# Bayes Filtering for Spatio-Temporal Poisson Cluster Point Processes

Daniel Clark and Anthony Swain
Joint Research Institute in Signal and Image Processing
Heriot-Watt University
UK

4th August 2009

# Motivation

Suppose that we recursively receive data sets that are observations of some *dynamic cluster process*.
Is it possible to derive optimal formulae (in a Bayesian sense) to estimate:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

# Motivation

Suppose that we recursively receive data sets that are observations of some *dynamic cluster process*.
Is it possible to derive optimal formulae (in a Bayesian sense) to estimate:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

# Motivation

Suppose that we recursively receive data sets that are observations of some *dynamic cluster process*.
Is it possible to derive optimal formulae (in a Bayesian sense) to estimate:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

# Motivation

Suppose that we recursively receive data sets that are observations of some *dynamic cluster process*.
Is it possible to derive optimal formulae (in a Bayesian sense) to estimate:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

1. Spatial Point Processes
   - i.i.d. Cluster Processes
   - Poisson Point Processes
   - Probability Generating Functionals
   - General Cluster Processes

2. Multi-Object Bayes Filtering
   - The Multi-Object Bayes Filter
   - p.g.fl. Bayes Filter
   - Probability Hypothesis Density (PHD) Filtering
   - Poisson Cluster PHD filters

3. Discussion

# Definition: Spatial Point Process

A *spatial point process* is defined as an unordered set of points

$$X = \{x_1, \ldots, x_N\} \tag{1}$$

located in some complete separable metric space $B$, such as $R^d$, known as the *state space*.

A *multi-object probability distribution* $p_X$ on measure space $(X, B)$ defines the distribution of the points.

A *cardinality distribution* $\{p_n\}$, determines the total number of points and satisfies

$$\sum_{n=0}^{\infty} p_n = 1. \tag{2}$$

# Example: i.i.d. cluster process

Let $p(n)$ be the cardinality distribution of the point process $X$ and let $f(y)$ be a probability density function on state space $\mathscr{X}$. For any random set $Y = \{y_1, \ldots, y_n\}$ with $|Y| = n$, define the multi-object probability distribution

$$f(Y) := n! \cdot p(n) \cdot f(y_1) \ldots f(y_n). \tag{3}$$

This is known as an *i.i.d. cluster process*.

# Example: Poisson point process

Suppose that in equation (3), that the cardinality distribution $p(n)$ is given by

$$p(n) = \frac{\exp(-\lambda)\,\lambda^n}{n!}, \quad (4)$$

and the multi-object distribution is

$$f(Y) := e^{-\lambda} \cdot \lambda^n \cdot \prod_{y_i \in Y} f(y_i). \quad (5)$$

This is known as a *Poisson point process*.

# Definition: Probability Generating Functional (p.g.fl.)

The *probability generating functional (p.g.fl.)* of a multi-object probability distribution $p_X$ with measure space $(X, B)$, is defined as the expectation value of symmetric test function

$$h^X := \prod_{x \in X} h(x), \tag{6}$$

so that

$$G_X[h] := E\left(h^X\right) = \int p_X(X) h^X \delta X, \tag{7}$$

where we adopt the concept of the *set integral*,

$$\int f(X)\delta X := f(\phi) + \sum_{n=1}^{\infty} \frac{1}{n!} \int f(\{x_1, \dots, x_n\}) dx_1 \dots dx_n. \tag{8}$$

# Example: Poisson point process

The p.g.fl. of a *Poisson point process* is given by

$$G_P[h] = \exp\left(\mu \int h(x)f(x)dx - \mu\right) \tag{9}$$

The intensity function, or PHD, of a Poisson point process is found by taking the functional derivative of (9), evaluated at $h = 1$, i.e.

$$D(x) := \frac{\delta}{\delta x} G_P[h]\bigg|_{h=1} = \mu \cdot f(x), \tag{10}$$

where $\mu$ gives the expected number of objects that are distributed according to $f(x)$.

