① Transmission Tree Evolution On Contact Networks

(with D. Welch, Auckland Comp Sci)

DOI: 10.1016/j.jtbi.2016.07.038

preprint http://goo.gl/fVPvzH

Outline.

1. Transmission Process as a Markov Chain on $\mathcal{C}_n \times \mathcal{T}_n$
2. 3 examples.
3. Beta Splitting Model
4. 3 examples again
5. Equiv classes in $\mathcal{C}_0$ of $\mathcal{B}^{\uparrow}$
6. Some pictures. [MLE fig for other models]

$$c=(\omega,s)\in\mathcal{C}_n := \boxed{2^{\overset{\text{weighted edge set of complete weighted directed graph (or network) } K_n}{\mathcal{W}_n}}} \times \underbrace{\{0,1\}}_{\text{status tag set}}^{\Pi_n}$$

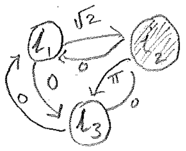
power set of $\mathcal{W}_n$

status tag set. 0 = Susceptible, 1 = Infected

Π_n → vertex set $\{\ell_1, \ell_2, \ldots, \ell_n\}$ of individual "hosts" in a pop. of size n.

"SI-tagged contact Networks" (SICN)

An example element of $\mathcal{C}_3$



▨ 1

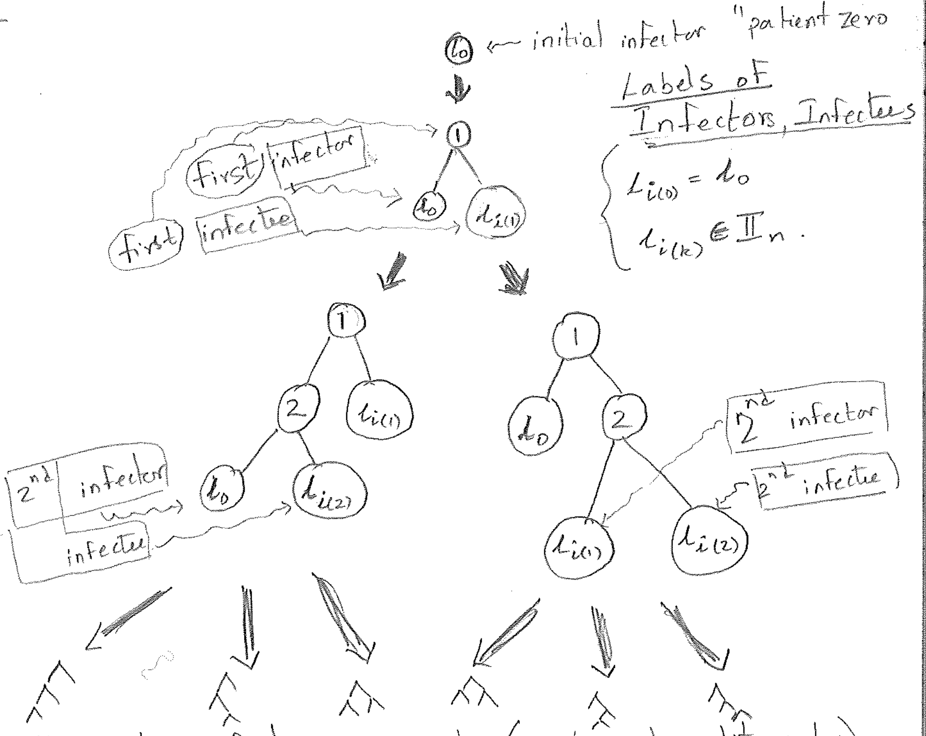
☐ 0

Note Hasse diagram on $2^{\mathcal{W}_n}$ with unit weights, given by posets under $\preceq \equiv \subseteq$

$$\tau \in \mathcal{T}_n = \{\text{rooted planar } \underset{\text{infection order}}{\underline{\text{ranked}}}\ \underset{\{\text{infected indivs.}\}\subseteq \Pi_n}{\underline{\text{leaf-labelled}}} \text{ binary trees}\}$$

"Transmission Trees"

Hasse-diagram on $\mathcal{T}_n$. (with generic leaf labels)



So with k infection events (internal split nodes) there are k! transmission trees.

note [Flajolet & Sedgewick, Anal Comb, "increasing tree-labelling" $\xrightarrow[\text{onto}]{1-1}$ permutations of $\{1,2,\ldots,k\}$].

