Retro-preferential Stochastic Mobility Models on Random Fractals Under Sporadic Observations

R. W. R. Darling

March 22, 2018

Abstract

In modelling a sequence of adaptive choices by an intelligent agent (e.g. places visited, web sites browsed), memory-less random walks are unsuitable, because of the formation of agent habits and preferences. Often these choices are only partially observed, and report times are sporadic and bursty, in contrast to regular or exponentially spaced times in classical models. The FRACTALRABBIT stochastic mobility simulator creates realistic synthetic sporadic waypoint data sets. It consist of three tiers, each based on new stochastic models:

1. An *Agoraphobic Point Process* generates a set V of points in $\mathbf{R}^d$, whose limit is a random fractal, representing sites that could be visited.
2. A *Retro-preferential Process* generates a trajectory X through V, with strategic homing and self-reinforcing site fidelity as observed in human/animal behavior.
3. A *Sporadic Reporting Process* models time points T at which the trajectory X is observed, with bursts of reports and heavy tailed inter-event times.

FRACTALRABBIT is being used to test algorithms applicable to sporadic waypoint data, such as (1) co-travel mining, (2) anomaly detection, and (3) extraction of maximal self-consistent subsets of corrupted data.

Keywords: mobility model, point process, Zipf Law, preferential attachment, fractal dimension

1. Introduction

1.1. Background

A classic stochastic model dominates the engineering literature on target tracking: a Markov process on a Euclidean state space, with known dynamics, is observed at regular times, through additive noise with a known statistical law. Texts such as Elliott, Aggoun, and Moore [6], describe an optimal state estimation paradigm based on recursive filters, such as the Kalman filter.

In applications such as:

- Animal motion, observed sporadically from fixed sensors,
- Sequence of web sites visited by one person during many weeks,
- Sequence of songs one person plays on their smartphone,

such low-dimensional hidden Markov models with regular observations may not be applicable. The plain Markov assumption is invalid because the entire history, not just the last event, influences the chance of the next event. Biologists observe *strategic homing* and *self-reinforcing site fidelity*, which are not compatible with memoryless Markov processes. These are also observed in human behaviour [8]. Barabasi et al. [1] point out that indications of location may be bursty and sporadic, rather than regular; intervals between successive observations, rather than being exponentially distributed as in a Poisson process, appear to have heavy-tailed distributions. Errors need not be additive; observations may be imprecise statements lacking a suitable probability model.

There are non-Markov models of animal motion, in a discrete or continuous state space, which exhibit preferential relocations to places visited in the past, see Boyer and Solis-Salas [2], and works cited therein. Fujihara & Miwa [7] propose a "homesick Lévy walk". There is also a probabilistic literature on random walk on graph vertices with reinforcement, such as Pemantle [11]. The present work has some conceptual similarities to both the last cited works, but our model displays quite different statistical properties.

The set of places on the Earth's surface often visited by humans seem to form a set of Hausdorff dimension less than 2, at least across suitable scales. For an example, see the *developed high intensity* cells in the National Land Cover Database 2011 [9]. In a companion paper [5] we develop Euclidean embeddings of random trees whose vertex sets form random fractals suitable for modelling places visited, and employ these embeddings in the Agoraphobic Point Process.

1.2. Overview

There are three types of stochastic model described in this paper, which may be assembled to give a model of motion under sporadic imprecise observation.

PLACES We construct recursively a set $V = \{v_0, v_1, \ldots, v_k\}$ in the unit ball of $\mathbf{R}^d$, called the **Agoraphobic Point Process**, whose limit is a random fractal. Such a V provides source data for a Retro-preferential Process whose $(P_{i,j})$ matrix is derived from inter-point distances.

VISITS Given a finite set $V = \{v_0, v_1, \ldots, v_k\}$, a matrix $(P_{i,j})_{0 \le i,j \le n}$, where $i \ne j$ implies $P_{i,j} > 0$, and a non-stationary Bernoulli process $Z_1, Z_2, \ldots$, we construct a random trajectory $(x_0, x_1, \ldots, x_n)$ consisting of elements of V, called the **Retro-preferential Process**.

REPORTS Finally we propose a **sporadic reporting model**, which involves the number of reports (possibly none) at each point x_j of a trajectory, the spacing between times of consecutive reports at each point, and the interval between the last report at x_j and the first report at x_{j+1}.

We propose to combine these models as follows:

- Use an Agoraphobic Point Process (where $d = 2$) to build a random set V of places with a highly inhomogenous spatial distribution, and then use a Retro-preferential Process to simulate a sequence of visits to some of these places, where the steps are dependent both on the distances between places, and on the history of places already visited. This will be a trajectory $(x_0, x_1, \ldots, x_n)$ in V.
- Use a sporadic reporting model to embed the trajectory in continuous time, and generate reports about location from time to time. Seasonality (for example, of day and night) can also be simulated. Interpretation of a report issued at x_j at time T_j^k may be that "the target is within distance r_j of point x_j at time T_j^k", where r_j is interpreted as the range of a sensor located at the point x_j. In other words, the possible continuous-time trajectories $t \mapsto y(t)$ of the target satisfy

$$\|y(T_j^k) - x_j\| \le r_j, \quad \forall j, k.$$

All these phases are implemented in the open source FRACTALRABBIT [3] software, which also constructs waypoint data sets for multiple travelers on the same fractal set V.

2. Point Processes with Random Fractal Limits

This section is devoted to construction of a finite subset $V \subset \mathbf{R}^d$ which will serve as points visited by a Retro-preferential Process.