# Joint p.g.fl.

A joint probability generating functional, $F_{X,Y}$ of point processes $X$ and $Y$ can be defined by

$$F_{X,Y}[g,h] := \int\int g^X \cdot h^Y \cdot f_{X,Y}(x,y)\delta X \delta Y. \tag{11}$$

The multi-object Bayes update is defined in terms of iterated functional derivatives of a joint p.g.fl.

# Conditioning

Clustering and doubly stochastic models are defined by a procedure that involves conditioning of a p.g.fl. as an intermediate step,

$$G_L[h|x] = \int h(z)g(z|x)dz. \tag{12}$$

Conditioning is used in both the prediction and update steps of the p.g.fl. Bayes filter for the Markov transition and object likelihood.

# Random deletion, or thinning

Random deletion, or thinning refers to the concept of point deletion within the point process. The p.g.fl. for the thinning process is

$$G_T[h] = 1 - p_D(x) + p_D(x) \cdot h(x). \tag{13}$$

Thinning is used in Bayes prediction to deal with the possibility of missed detections.

# Superposition

Let $X_1$ and $X_2$ be two independent point processes. We can define a third point process to be the superposition of these two, The resulting p.g.fl. is simply the multiplication of the two independent p.g.fl.s,

$$G_X[h] = G_{X_1}[h] \cdot G_{X_2}[h]. \quad (14)$$

# General Cluster Processes

The General Cluster Process is characterised by a daughter process p.g.fl. $G_m$ within the parent cluster centre p.g.fl. $G_c$,

$$G_c[G_m[h|\cdot]], \tag{15}$$

where $G_m[h|\cdot]$, treated as an argument of $G_c$, is the p.g.fl. of the daughter process for any particular realisation of the parent process.

# Example: The Poisson Cluster Process

The Poisson cluster process, is a doubly-stochastic Poisson process. Its p.g.fl. is given by

$$G_c[G_m[h|\cdot]] = \exp\left(\mu \int s_1(u) \exp\left(\nu(u) \int s_2(w|u)\, h(w|u)\, dw - \nu(u)\right) du - \mu\right) \tag{16}$$

# Example: The Poisson Cluster Process

The intensity function can be found by taking the functional derivative

$$\frac{\delta}{\delta(c,y)} G_c[G_m[h|\cdot]]\Big|_{h=1} = \mu s_1(c) \cdot \nu(c) s_2(y|c), \tag{17}$$

where $y$ is on the space of the daughter process and $c$ is on the parent space.

$$\begin{aligned} \mu s(c) &= \text{Poisson intensity function on parent space,} \\ \nu(c) s_2(y|c) &= \text{Poisson intensity function on daughter space.} \end{aligned}$$

# Matern Cluster Process (Uniform circular daughter)

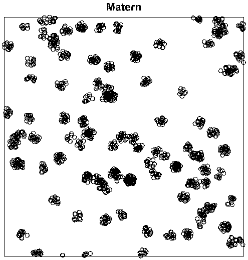


Figure: Realisation of a Matérn cluster process with parent Poisson rate 100, and daughter process Poisson rate 15 with radius 0.02.

# Modified Thomas Process (circular Gaussian daughter)

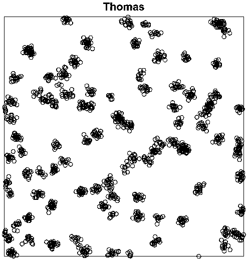


Figure: Realisation of a modified Thomas cluster process with parent Poisson rate 100, and daughter process Poisson rate 15 with standard deviation 0.01.

# Spatial Point Processes - Summary

In summary:

- Spatial point processes model random spatial point patterns
- Uniquely specified in terms of p.g.fl.s
- Can be approximated with its intensity function

In the next section, I discuss spatio-temporal point processes.