So Our Transmission Proc. M.C. (jump chain) has as its state at current discrete time z:

$$(\tau(z), c(z)) \in \mathcal{T}_n \times \mathcal{C}_n$$

Transition Probs. of the jump M.C.

$$\Pr\{(\tau(z+1), c(z+1)) \mid (\tau(z), c(z))\}$$

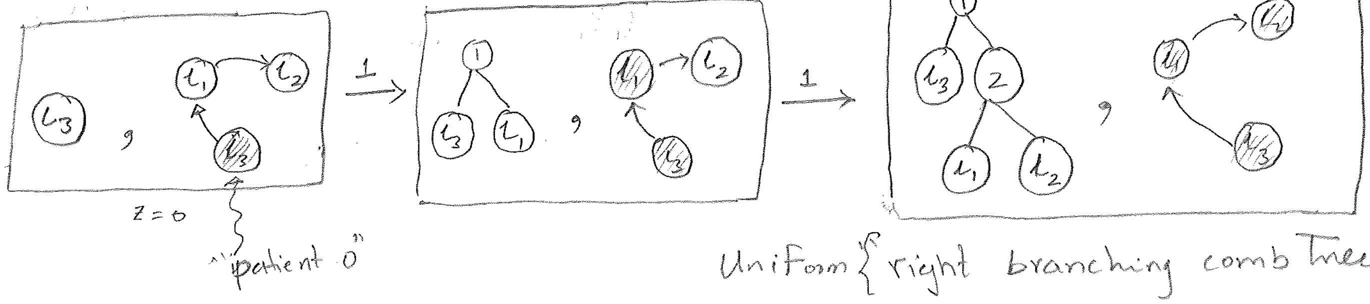
$$= \frac{\text{the weight of edge from } (z+1)\text{-th infector to the } (z+1)\text{-th infectee}}{\text{Sum of edge weights from every potential infector to every potential infectee within its susceptible out neighborhood.}} \quad (*)$$

[letting time for each infection event $\overset{iid}{\sim}$ Exponential(λ) we can get the C.T.M.C's generators as usual — ignored here].

Note. We limit to only connected networks in 2^{ω_n} & typically have unit weights (to focus on discrete structures) and undirected (bidirected) edges.

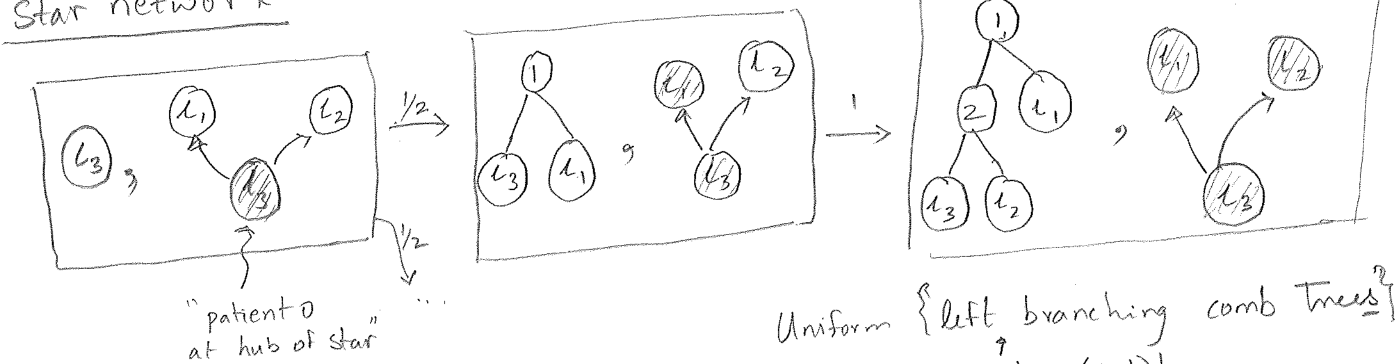
$\boxed{2}$ Three Example Networks. [special cases of $(*)$]

