Nearest-neighbour Markov point processes on graphs with Euclidean edges

Marie-Colette van Lieshout

CWI and University of Twente
The Netherlands
Renewal process

Let \( U_i, i = 1, 2, \ldots \), be i.i.d. \textbf{inter-arrival times} with p.d.f. \( \pi \), c.d.f. \( F \). Set \( S_0 \equiv 0 \) and define

\[
S_i = S_{i-1} + U_i
\]

for \( i \in \mathbb{N} \). Those \( S_i \) that fall in \((0, T]\) form a \textbf{renewal process} \( X \) with time horizon \( T \).
Stochastic intensity

The stochastic intensity

\[ h^*(t) \, dt = \mathbb{P}( \text{point in } dt | X \cap (0, t)) \]

of a renewal process is

\[ h^*(t) = h(V_{t-}) \]

where \( V_t \) is the backward recurrence time at \( t \), i.e. the time elapsed since the last event before or at time \( t \), and

\[ h(t) = \frac{\pi(t)}{1 - F(t)} \]

is the hazard rate of \( \pi \).

Disadvantage:

- \( h^* \) is asymmetric in that it depends on the past of the process;
- \( h^* \) does not generalise easily to networks without chronology.
The probability distribution of $X$ (and of any simple point process on the half-line) can be specified by giving

- a (discrete) probability mass function $q_n, n \in \mathbb{N}_0$, for the number of points in $[0, T]$,
- for each $n \in \mathbb{N}$, a symmetric p.d.f. $p_n$ for the locations of the points given there are $n$ of them,

which combine in (un-normalised) **Janossy densities**

$$j_n(t_1, \ldots, t_n) = n! q_n p_n(t_1, \ldots, t_n)$$

for $n \in \mathbb{N}$, with $j_0 = q_0$. 
Interpretation

Janossy densities

- are invariant under permutations of the \( t_i \);
- \( j_n(dt_1, \ldots, dt_n) \) is the probability of finding exactly \( n \) points in \([0, T]\), one each in \( dt_1, \ldots, dt_n \).

Upon normalisation,

\[
p(\{t_1, \ldots, t_n\}) = e^T n! q_n p_n(t_1, \ldots, t_n) \\
= e^T j_n(t_1, \ldots, t_n)
\]

is a probability density w.r.t. the law of the unit rate Poisson process on \([0, T]\).
Renewal process: Probability density

Order the points chronologically:

\[ t_1 < \cdots < t_n. \]

Then

\[
p(\{t_1, \ldots, t_n\}) = e^T \exp \left[ - \int_0^T h^*(t) \, dt \right] \prod_{i=1}^n h^*(t_i)
\]

\[
= e^T (1 - F(T - t_n)) \prod_{i=1}^n \pi(t_i - t_{i-1})
\]

w.r.t. the law of a unit rate Poisson process on \([0, T]\). By convention, an empty product is one and \(t_0 = 0\).
Papangelou conditional intensity

Consider a point $t$ such that

$$t_1 < \cdots < t_{i-1} < t < t_i < \cdots < t_n$$

and define

$$\lambda(t|\{t_1, \ldots, t_n\}) = \frac{p(\{t, t_1, \ldots, t_n\})}{p(\{t_1, \ldots, t_n\})}$$

provided $p(\{t_1, \ldots, t_n\}) > 0$.

For a renewal process,

$$\lambda(t|\{t_1, \ldots, t_n\}) = \frac{\pi(t - t_{i-1}) \pi(t_i - t)}{\pi(t_i - t_{i-1})}$$

with suitable modifications at the end points.

**Important observation:** $\lambda(t|\{t_1, \ldots, t_n\})$ depends only on the nearest neighbours $t_{i-1}$ and $t_i$ of $t$. 
Nearest-neighbour Markov point process

Baddeley and Møller (1989).

Point process $X$ on $[0, T]$ with p.d.f. $p$ is **nearest-neighbour Markov** if whenever $p(\{t_1, \ldots, t_n\}) > 0$.

(a) $p(y) > 0$ for all $y \subseteq \{t_1, \ldots, t_n\}$;

(b) for all $t \notin \{t_1, \ldots, t_n\}$, $\lambda(t|\{t_1, \ldots, t_n\})$ depends only on $u$, its nearest neighbours in $\{t_1, \ldots, t_n\}$ and their mutual nearest neighbour relations.