# Spatial Point Processes - Summary

In summary:

- Spatial point processes model random spatial point patterns
- Uniquely specified in terms of p.g.fl.s
- Can be approximated with its intensity function

In the next section, I discuss spatio-temporal point processes.

# Spatial Point Processes - Summary

In summary:

- Spatial point processes model random spatial point patterns
- Uniquely specified in terms of p.g.fl.s
- Can be approximated with its intensity function

In the next section, I discuss spatio-temporal point processes.

# Multi-Object Bayes Filtering

The optimal multi-target Bayes filter propagates the multi-target posterior density $p_k(\cdot|Z_{1:k})$ conditioned on the sets of observations up to time $t$, $Z_{1:k}$, with the following recursion

$$p_{k|k-1}(X_k|Z_{1:k-1}) = \int f_{k|k-1}(X_k|X)p_{k-1}(X|Z_{1:k})\delta X, \qquad (18)$$

$$p_k(X_k|Z_{1:k}) = \frac{g_k(Z_k|X_k)p_{k|k-1}(X_k|Z_{1:k-1})}{\int g_k(Z_k|X)p_{k|k-1}(X|Z_{1:k-1})\delta X}, \qquad (19)$$

where

$$\begin{aligned} f_{k|k-1}(X_k|X_{k-1}) &= \text{multi-object Markov density}, \\ Z_k &= \text{measurement set at time } k, \\ g_k(Z|X) &= \text{multi-object measurement likelihood at time } k \end{aligned}$$

# p.g.fl. Bayes filter

An equivalent representation of the multi-object Bayes filter can be written in terms of p.g.fl.s. The p.g.fl. form of the Bayes filter is given by

$$G_{k|k-1}[h] = \int\int h^Y f_{k|k-1}(Y|X) p_{k-1}(X|Z_{1:k}) \delta Y \delta X, \tag{20}$$

$$G_k[h] = \frac{\int h^Y g_k(Z_k|Y) p_{k|k-1}(Y|Z_{1:k-1}) \delta Y}{\int g_k(Z_k|X) p_{k|k-1}(X|Z_{1:k-1}) \delta X}, \tag{21}$$

where

$$\begin{aligned} G_{k|k-1}[h] &= \text{p.g.fl. prediction,} \\ G_k[h] &= \text{p.g.fl. Bayes update} \end{aligned}$$

# Probability Hypothesis Density (PHD) Filtering

The intensity, or PHD, recursion is found by taking the first moment of the p.g.fl. Bayes filter

$$D_{k|k-1}(x) = \frac{\delta}{\delta x} \int \int h^Y f_{k|k-1}(Y|X) p_{k-1}(X|Z_{1:k}) \delta Y \delta X, \tag{22}$$

$$D_k(x) = \frac{\delta}{\delta x} \frac{\int h^Y g_k(Z_k|Y) p_{k|k-1}(Y|Z_{1:k-1}) \delta Y}{\int g_k(Z_k|X) p_{k|k-1}(X|Z_{1:k-1}) \delta X}, \tag{23}$$

where

$$\begin{aligned} D_{k|k-1}(x) &= \text{PHD prediction}, \\ D_k(x) &= \text{PHD update}, \end{aligned}$$

# Example: PHD update

Define a joint p.g.fl.

$$F[g,h] = \int \int h^X g^Z g_k(Z|X) p_{k|k-1}(X|Z_{1:k-1}) \delta X \delta Z \tag{24}$$

