem}} &= \{7\}
\end{align*}
\]
The data

All farms who involved the epidemic according to our data:

• 2026 farms, 235 days
• Information available for each farm includes:
  - Report date, Infectious data, Cull date
  - Amount and density of livestock (sheep and cattle)
  - Region (shown by colors in this plot)
  - X – Y coordinates
  - Diagnosis type
• The data has been polished by removing farms with missing information.

A very intensive region
Normal test statistics
(H0: parameter = 0)

\[ \ell(\theta) = \sum_{k=1}^{K-1} \left[ \log \left( \sum_{i \in V_k^{\text{inf}}} h_{ij_k} \right) - T_{k+1} \sum_{j \in V_k^{\text{sus}}} \sum_{i \in V_k^{\text{inf}}} h_{ij} \right] \]

<table>
<thead>
<tr>
<th>hazard</th>
<th>parameters</th>
<th>ML estimates</th>
<th>Normal test statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>( h_{ij} = \exp{-\alpha d_{ij}^\beta} )</td>
<td>( \alpha )</td>
<td>12.5873576</td>
<td>38.4376552</td>
</tr>
<tr>
<td></td>
<td>( \beta )</td>
<td>0.2141929</td>
<td>22.2245189</td>
</tr>
<tr>
<td>( h_{ij} = n_i^\gamma \exp{-\alpha d_{ij}^\beta} n_j^\delta )</td>
<td>( \alpha )</td>
<td>12.01374599</td>
<td>31.208301</td>
</tr>
<tr>
<td></td>
<td>( \beta )</td>
<td>0.18493832</td>
<td>19.552151</td>
</tr>
<tr>
<td></td>
<td>( \gamma )</td>
<td>0.10886620</td>
<td>-1.359490</td>
</tr>
<tr>
<td></td>
<td>( \delta )</td>
<td>0.04010013</td>
<td>-1.275928</td>
</tr>
</tbody>
</table>
Simulation and effects of control policies

- Reducing the infectious period for each farm
- Decreasing the number of infectious (active) farms at each time point
Simulation results

After 50 transmissions
After 1000 transmissions
After 1500 transmissions
Simulation results

Cumulative number of infectious farms

Date (day)
How infectious and resistance each farm was during the epidemic?

- All in region 2!
- Density of livestock is 78.57 on average with 47.6 standard deviation.

- Mostly in regions: 2, 6, 4, and 1
- Density of livestock is 137.30 on average with 113.9 standard deviation.

Highly infectious farms

Highly resistant farms
Comparing the two groups

Highly infectious farms

Highly resistant farms

Epidemic data

The kernel density estimate of sheep ratio in the livestock
Thanks...