2.1. Sequences of Random Rooted Trees

Fix a dimension $d \ge 2$ and a desired fractal dimension $\alpha \in (0, d)$. Take independent, identically distributed mean zero random vectors $Z_1, Z_2, \ldots$ in $\mathbf{R}^d$ with finite second moments, where

$$\mathrm{Cov}[Z_n] = \psi I_d \tag{1}$$

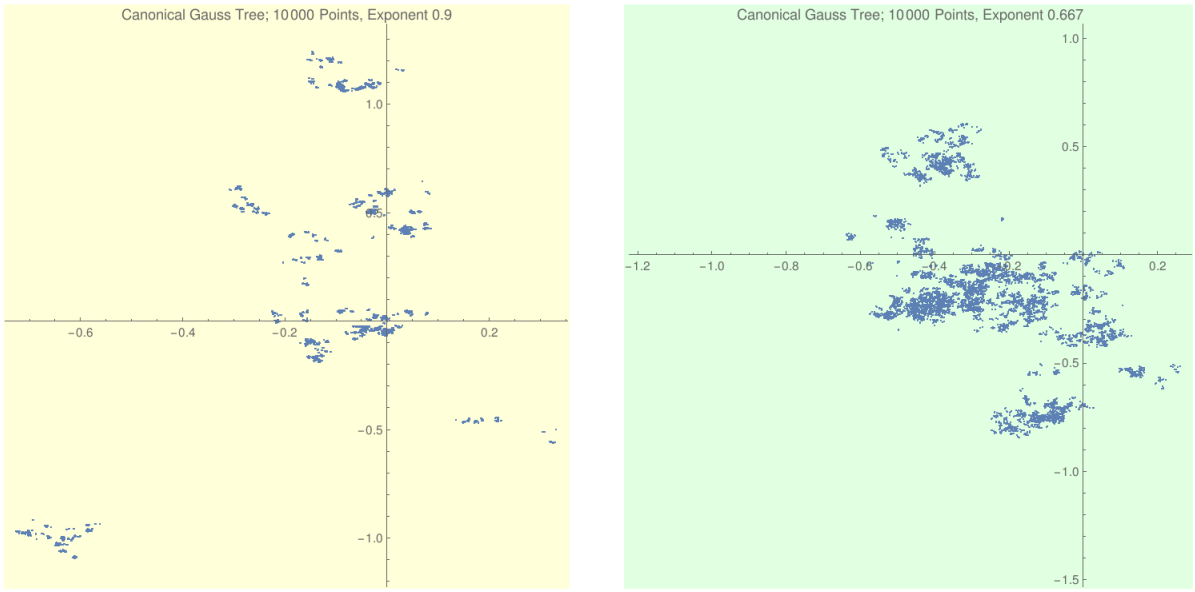
for some $\psi > 0$. Examples include:

- $Z_i \sim$Normal$(\mathbf{0}, I_d)$ (canonical Gaussian tree case), and
- Z_i uniformly distributed in the unit ball of $\mathbf{R}^d$ (hard threshold case).

The precise distribution will not play a part in the proofs.

Construct a nested sequence of directed random rooted trees $(V_n, E_n)_{n\ge0}$ as follows. Intially $V_0 = \mathbf{0}$ and E_0 is empty. Suppose (V_{n-1}, E_{n-1}) exist with $|V_{n-1}| = n$.

Figure 1: *Vertices of a Canonical Gaussian Tree in* $\mathbf{R}^2$: Two simulations of 10,000 points each, with different exponents $1/\alpha$ in (2). Limiting Hausdorff dimensions are 10/9 and 3/2, respectively.



1. Select $v \in V_{n-1}$ uniformly at random.
2. Append to V_{n-1} the new vertex
$$v' := v + n^{-1/\alpha} Z_n. \tag{2}$$
3. Append to E_{n-1} the directed edge $v \to v'$.

Figure 1 shows two instances of V_n for different values of $1/\alpha$, in the Gaussian case.

Theorem 2.1. *The nested sequence* $(V_n, E_n)_{n\geq 0}$ *of rooted trees has a direct limit* (V_∞, E_∞)*, which is an infinite tree rooted at* $\mathbf{0}$*, with* $V_\infty \subset \mathbf{R}^d$*. The elements of* V_∞ *are limits in* L^2 *of directed paths from the root.*

A proof will be given in Section 2.3. A more general branching model in continuous time is analysed in a companion paper [5], which contains a proof that the random set V_∞ has Hausdorff dimension α. A heuristic fractal dimension argument is provided in Section 2.5. Furthermore [5] shows that the diameter of V_∞ is finite with probability one,

2.2. The Agoraphobic Point Process in the Unit Ball

Fix a dimension $d \geq 2$, and a desired fractal dimension $\alpha < d$.

In the model (2), a single parameter α controls both the limiting Hausdorff dimension and the statistics of the number of clumps, For flexibility in modelling, we add an **innovation parameter** $\theta \geq 0$ which controls the rate of insertion of new cluster seeds, and thus affects the number of clumps.

Let B_d denote the closed unit ball[1] in R^d. Unlike the construction (2), this one is designed to create a random tree whose vertices are points in R^d with norm not exceeding 1.

A nested increasing sequence of finite random rooted trees $(\xi_n, E_n)_{n\geq 0}$ will be constructed according to the following rules. Take $\xi_0 := \{\mathbf{0}\}$, i.e. the single point at the origin, which serves as the root, and $E_0 := \emptyset$.