**Note:** in higher dimensions, points are nearest neighbours (**Delaunay neighbours**) if their Voronoi cells overlap.
Hammersley–Clifford theorem

A point process $X$ with p.d.f. $p$ is nearest-neighbour Markov iff $p({t_1, \ldots, t_n})$ can be factorised up to normalisation as

$$
\begin{align*}
\begin{cases}
\prod_{i=1}^n \gamma({t_i}) \prod_{i=2}^n \gamma({t_{i-1}, t_i}) & \text{if } \gamma({t_i, t_j}) > 0 \text{ for all distinct } t_i, t_j \\
0 & \text{otherwise}
\end{cases}
\end{align*}
$$

where $t_1 < \cdots < t_n$ and $\gamma \geq 0$ is an interaction function:

if $\gamma({t_1, \ldots, t_n}) > 0$,

- $\gamma(y) > 0$ for subsets $y$ of $\{t_1, \ldots, t_n\}$;
- if $\gamma({t_{i-1}, t, t_i}) > 0$, $t_{i-1} < t < t_i$, then also $\gamma({t, t_1, \ldots, t_n}) > 0$.

Example: if $\pi > 0$, the renewal process is hereditary and

$$
\begin{align*}
\gamma({t_i}) &= \pi(t_i)(1 - F(T - t_i)), \\
\gamma({t_{i-1}, t_i}) &= \frac{\pi(t_i - t_{i-1})}{\pi(t_i)(1 - F(T - t_{i-1}))}.
\end{align*}
$$
Objective in this talk

**Aim:** develop analogues of renewal processes on networks, exploiting their one-dimensional nature.

To do so, we study the Delaunay relation on networks, which define nearest-neighbour Markov point process in the sense of Baddeley and Møller (1989) by a probability density of the form

\[ p(x) \propto 1\{\gamma(x) > 0\} \prod_{x_i \in x} \gamma(\{x_i\}) \prod_{i < j : x_i \sim x_j} \gamma(\{x_i, x_j\}), \]

where \( \gamma \geq 0 \) is an interaction function, at least on a tree. In the general case, a local Delaunay relation is proposed.
Features:

- inherent network geometry;
- the network may not be closed under translations;
- focus on first and second order summary statistics, see Rakshit et al. (2017) and references therein.
Models on linear networks

Examples:

- Baddeley et al. (2017): Cox, Switzer-type and cell process.
- Anderes et al. (2016): expand the modelling framework in various directions:
  - a more general definition in terms of parametrised curves that may or may not overlap;
  - parametrisations naturally define a weighted shortest path distance;
  - investigation of alternative definitions of distance;
  - construction of log-Gaussian Cox processes in terms of a Gaussian process on the network specified by an isotropic covariance function.
A graph with Euclidean edges is a triple $G = (V, E = (e_i)_i, \Phi = (\phi_i)_i)$ s.t.

- $(V, E)$ is a finite, simple connected graph (no loops/multiple edges);
- every edge $e_i = \{v^1_i, v^2_i\} \in E, v^1_i, v^2_i \in V$, is parametrised by the inverse of a homeomorphism $\phi_i$, and

$$
\phi_i^{-1} : J_i \rightarrow \mathbb{R}^2
$$

is $C^1$ on an open interval $\emptyset \neq J_i \subseteq \mathbb{R}$ with endpoints $\phi_i(v^j_i), j = 1, 2$, and has induced length $|\phi_i(v^1_i) - \phi_i(v^2_i)|$.

If the edge interiors are disjoint, define the network

$$
L = V \cup \bigcup_{i=1}^{n(E)} \phi_i^{-1}(J_i)
$$

equipped with weighted shortest path distance $d_G$. 
Example: Chicago crime data

![Map of Chicago crime data with various crime types represented by different symbols.]

- assault
- burglary
- car theft
- damage
- robbery
- theft
- trespass
The Delaunay relation on a network

For a finite configuration $x \subseteq L$ of distinct points define

$$x_i \sim_x x_j \iff C(x_i|x) \cap C(x_j|x) \neq \emptyset$$

where

$$C(x_j|x) = \{y \in L : d_G(y, x_j) \leq d_G(y, x) \text{ for all } x \in x\}$$

is the Voronoi cell of $x_j$ in $x$.

The $x$–neighbourhood of a subset $y \subseteq x$ is defined as

$$N(y|x) = \{x \in x : x \sim_x y \text{ for some } y \in y\}.$$ 

The configuration $y$ is an $x$-clique if for each $y_i, y_j \in y$, $y_i \sim_x y_j$. 
Properties

• Voronoi cells are not necessarily line segments;
• points may have more than two neighbours;
• cliques need not be pairs.

The **clique indicator function** satisfies

• if $\chi(y|x) = 1$, then $\chi(y|z) = 1$ for all $y \subseteq z \subseteq x$;
• if $\chi(y|z) = 0$, then $\chi(y|x) = 0$ for all $y \subseteq z \subseteq x$.
Definition

A graph with Euclidean edges $(V, E, \Phi)$ is a **tree** if the graph $(V, E)$ is. Specifically, $(V, E)$ has no **cycles**, i.e. no closed path $(v_0, v_1, \ldots, v_p, v_0)$, $v_i \in V$, over different edges having positive length ($p > 0$).