(i) Path network $(n=3)$



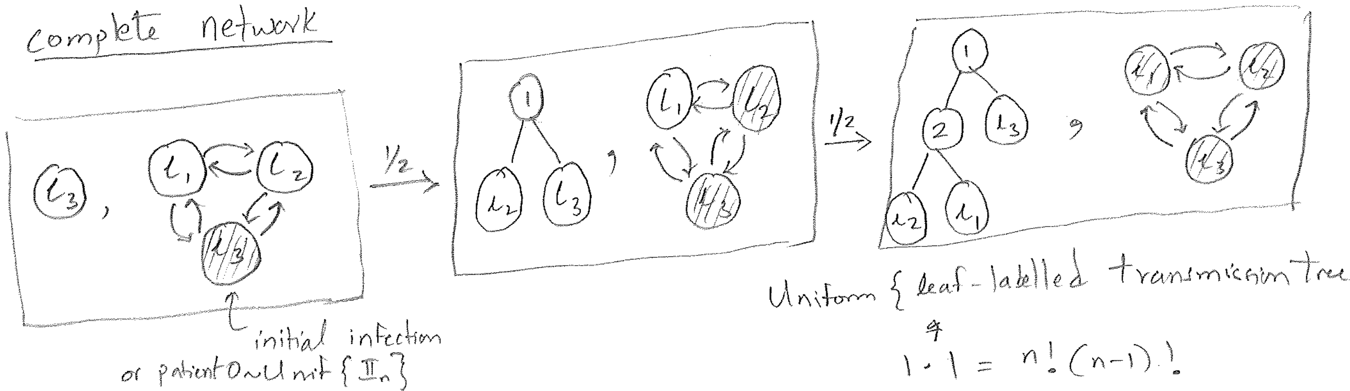
Uniform{right branching comb Tree}.

(ii) Star network



Uniform {left branching comb Trees}
$1 \cdot 1 = (n-1)!$

(iii) complete network



Uniform {leaf-labelled transmission Trees}
$1 \cdot 1 = n!\,(n-1)!$

$\boxed{3}$ Beta-splitting Model.

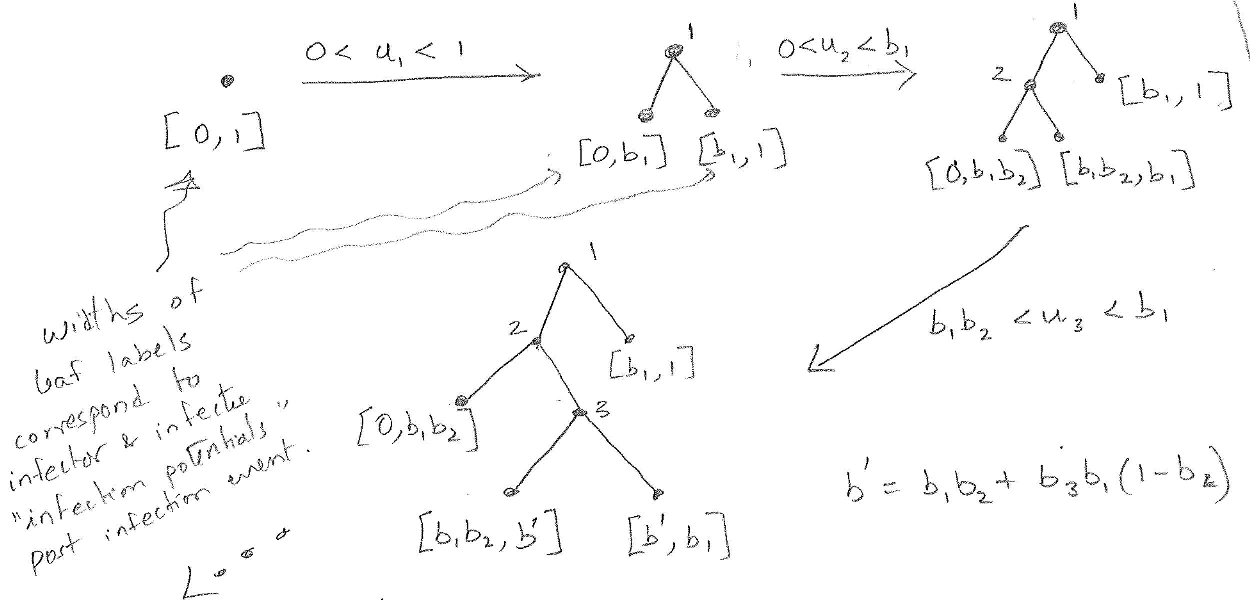
$\boxed{\text{Let}}$ $U_1, U_2, \ldots \overset{iid}{\sim}$ Uniform $[0,1]$