Then the p.g.fl. update and PHD update can be written as

$$G_k[h] = \frac{\frac{\delta F}{\delta Z}[0,h]}{\frac{\delta F}{\delta Z}[0,1]} \tag{25}$$

$$D_k(x) = \frac{\delta}{\delta x} G_k[h] \tag{26}$$

# Example: PHD update

The joint p.g.fl. can be written as a composition of p.g.fl.s,

$$F[g,h] = G_{\Theta}[g] \cdot G_P\left[h \cdot G_T\left[G_L\left[g|\cdot\right]\right]\right], \tag{27}$$

where

$$\begin{aligned} G_P[h] &= \text{Poisson predicted p.g.fl.,} \\ G_{\Theta}[g] &= \text{false alarm p.g.fl. (superposition),} \\ G_T[h] & \quad \text{missed detection p.g.fl. (thinning),} \\ G_L[g|\cdot] &= \text{single-object likelihood p.g.fl. (conditioning).} \end{aligned}$$

# Example: PHD update

Assume that the multi-object prediction $p_{k|k-1}(X|Z_{1:k-1})$ and false alarm processes are Poisson, i.e.

$$G_{k|k-1}[h] = \exp(\mu \int s(x) \cdot h(x) dx - \mu) \quad (28)$$

Then the PHD update equation is

$$D_k(x) = (1 - p_D(x)) D_{k|k-1}(x) + \sum_{z \in Z} \frac{p_D(x) g_k(z|x) D_{k|k-1}(x)}{\lambda_k c_k(z) + \int p_D(\xi) g_k(z|\xi) D_{k|k-1}(\xi) d\xi} \quad (29)$$

where

$$\begin{array}{lcl} Z_k & = & \text{measurement set at time } k, \\ g_k(\cdot|x) & = & \text{single target measurement likelihood at time} k \\ p_{D,k}(x) & = & \text{probability of target detection at time } k \\ \lambda_k c_k(\cdot) & = & \text{Poisson intensity of clutter measurements at time } k, \end{array}$$

# Poisson Cluster PHD filters

We now consider a Bayes update for Poisson cluster processes. Similar to the Poisson PHD update, the Poisson cluster PHD update can be written as

$$G_k[h] = \frac{\frac{\delta F}{\delta Z}[0,h]}{\frac{\delta F}{\delta Z}[0,1]} \tag{30}$$

$$\mu s_k(u) D_k(y|u) = \frac{\delta}{\delta(u,y)} G_k[h], \tag{31}$$

where

$$\begin{aligned} \mu s_k(u) D_k(y|u) &= \text{Poisson cluster intensity,} \\ \mu s_k(u) &= \text{Poisson parent process,} \\ D_k(y|u) &= \text{Poisson daughter conditional intensity,} \end{aligned}$$

# Poisson Cluster p.g.fl. update

For simplicity, we do not consider clutter or missed detections. The joint p.g.fl. can be written as the composition

$$F[g,h] = G_C\left[G_P\left[h \cdot G_L\left[g|\cdot\right]|\cdot\right]\right], \tag{32}$$

where

$$\begin{aligned} G_C[h] &= \text{parent (Poisson) p.g.fl.,} \\ G_P[h|\cdot] &= \text{daughter Poisson p.g.fl. (conditioning),} \\ G_L[g|\cdot] &= \text{single-object likelihood p.g.fl. (conditioning).} \end{aligned}$$

# Poisson Cluster p.g.fl. update

For this case we consider a doubly-Poisson process with no clutter and no missed detections, the p.g.fl. is

$$
\begin{aligned}
F[g,h] &= G_c[G_y[h \cdot G_L[g|\cdot]|\cdot]] \\
&= \exp\left\{\mu \int s_1(u) \exp\left\{\nu(u) \int s_2(w|u)\, h(w|u)\right.\right. \qquad (33) \\
&\quad \times \left.\left.\left(\int g(z) f(z|w) f(w|u)\, dz\right) dw - \nu(u)\right\} du - \mu\right\}
\end{aligned}
$$

where

$$
\begin{aligned}
G_C[h] &= \text{parent (Poisson) p.g.fl.,} \\
G_P[h|\cdot] &= \text{daughter Poisson p.g.fl. (conditioning),} \\
G_L[g|\cdot] &= \text{single-object likelihood p.g.fl. (conditioning).}
\end{aligned}
$$

