

The SIS/SEIRS dynamic network epidemic models for individual risk prediction

Konstantin Vorontsov¹, Anastasia Ianina¹, Vasiliy Alekseev¹,
Andrey Stepanov², Alexander Kozhevnikov², Alexander Mitroshkin²

¹MIPT, ²Softtree LLC

April 26, 2020

We propose a new algorithm RiskPropNet that models how infection risks propagate through the contact network. Our risk-propagation algorithm combines SIS/SEIR (susceptible-exposed-infected-recovered) epidemic models (fig. 1) and machine learning approaches in order to predict the probability of infection for individual users.

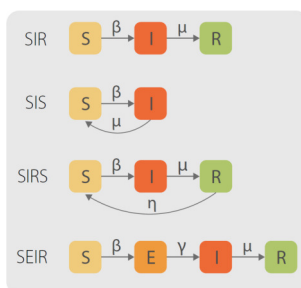


Figure 1: Diagrammatic representation of epidemic models in terms of reaction-diffusion processes. Boxes stand for different user states (compartments), while the arrows represent transitions between compartments, happening stochastically according to their respective rates [1].

The algorithm RiskPropNet processes in real-time a sequence of records of two types. *Contact record* $\langle t, u, v \rangle$ means that users u and v contacted at time t . *State record* $\langle t, x, y \rangle$ means that user x has state y at time t . The data structure above can be generated by ContactTracer mobile application. The states y are called *compartments* in epidemic models. Two basic states used in SIS model [1] are S (susceptible) and I (infected).

When the contact happens, RiskPropNet estimates the probability of infection transmission and updates the infection probability for both users who have been in contact.

When the state of the user changes, RiskPropNet updates the probabilities of infection for all users who have contacted him or her in the recent past. This update spawns a fading wave of probability updates that propagates throughout the contact graph. Also RiskPropNet updates the parameters of two predictive models: the transmission probability and the infection probability. Note that all the above updates may require intensive computations.

To the best of our knowledge, the individual-based epidemic models with learnable parameters for dynamic networks have not previously been proposed in the literature.

Further we describe the RiskPropNet algorithm in more detail.

First you need to agree on the representation of time in our model. We assume that time is discrete with a tick equal to 1 for all model parameters that are computed and stored over time. Nevertheless, we will label sequential data records by continuous time in order to preserve their proper ordering. We denote by \bar{t} the discrete time interval for continuous time label t .

Infection probability model. The change over time of the probability of infection $\rho_t(x)$ of user x equals the average of two competing variables: (a) if x is infected, then the probability decreases with rate equal to the *curing rate* μ and (b) if x is healthy, it can be infected with *infection rate* β from each infected neighbor [1]:

$$\rho_{t+1}(x) = (1 - \mu)\rho_t(x) + \beta(1 - \rho_t(x)) \sum_{\langle s < t, x, v \rangle} a_{st}(x, v)\rho_s(v), \quad (1)$$

where $a_{st}(x, v)$ is asymmetric adjacency matrix, whose elements are probabilities that the infection is transmitted from v to x .

Transmission probability model The elements of the adjacency matrix may depend on many factors including the time difference $t - s$, contact duration, a distance between v, x locations, a distance between beacon lists in these locations, etc. Denote $f_j(x, v)$, $j = 1, \dots, n$ the proximity features whose values are greater, the closer the contact and the higher the probability of transmission. Understanding that it is a combination of these features that increases the probability, we propose a multiplicative linear model with learnable parameters $\alpha_j \geq 0$:

$$a_{st}(x, v) = \exp\left(\sum_{j=1}^n \alpha_j f_j(x, v)\right).$$

Contact record processing. The contact record $\langle t, u, v \rangle$ adds one term to the sum of the infection probability models for both u and v :

$$\begin{aligned} \rho_{\bar{t}+1}(u) &+= \beta(1 - \rho_{\bar{t}}(u))a_{tt}(u, v)\rho_{\bar{t}}(v); \\ \rho_{\bar{t}+1}(v) &+= \beta(1 - \rho_{\bar{t}}(v))a_{tt}(v, u)\rho_{\bar{t}}(u). \end{aligned}$$

State record processing. The state record $\langle t, x, y \rangle$ with known $y \in \{I, S\}$ gives us an accurate value of probability $\rho_t(x) = [y=I]$ at time t . However, we cannot reliably estimate the probabilities $\rho_i(x)$ for $i < t$. The deeper into the past, the greater the probability $\rho_i(x)$ may deviate from the known binary value $\rho_t(x)$. We propose to modify the previously calculated probabilities by a mixture of probabilities with forgetting rate λ :

$$\rho_i(x) := (1 - \lambda^{i-t})\rho_i(x) + \lambda^{i-t}\rho_t(x), \quad i = 1, \dots, t.$$

This modification will affect the infection probabilities for all users u who have contacted user x in the past. Probabilities $\rho_i(u)$ must be recalculated according to (1) for all $i < t$ such that a record $\langle i, x, u \rangle$ exists. The recalculation of probabilities for the set of users u will affect the infection probabilities of even more users.

Thus, the iterative recalculation of probabilities can reveal the ways of possible spread of the infection, which could have occurred in the past before the user x found out about his infection.

The state records $\langle t, x, y \rangle$ give us another great opportunity to apply log-likelihood maximization in order to determine the parameters of a probabilistic model:

$$\sum_{i=1}^t \lambda^{i-t} \left([y = I] \ln \rho_i(u, \omega) + [y = S] \ln(1 - \rho_i(u, \omega)) \right) \rightarrow \max_{\omega},$$

where ω is a vector of parameters for both infection probability model and transmission probability model.

An optimization procedure can be used for this problem similar to the stochastic gradient for machine learning models, such as recurrent neural networks. Each iteration of the procedure is one pass through a sequence of contact/state records. State records induce recalculation loops for retrospective contact tracing in cases where new information appears, for example, from infection test results.

Conclusion. We proposed the RiskPropNet algorithm that learns the infection probability model for individual risk prediction. The RiskPropNet is a symbiosis of two algorithms that work in one common loop. The first one is a data-driven dynamic network epidemic model in which new vertices and connections can appear. The second one is a gradient-based optimization technique that opens up a possibility of using learnable models with high-dimensional parameter space.

References

- [1] *Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani.* Epidemic processes in complex networks. 2014.
- [2] *Shi Pengpeng, Cao Shengli, and Feng Peihua.* SEIR transmission dynamics model of 2019 nCoV coronavirus with considering the weak infectious ability and changes in latency duration. 2020.