Suppose $n \geq 1$, $\xi_{n-1} \subset B_d$ is a set consisting of n elements, and E_{n-1} is a collection of directed edges so (ξ_{n-1}, E_{n-1}) is a tree rooted at $\mathbf{0}$. To generate ξ_n, perform two random experiments, independent of each other and of ξ_{n-1}:

1. Sample X uniformly at random in B_d.
2. Perform a Bernoulli$(\theta/(\theta + n - 1))$ trial (here $0/0 = 1$).

If the trial is a **success**, define $\xi_n := \xi_{n-1} \cup \{X\}$, declare the parent of X in the tree to be $\mathbf{0}$, and define E_n to be E_{n-1} together with this extra directed edge. In effect we are seeding a new clump, like adding a ball of a new color in the Polya urn of Section 3.1.

If the trial is a **failure** (which cannot occur when $n = 1$) compute

$$\Delta := \min_{x\in\xi_{n-1}\setminus\{\mathbf{0}\}} \|X - x\|. \tag{3}$$

If ξ_{n-1} were a uniform random sample of n points in B_d, this minimum distance Δ would scale like $n^{-1/d}$. To force points to clump together we would like the distance of a new point to the closest previous point to scale like $n^{-1/\alpha}$, where by assumption $\alpha < d$. In the **smooth minimum distance** formulation, we accept X with probability

$$e^{-\Delta n^{1/\alpha}}. \tag{4}$$

If X is accepted, define $\xi_n := \xi_{n-1}\cup\{X\}$, and define E_n to be E_{n-1} together with the directed edge $x \to X$, where $x \in \xi_{n-1} \setminus \{\mathbf{0}\}$ is the minimizing point in (3) (almost surely unique). If X is not accepted, keep sampling X again in the same way until acceptance occurs. The conditional law of ξ_n given ξ_{n-1} is independent of $\xi_0, \xi_1, \ldots, \xi_{n-2}$, and so $\{\xi_n\}_{n\geq 0}$ is a Markov process.

Except when n is small, the formula (4) has the effect of making rejection likely when the proposed $(n + 1)$-st point is further than $n^{-1/\alpha}$ from any existing point in $\xi_{n-1} \setminus \{\mathbf{0}\}$. Clumps form, as shown in Figure 2. Inserting $\xi_0 := \{\mathbf{0}\}$ serves to create a centering effect. The number of distinct clumps, which is like the number of colors present in the Polya urn of Section 3.1, will grow according to (12), namely $\theta \log \frac{n+\theta}{\theta}$ after n points.

[1] This construction generalizes to any compact metric space carrying a probability measure.

2.2.1. Hard Threshold Formulation of Agoraphobic Point Process

Simulation is slow for the Minimum Distance Formulation when n reaches the hundreds, because of high rejection rates. An inequivalent alternative formulation in the case of failure at Step 2., with much lower rejection rate, is given in [5]. The idea is to replace $e^{-\Delta n}$ in the acceptance rate formula (4), which in (3) is a continuous function of the sampled point X, by the **hard threshold**

$$1_{\{\Delta \le n^{-1/\alpha}\}}.$$

To achieve this efficiently, [5] proposes to generate a new point which must lie within distance $n^{-1/\alpha}$ of an existing point, as follows:

1. Select z uniformly from the $n-1$ points in $\xi_{n-1} \setminus \{\mathbf{0}\}$.
2. Sample Y uniformly from the ball of radius $n^{-1/\alpha}$, centered at z.
3. Accept Y with probability $1/k$, where
$$k := \sum_{x \in \xi_{n-1}} 1_{\{\|Y-x\| \le n^{-1/\alpha}\}}.$$
4. If Y is accepted, set $\xi_n := \xi_{n-1} \cup \{Y\}$, and declare z to be the parent of Y in the random tree. If Y is not accepted, keep sampling Y again in the same way until acceptance occurs.

This formulation is reminiscent of the random tree of Section 2.1, except that step 3. prevents oversampling in areas which are already dense. As described above, not all the points need lie inside the unit disk. See Figure 2. The hard threshold method is the default in FRACTALRABBIT [3]. Whichever method we use, we obtain a random rooted tree (ξ_n, E_n) for each n.

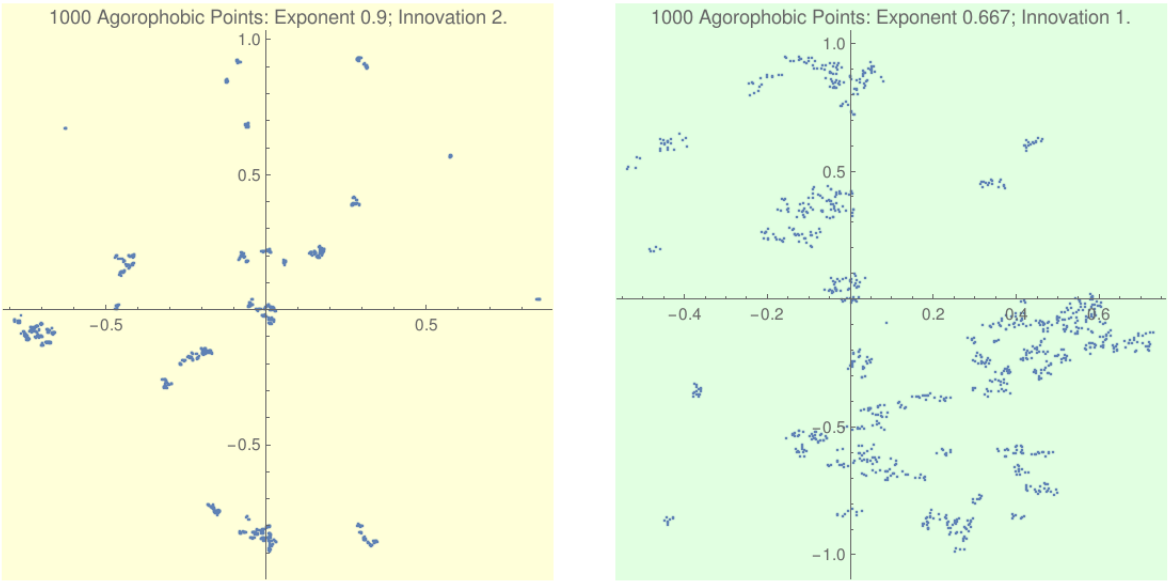
2.3. Limit Set in Random Embedded Rooted Trees