$B_1, B_2, \ldots \overset{iid}{\sim}$ Beta $(\alpha+1, \beta+1)$,

$\alpha, \beta > -1$

$f_{\alpha,\beta} \propto x^{\alpha}(1-x)^{\beta}$

$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$



$$b' = b_1 b_2 + b_3 b_1 (1 - b_2)$$

widths of leaf labels correspond to infector & infectee "infection potentials" post infection event.

L…

Now integrate out the interval-valued realizations of leaf labels

Thm 1

transmission Tree with m splits, $m \leq n-1$

# of splits in left subtree of node z

" " " right " " "

$$Pr\{\tau(m)\} = \prod_{z=1}^{m} \frac{B(\delta_z^L + \alpha + 1, \delta_z^R + \beta + 1)}{B(\alpha+1, \beta+1)} \times Pr\{\text{leaf labels}\}.$$

simplifies to just multiplication and divisions in $\mathbb{R}$.

proof

straight up integration of Beta's pdf. over all split internal nodes $z = 1, \ldots, m$.

$\boxed{4}$. Three Examples revisited with Beta-splitting Model.

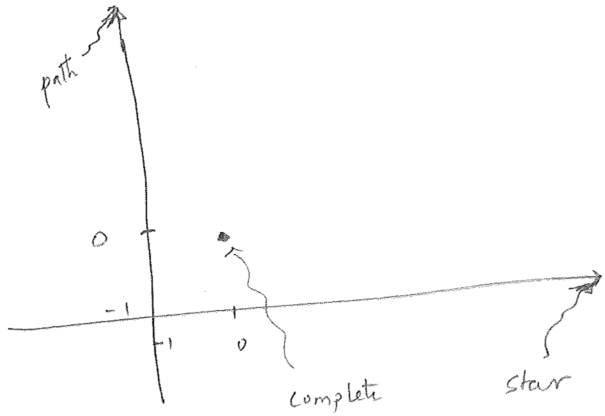
Equivalence of Pr{Transmission Trees}

[NO contact network here!]

case (iii) complete network $\equiv (\alpha, \beta) = (0, 0)$

case (ii) star network. $\equiv (\alpha, \beta) \to (\infty, -1)$

case (i) path network. $\equiv (\alpha, \beta) \to (-1, \infty)$.



?ns – what values of (α, β) correspond to other networks?

– is there an equivalence class of initial SICNs that have same α, β specified distrns on Trans. Trees?

$\boxed{5}$ MLE, Sufficiency & Equiv. Classes.

Let $\mathcal{C}_n^0$ = {all initial distrns starting from a single indiv.}

↑ initial SICNs

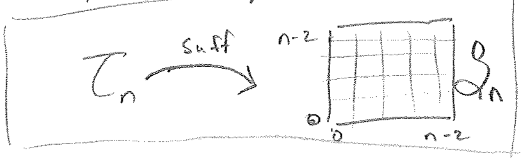
\# of transmission trees (indep. sampled)

Let $(\hat{\alpha},\hat{\beta}) = \operatorname{argmax}_{(\alpha,\beta)\in(-1,\infty)^2} \prod_{i=1}^{r} P_r(\tau_i;\alpha,\beta)$

↑ i-th transmission Tree sampled from the transmission process from an initial SICN.

Thm 2: $P_r(\tau_i;\alpha,\beta)$ only depends on the minimal sufficient statistics of split-pair frequencies:

$$\{f(s^L,s^R) : (s^L,s^R)\in\mathcal{S}_n\},$$



$$\mathcal{S}_n = \{(s^L,s^R)\in(0,1,\dots,n-2)^2 ; s^L+s^R\le n-2\}$$

proof look at lkd. expression $Pr(\tau)$ in Thm 1.