# Poisson Cluster PHD update

The denominator of the intensity of this case is found to be

$$e^{\langle D_1, e^{-\nu}-1\rangle} \sum_{P \in \mathscr{P}(Z)} \prod_{W \in P} \langle D_1, e^{-\nu} \hat{L}_W^c \rangle \tag{34}$$

where the sum is over all partitions of the measurement sets, and the (outmost) product is over all subsets, and the (innermost) product is over measurements.

$$\begin{aligned} D_1(c) &= \mu s_1(c), \\ \hat{L}_W^c &= \nu(c) s_2(y|c), \\ L_{\mathbf{z}}^y(y|c) &= f(\mathbf{z}|y) f(y|c), \end{aligned}$$

# Poisson Cluster PHD update

The numerator of the intensity of this case is found to be

$$e^{\langle D_1, e^{-\nu} - 1 \rangle} \sum_{P \in \mathscr{P}(Z)} \sum_{W \in P} \hat{L}^c_W(c) D_1(c) e^{-\nu} \times \prod_{W' \in P - W} \langle D_1, e^{-\nu} \hat{L}^c_{W'} \rangle \sum_{\mathbf{z} \in W} \frac{L^y_{\mathbf{z}}(y|c) D_2(y|c)}{\langle D_2, L^y_{\mathbf{z}} \rangle} \quad (35)$$

# Poisson Cluster PHD update

The updated intensity becomes

$$\sum_{P \in \mathscr{P}(Z)} \omega_P \sum_{W \in P} \frac{\hat{L}_W^c(c) D_1(c) \mathrm{e}^{-\nu}}{\langle D_1, \mathrm{e}^{-\nu} \hat{L}_W^c \rangle} \sum_{z \in W} \frac{L_z^y(y|c) D_2(y|c)}{\langle D_2, L_z^y \rangle} \tag{36}$$

where

$$\omega_P = \frac{\prod_{W \in P} \langle D_1, \mathrm{e}^{-\nu} \hat{L}_W^c \rangle}{\sum_{Q \in \mathscr{P}(Z)} \prod_{W \in Q} \langle D_1, \mathrm{e}^{-\nu} \hat{L}_W^c \rangle} \tag{37}$$

is the weight of partition $P$.

# Complexity of Poisson Cluster Process

| No. measurements | Numerator, $nB_n$ terms | Denominator, $B_n$ terms |
|---|---|---|
| 1 | 1 | 1 |
| 2 | 4 | 2 |
| 3 | 15 | 5 |
| 4 | 60 | 15 |
| 5 | 260 | 52 |
| 6 | 1,218 | 203 |
| 7 | 6,139 | 877 |
| 8 | 33,120 | 4,140 |
| 9 | 190,323 | 21,147 |
| 10 | 1,159,750 | 115,975 |

Figure: Number of terms for the doubly-Poisson PHD filter with no clutter or missed detections.

# Discussion

Presented a new class of filters for Bayes-optimal dynamic clustering for estimating:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

If the complexity issues could be resolved, it is anticipated that it could be useful for engineering applications in target tracking and statistical applications in epidemiology.

# Discussion

Presented a new class of filters for Bayes-optimal dynamic clustering for estimating:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

If the complexity issues could be resolved, it is anticipated that it could be useful for engineering applications in target tracking and statistical applications in epidemiology.

# Discussion

Presented a new class of filters for Bayes-optimal dynamic clustering for estimating:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

If the complexity issues could be resolved, it is anticipated that it could be useful for engineering applications in target tracking and statistical applications in epidemiology.

# Discussion

Presented a new class of filters for Bayes-optimal dynamic clustering for estimating:

- How many clusters there are?
- Where are the clusters located?
- How many members within each cluster and their locations?
- How the clusters and elements evolve over time?

If the complexity issues could be resolved, it is anticipated that it could be useful for engineering applications in target tracking and statistical applications in epidemiology.