The remainder of this section demonstrates some properties of the model (2). Existence of the limit set V_∞, described in Theorem 2.1, will be established by proving that, for every vertex $v \in V_\infty$, the second moment of the length of the vector v is finite. To prove Theorem 2.1, introduce the notion of the path of **primogeniture** from the root. Based upon $(V_n, E_n)_{n \ge 0}$, we will construct an increasing sequence of random integers $(\eta(t))_{t \ge 0}$, where $\eta(0) = 0$ and $\eta(1) = 1$. Write V_n as $\{v_0, v_1, \ldots, v_n\}$. Given $\eta(t-1) = m$, define $\eta(t)$ to be $m + k$ if the first child of v_m is v_{m+k}. In other words, the directed path

$$v_{\eta(0)} \to v_{\eta(1)} \to v_{\eta(2)} \to \cdots \tag{5}$$

is the line of primogeniture from the ancestor to its first child, then to the first grandchild, and so on. Furthermore (2) implies that the vector in $\mathbf{R}^d$ pointing to the primogeniture

Figure 2: **Agoraphobic Point Process:** *Two instance of the Hard Threshold Version, 1000 points, with different exponents.*



descendant $v_{\eta(t)}$ at depth t from the ancestor is

$$v_{\eta(t)} = \sum_{s=1}^{t} \frac{1}{\eta(s)^{1/\alpha}} Z_{\eta(s)}. \tag{6}$$

Lemma 2.2. *The labels $(\eta(t))_{t\geq 1}$ along the path of primogeniture (5) satisfy $\eta(t) \geq t$, and exhibit exponential growth in the sense that for any $\kappa > 0$,*

$$\lim_{t\to\infty} \mathbf{E}[\left(\frac{\eta(t-1)}{\eta(t)}\right)^{\kappa} \mid \eta(t-1)] = \frac{1}{1+\kappa}. \tag{7}$$

As for moment bounds, $\eta(1) = 1$, and $\mathbf{E}[\eta(t)^{-\kappa}] < t^{-\kappa}$ for $t \geq 2$.

Proof. Recall that when v_{n+1} is being created, its parent is selected uniformly from V_n, where $|V_n| = n + 1$. The chance that the next vertex is not a child of the last one created is

$$\mathbf{P}[\eta(t) - \eta(t-1) > 1 \mid \eta(t-1) = n] = \frac{n}{n+1}, \quad t \geq 2,\, n \geq 1.$$

Since successive choices of parent are independent,

$$\mathbf{P}[\eta(t) - \eta(t-1) > k \mid \eta(t-1) = n] = \frac{n}{n+1} \cdot \frac{n+1}{n+2} \cdots \frac{n+k-1}{n+k} = \frac{n}{n+k}.$$

Take differences to obtain

$$\mathbf{P}[\eta(t) = n + k \mid \eta(t-1) = n] = \frac{n}{n+k-1} - \frac{n}{n+k} = \frac{n}{(n+k)(n+k-1)}.$$

For any $\kappa > 0$, and $t \geq 2$,

$$\mathbf{E}[\left(\frac{\eta(t-1)}{\eta(t)}\right)^{\kappa} \mid \eta(t-1) = n] = \sum_{k \geq 1} \frac{n^{1+\kappa}}{(n+k)^{1+\kappa}(n+k-1)} < n^{-1} \sum_{j \geq n} (j/n)^{-2-\kappa}.$$

The following integral bound applies when $n \geq 2$:

$$n^{-1} \sum_{j \geq n} (j/n)^{-2-\kappa} \leq \int_{1-1/n}^{\infty} x^{-2-\kappa} dx = \frac{(1-1/n)^{-1-\kappa}}{1+\kappa}.$$

There is a similar lower bound, leading to the limiting result (7). □

2.4. Proof of Theorem 2.1

Consider the primogeniture descendant $v_{\eta(t)}$ at depth t from the ancestor as in (6). This has mean zero, and has conditional covariance, given the branching times:

$$\mathrm{Cov}[v_{\eta(t)} \mid (\eta(s))_{0 \leq s \leq t}] = \psi I_d \sum_{s=1}^{t} \frac{1}{\eta(s)^{2/\alpha}}.$$

Lemma 2.2 with $\kappa = 2/\alpha$ implies that, given $\delta \in (0, 2)$, there exists $t \geq 1$ such that

$$\mathbf{E}[\eta(s)^{-2/\alpha}] < \left(\frac{\delta + \alpha}{2 + \alpha}\right)^{s-t} \mathbf{E}[\eta(t)^{-2/\alpha}], \quad \forall s > t.$$

Sum a geometric series, and apply moment bounds, to obtain

$$\mathbf{E}\left[\sum_{s=t}^{\infty} \frac{1}{\eta(s)^{2/\alpha}}\right] < \left(\frac{2+\alpha}{2-\delta}\right) \mathbf{E}[\eta(t)^{-2/\alpha}] \leq \left(\frac{2+\alpha}{2-\delta}\right) t^{-2/\alpha}. \tag{8}$$

On the other hand Lemma 2.2 shows that the corresponding moment sum from $s = 1$ to $s = t - 1$ is bounded by a linear function of t. It follows from (6) that the trace of $\mathrm{Cov}[v_{\eta(t)}]$ is bounded above as $t \to \infty$. Hence $(v_{\eta(t)})_{t \geq 0}$ is an L^2 bounded d-dimensional martingale, and has a limit almost surely.