Lemma

A graph with Euclidean edges is a tree iff there is exactly one path between any two points in its associated network $L$. 
The Delaunay relation on a tree: Paths

**Lemma**

Let \((V, E, \Phi)\) be a tree, \(L\) its associated network and \(y = \{y_1, y_2, y_3\} \subseteq L\). Then there exist unique paths between the elements of \(y\) that

- either form a star with three stokes of strictly positive length emanating from a vertex \(v \in L\),
- or combine into a single path.
The Delaunay relation on a tree: Cliques

$y \subseteq L$ is in **general position** if no three points lie on the boundary of the same $d_G$-ball.

**Lemma**

Clique sizes are at most two on the class of configurations in general position.

Moreover, for all $y = \{y_1, y_2\} \subseteq x$ with $x$ in general position, $\chi(y|x) = 1$ iff the midpoint of $y$ w.r.t. $d_G$ along the unique path between $y_1$ and $y_2$ lies in $C(y_1|x) \cap C(y_2|x)$.

The blue points form a clique iff the stokes are equally long.
**Consistency conditions**

Baddeley and Møller (1989).

For all finite point configurations $y \subseteq z \subseteq L$ and all $u, v \in L$ such that $u, v \notin z$,

(C1) $\chi(y|z) \neq \chi(y|z \cup \{u\})$ implies $y \subseteq N(\{u\}|z \cup \{u\});$

(C2) if $u \sim_x v$ for $x = z \cup \{u, v\}$, then

$$\chi(y|z \cup \{u\}) + \chi(y|z \cup \{v\}) = \chi(y|z) + \chi(y|x).$$

**Theorem**

Let $(V, E, \Phi)$ be a graph with Euclidean edges that is a **tree** and $L$ its associated network. Then the Delaunay relation satisfies (C1)–(C2) on the family of configurations in **general position**.
Let $z$ contain the blue points. Then $\chi(z|z) = 1$. Adding any $u \notin z$ would separate two points of $z$ ($\chi(z|z \cup \{u\}) = 0$) while $u$ is no neighbour of the third point $z$ ($z \notin N(\{u\}|z \cup \{u\})$).

Adding any $v$ s.t. $u \not\sim_{z \cup \{u, v\}} v$,

$$\chi(z|z \cup \{u, v\}) = \chi(z|z \cup \{u\}) = \chi(z|z \cup \{v\}) = 0$$

although $\chi(z|z) = 1$. 

Counterexample
Hammersley–Clifford theorem for trees

Consider \((V, E, \Phi)\) for which \((V, E)\) is a tree. Let \(p \geq 0\) be a function of finite configurations in general position. Then \(p\) is **Markov wrt the Delaunay relation** iff \(p(x)\) is proportional to

\[
\begin{cases}
\prod_{x_i \in x} \gamma(\{x_i\}) \prod_{i < j : x_i \sim x \; x_j} \gamma(\{x_i, x_j\}), & \text{if } \gamma(\{x_i, x_j\}) > 0 \text{ for all distinct } x_i, x_j \\
0, & \text{otherwise}
\end{cases}
\]

for some function \(\gamma \geq 0\) (extended to \(x\) with \(n(x) \geq 3\) by \(\gamma(x) = 1\{\gamma(y) > 0 \ \forall y \subseteq x, n(y) \leq 2\}\)) that is an **interaction function**: if \(\gamma(x) > 0\), then

- \(\gamma(y) > 0\) for all \(y \subseteq x\);
- if \(\gamma(N(\{u\}|x \cup \{u\})) > 0\) then \(\gamma(x \cup \{u\}) > 0\).
The local Delaunay relation

Define $\sim_E$ on $L$ as follows:

- points on edges are neighbours if their edges share a common vertex;
- vertices are neighbours of points on their incident (closed) edges;

Write $\sim_{i,j}^x$ for the Delaunay relation restricted to

$$L \cap \left( \phi_i^{-1}(\bar{J}_i) \cup \phi_j^{-1}(\bar{J}_j) \right).$$

The local Delaunay relation $\sim_{Z}^{E}$ is defined as follows:

- for $x, y$ on edges $i, j$ that share a common vertex $x \sim_{Z}^{E} y$ iff
  
  $$x \sim_{Z}^{i,j} y;$$

- for $x, y$ on the same (closed) edge $e_i$, $x \sim_{Z}^{E} y$ iff $x \sim_{Z}^{i,i} y.$
Theorem
Let $G = (V, E, \Phi)$ be a graph with Euclidean edges and $L$ its associated network. If $G$ does not contain any triangles, then the local Delaunay relation $\sim^E_x$ satisfies (C1)–(C2). By adding an artificial extra vertex in every triangle, (C1)–(C2) can be made to hold generally.