Take any path $(v_{\theta(t)})_{t \geq 0}$ in (V_∞, E_∞), starting from the root, and compare it with the primogeniture path. We claim that, for any $\kappa > 0$, the upper bounds of Lemma 2.2 hold. For example:

$$\limsup_{t \to \infty} \mathbf{E}[\left(\frac{\theta(t-1)}{\theta(t)}\right)^{\kappa} \mid \theta(t-1)] \leq \frac{1}{1+\kappa}.$$

This is because, by definition of the primogeniture path,

$$\mathbf{E}\left[\frac{1}{\theta(s)^\kappa} \mid \theta(s-1)=n\right] \le \mathbf{E}\left[\frac{1}{\eta(s)^\kappa} \mid \eta(s-1)=n\right].$$

So $(v_{\theta(t)})_{t\ge 0}$ is also an L^2 bounded d-dimensional martingale, with a limit almost surely. We have now proved that directed paths in (V_∞, E_∞) have a limit almost surely in $\mathbf{R}^d$, as claimed.

2.5. Hausdorff Dimension Heuristics

Take any path $(v_{\theta(t)})_{t\ge 0}$ in (V_∞, E_∞), starting from the root. Define σ_n^2 to be the conditional second moment

$$\sigma_n^2 := \mathbf{E}[\| \lim_{s\to\infty} v_{\theta(s)} - v_{\theta(t)}\|^2 \mid \theta(t)=n]$$

By the reasoning of (8), for all sufficiently large t,

$$\sigma_n^2 = \psi d \mathbf{E}\left[\sum_{s=t}^{\infty} \frac{1}{\theta(s)^{2/\alpha}} \mid \theta(t)=n\right] \le \frac{\psi d}{n^{2/\alpha}}\left(\frac{2+\alpha}{2-\delta}\right)$$

This indicates that the offspring of v_n typically lie with a radius $O(n^{-1/\alpha})$ of v_n, for large n. In other words, the whole of V_∞ is covered by $O(n)$ balls of radius $n^{-1/\alpha}$, indicating the Hausdorff dimension to be α. Rigorous arguments are found in [5].

3. Processes with Historical Memory

3.1. One-parameter Chinese Restaurant Process

Recall a specific Polya Urn Scheme which Pitman [10] elaborates as the Chinese Restaurant Process, with a parameter $\theta > 0$. Create an infinite sequence of independent Bernoulli trials, where the probability of success at the n-th trial is $\theta/(n+\theta)$. Start with an urn containing one ball, on which the number 0 is inscribed. After $n-1$ iterations of the rule we next describe, the urn will contain n balls, each bearing a number, and there will some integer $k \ge 1$ such that, for every $0 \le j \le k$, there is some ball bearing the number j. Here is the rule for the n-th step:

- If the n-th trial is a success, take a new ball, inscribe on it the number $k+1$, and place it in the urn.
- If the n-th trial is a failure, select a ball in the urn uniformly at random, note the number on it, and return it to the urn together with a new ball bearing the same number.

The essential property of the process, and which we wish to imitate, is that popular numbers become even more popular, as in the saying "the rich get richer".

3.2. Retro-preferential Processes

Take a finite set $V = \{v_0, v_1, \ldots, v_n\}$ of $n+1$ **points**, and a matrix $(P_{i,j})_{0 \le i,j \le n}$ where $i \neq j$ implies $P_{i,j} > 0$. The state v_0 is distinguished. We are going to create a random **trajectory** $(L_t)_{0 \le t \le \tau}$, where

$$L_t := (X_0, X_1, X_2, \ldots, X_t).$$

Here $X(t) \in V$ denotes the point visited at time t, $X_0 := v_0$, and τ is a terminal time. The construction will involve the vector of frequencies

$$F(t) := (F_0(t), F_1(t), \ldots, F_n(t))$$

where $F_j(t)$ counts the number of times $s \in \{0, 1, \ldots, t\}$ at which $X_s = v_j$. In other words

$$F_j(t) := \sum_{s=0}^{t} 1_{\{X_s = v_j\}}. \tag{9}$$

The **unvisited points** at time t consist of $\{j : F_j(t) = 0\}$. The **terminal time** τ occurs when the set of unvisited points is empty, i.e.

$$\tau := \min\{t : \min_j F_j(t) \ge 1\}.$$

The update rule depends on a sequence of functions

$$q_t : V^t \to [0,1], \quad t = 1, 2, 3, \ldots \tag{10}$$

called the **exploration rates**. Let $Z_1, Z_2, \ldots$ be a non-stationary Bernoulli process, where

$$\mathbf{P}[Z_t = 1 \mid L_{t-1}] := q_t(X_0, X_1, X_2, \ldots, X_{t-1}), \tag{11}$$

and Z_t is conditionally independent of $Z_1, Z_2, \ldots, Z_{t-1}$, given L_{t-1}. Here is the randomized update rule which extends L_{t-1} to L_t:

- If $Z_t = 1$ and $X(t) = v_j$, restrict the transition from v_j to **unvisited** points, meaning sample X_{t+1} from $\{v_k; F_k(t) = 0\}$, according to

$$\mathbf{P}[X_{t+1} = v_k] := \frac{P_{j,k} 1_{\{F_k(t)=0\}}}{\sum_{i:F_i(t)=0} P_{j,i}}.$$

- If $Z_t = 0$ and $X(t) = v_j$, restrict the transition from v_j to **visited** points, weighted by the number of previous visits, meaning sample X_{t+1} from $\{v_k; F_k(t) \ge 1\}$, according to

$$\mathbf{P}[X_{t+1} = v_k] := \frac{F_k(t) P_{j,k}}{\sum_i F_i(t) P_{j,i}}.$$

As a detail, note that if $q_1 = 1$, the list L_1 has the form (v_0, v_j) for some $j \neq 0$. If $P_{j,j} = 0$ for all j, then no two consecutive elements of L_t are identical.

3.3. Statistical Properties

Let Y_t denote the number of distinct points other than v_0 visited up to time t. In terms of (9),

$$Y_t := \sum_{s=1}^{t} Z_s = \sum_{i=1}^{n} 1_{\{F_i(t)\geq 1\}}.$$

The terminal time satisfies:

$$\mathbf{P}[\tau > t] = \mathbf{P}[Y_t < n] = \mathbf{P}[\sum_{s=1}^{t} Z_s \leq n-1].$$

In the simplest case where the sequence (q_t) does not depend on the trajectory, Y_t is a sum of independent Bernoulli random variables, with mean $\sum_{s=1}^{t} q_s$, and variance $\sum_{s=1}^{t} q_s(1-q_s)$. Estimates for $\mathbf{P}[\tau > t]$ may be derived from Chebyshev's inequality, or otherwise.

3.4. Examples of Exploration Rates Depending on One Parameter

Choose a **site exploration parameter** $\varphi > 0$.

3.4.1. Chinese Restaurant Process

If $P_{i,j} = 1$ for all $i \neq j$, and if $q_t = \varphi/(t-1+\varphi)$, the Retro-preferential Process coincides with the Chinese Restaurant Process, at least until the former reaches its terminal time. In this case, neglecting effects of $\tau < t$,

$$\mathbf{E}[Y_t] = \sum_{s=0}^{t-1} \frac{\varphi}{s+\varphi} \approx \varphi \log \frac{t+\varphi}{\varphi}. \tag{12}$$

3.4.2. Classic Retro-preferential Process

The **Classic Retro-preferential Process** (V, P, φ) is characterized by success rates:

$$\mathbf{P}[Z_t = 1 \mid L_{t-1}] := q_t := \frac{\varphi}{Y_{t-1}+\varphi}. \tag{13}$$

Observe $Y_0 = 0$, so $q_1 = 1$. Under the rule (13), the count $(Y_t)_{t\geq 0}$ of points visited coincides (up to the terminal time) with a pure birth Markov Chain, i.e. $Y_t - Y_{t-1} \in \{0, 1\}$, with $Y_0 = 0$ and

$$\mathbf{P}[Y_t = m+1 \mid Y_{t-1} = m] = \frac{\varphi}{m+\varphi}. \tag{14}$$

In the continuum limit where space and time are rescaled by a factor of $1/n$, using methods described in detail in Darling and Norris [4], the trajectory of the birth process is approximated by the solution of the ordinary differential equation

$$\frac{dy}{dt} = \frac{\varphi}{y+\varphi}; \quad y(0) = 0,$$

namely

$$y(t) = \sqrt{(\varphi(2t+\varphi))} - \varphi \tag{15}$$

as illustrated in Figure 3. This $t^{1/2}$ behavior should be contrasted with the $\log t$ behavior in (12). A precise non-asymptotic result is as follows.

Lemma 3.1. *The Pure Birth Markov Chain $(Y_t)_{t\geq 0}$ in (14) satisfies*

$$\mathbf{E}[(Y_t + \varphi - 0.5)^2] = 2\varphi t + (\varphi - 0.5)^2. \tag{16}$$

Furthermore the stopping time $\tau_m := \min\{t : Y_t = m\}$ has mean and variance as follows:

$$\mathbf{E}[\tau_m] = \frac{m(m+2\varphi-1)}{2\varphi}; \quad Var[\tau_m] = \frac{m(m-1)(2m-1+3\varphi)}{6\varphi}, \quad m = 1, 2, \ldots.$$

Remark: τ_m has a standard deviation $\approx m^{1.5}/\sqrt{3\varphi}$ and mean $\approx m^2/(2\varphi)$, for large m.

Proof. Consider the process

$$M_t := Y_t^2 + (2\varphi - 1)Y_t - 2\varphi t, \quad t = 0, 1, 2, \ldots. \tag{17}$$

Then

$$M_t - M_{t-1} = Y_t^2 - Y_{t-1}^2 + (2\varphi - 1)(Y_t - Y_{t-1}) - 2\varphi.$$

By (14),

$$\mathbf{E}[M_t - M_{t-1} \mid Y_{t-1} = m] = \frac{\varphi((m+1)^2 - m^2 + 2\varphi - 1)}{m + \varphi} - 2\varphi = 0.$$

Consequently $(M_t)_{t\geq 0}$ is an integrable martingale, and so

$$\mathbf{E}[M_t] = \mathbf{E}[M_0] = 0, \quad t = 0, 1, 2, \ldots.$$

Rewrite the last equation in terms of (17) to obtain (16). Optional stopping of $(M_t)_{t\geq 0}$ at τ_m gives

$$m^2 + (2\varphi - 1)m - 2\varphi\mathbf{E}[\tau_m] = 0$$

which gives the formula for $\mathbf{E}[\tau_m]$. As for the variance, we can write

$$\tau_m = G_0 + G_1 + \cdots + G_{m-1}$$

where the (G_j) are independent, and G_j has the Geometric distribution

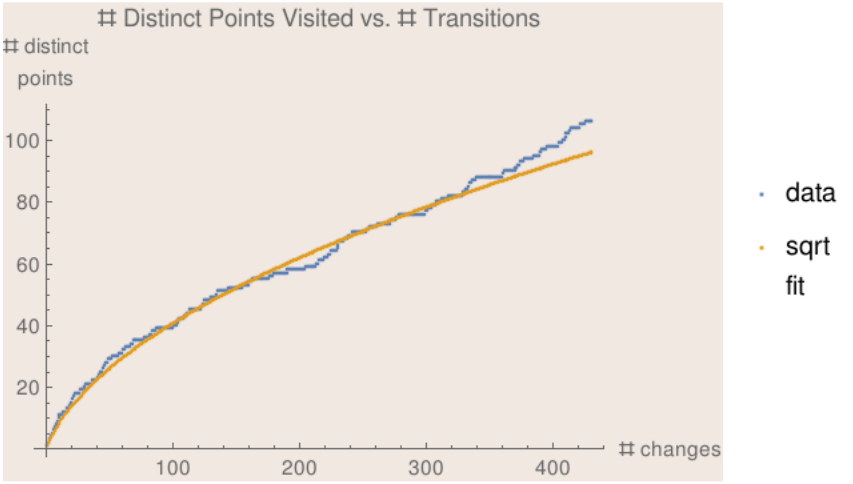
$$\mathbf{P}[G_j = k] = \frac{\varphi}{\varphi + j}\left(\frac{j}{\varphi + j}\right)^{k-1}, \quad k = 1, 2, \ldots.$$

The variance of G_j is $j(j+\varphi)/\varphi$. Invoke independence to obtain

$$\mathrm{Var}[\tau_m] = \sum_{j=0}^{m-1} \mathrm{Var}[G_j] = \sum_{j=0}^{m-1} j + \frac{1}{\varphi}\sum_{j=0}^{m-1} j^2 = \frac{m(m-1)}{2} + \frac{m(m-1)(2m-1)}{6\varphi},$$

which on simplification gives the desired result. □

Figure 3: *Birth Process vs. ODE Limit:* The pure birth process (14) is shown in the blue jagged curve, and its continuun limit (15) in the orange smooth curve.



3.5. Related Literature

It is surprising that a construction as simple as the Retro-preferential Process does not appear explicitly in the literature. Two close relatives are as follows:

- Pemantle [11] does not use the auxiliary process (Z_t). Instead he uses the transition probabilities
$$\mathbf{P}[X_{t+1} = v_k \mid X_t = v_j] := \frac{(1 + F_k(t))P_{j,k}}{\sum_i (1 + F_i(t))P_{j,i}}.$$
The asymptotic law of this vertex-reinforced random walk are described precisely in [11].

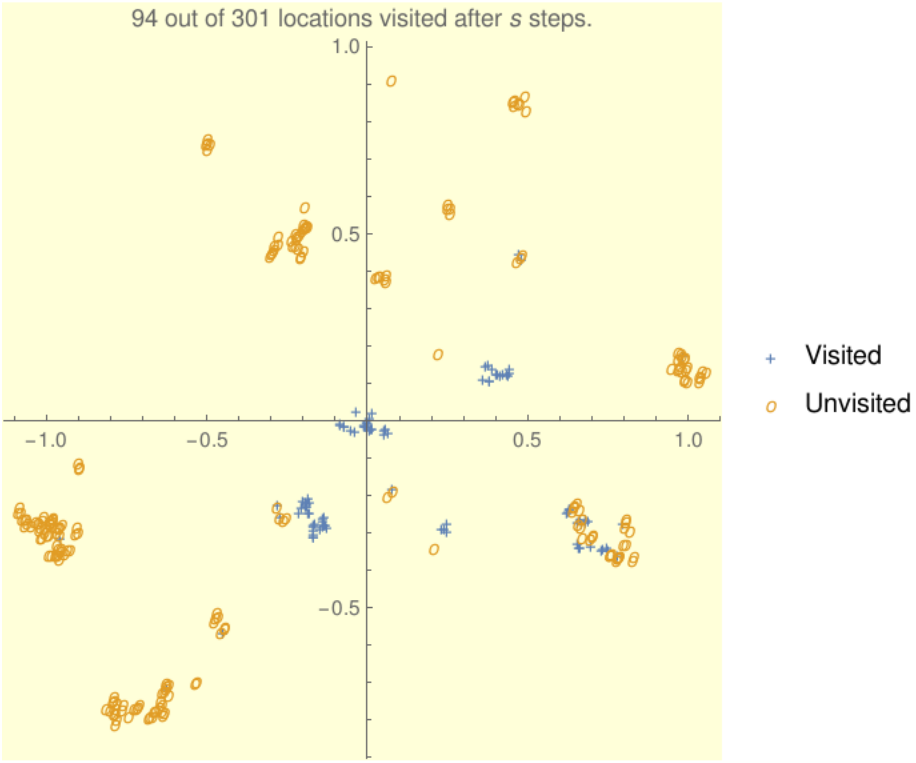
- Boyer and Solis-Salas [2] employ the auxiliary process (Z_t), but with constant $q_t := q$. The transition mechanism is the same as ours, but they limit attention to the graph with vertex set $\mathbf{Z}^d$, where $P_{j,k} = 1$ if v_j and v_k are adjacent sites in the lattice $\mathbf{Z}^d$ and $P_{j,k} = 0$ otherwise.

4. Retro-preferential Process on Spatial Points

4.1. Combining the Models

FRACTALRABBIT [3] uses a Classic Retro-preferential Process (13), on a set $V = \{v_0, v_1, \ldots, v_n\}$ of points in $\mathbf{R}^d$ generated as vertices of an Agoraphobic Point Process, where the rooted tree has edge set E. The default choice of matrix $(P_{i,j})_{0 \le i,j \le n}$ with positive

Figure 4: **Retro-preferential Random Trajectory:** *300 points were generated in the unit ball in* $\mathbf{R}^2$ *using the Agoraphobic Point Process* ($\theta = 2.0$, $1/\alpha = 0.8$)*. A Classic Retro-preferential Process* ($\varphi = 13.75$) *ran for 430 steps on these points, with transition matrix given by (18). 207 points remained unvisited. The 94 visited points tended to be clumped together.*



off-diagonal entries is the one where $P_{i,i} = 0$, and

$$P_{i,j} = \frac{1}{\|v_i - v_j\|^d}, \quad j \neq i. \tag{18}$$

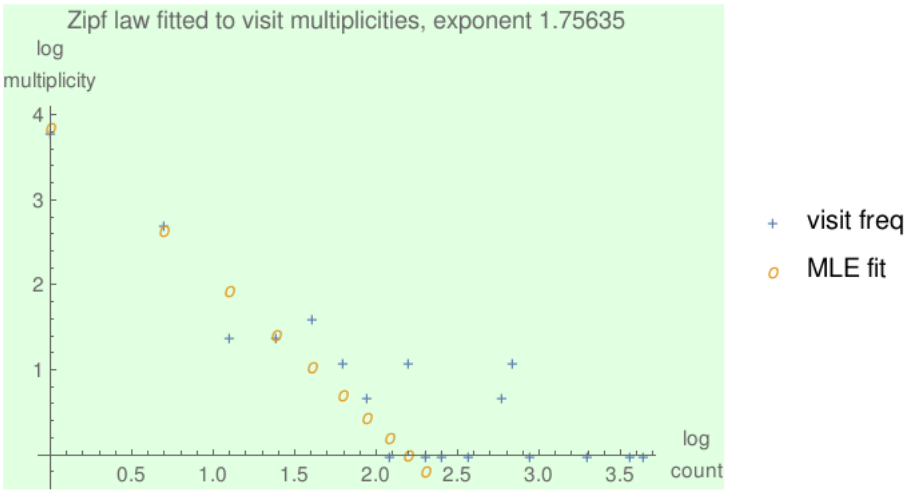
The choice of d in the exponent was made on the basis of experiments with $d = 2$. It strongly favors transitions to nearby points. This is not the only choice of P. For example $1/P_{i,j}$ could be taken as the number of edges along the shortest path from v_i to v_j in the tree (V, E), or as the sum of edge lengths along this path.

Figure 4 shows an example of a Retro-preferential Process on a realization of an Agoraphobic Point Process.

4.2. Power Law Observed in Visit Frequencies

Suppose $(X_0, X_1, X_2, \ldots, X_t)$ is a trajectory of a Retro-preferential Process on a realization $V = \{v_0, v_1, \ldots, v_n\}$ of an Agoraphobic Point Process. At the conclusion, point v_j has

Figure 5: **Power Law of Frequencies of Visits:** *The 431 point visits shown in Figure 4 are broken into 94 categories, according to which vertex was visited. For example, one point was visited 45 times, i.e. multiplicity 45, but eight of the points were visited only once. Logs of these counts are shown on the x axis, and logs of the count frequencies on the y axis. For example* $(0, log45)$ *is the leftmost point. This log-log plot suggests a power law, to which the Zipf distribution was fitted by maximum likelihood.*



been visited $F_j(t)$ times, as in (9). Summarize these statistics by defining

$$d_k := |\{v_j : F_j(t) = k\}|.$$

For example, in the example of Figure 5, only one point was visited 45 times, so $d_{45} = 1$, but eight of the points were visited just once, so $d_1 = 8$. Figure 5 shows a plot of

$$\{(\log k, \log d_k) : k \geq 1, d_k \geq 1\}. \tag{19}$$

In all experiments we performed, log-log plot displayed linear features typical of power law distributions. A Zipf distribution of the form

$$\mathbf{P}[Z = m] \propto m^{-s}, \quad m = 1, 2, 3, \ldots$$

was fitted by maximum likelihood, and is plotted for comparison. The fitted exponent s is shown.

Open Problem: Consider a sequence $(V^{(n)}, P^{(n)}, \varphi)_{n\geq 1}$ of Classic Retro-preferential Processes, where $|V^{(n)}| \to \infty$, and the row sums of the matrices $P^{(n)}$ are uniformly bounded. Assume (a) φ is fixed, (b) $V^{(n)} \subset V^{(n+1)}$, and (c) $P^{(n+1)}$ agrees with $P^{(n)}$ on a pair of vertices which are both in $V^{(n)}$. In this context can one prove rigorously the existence of power laws in vertex visit frequencies (19)?

Sojourn j	0	1	2	3	4	5	6	7	8	9
Zone x_j	a	b	a	c	d	a	c	a	e	a
Count R_j	2	3	4	1	1	4	0	5	4	1
Reports $(T_j^k)_{k\geq 0}$	0.1 0.12	0.17 0.17 0.21	0.4 0.42 0.42 0.5	0.7	1.2	1.3 1.3 1.5 1.51		1.8 1.9 1.9 1.9 1.98	2.1 2.25 2.27 2.27	3.0

Table 1: *Sojourns and sporadic reports:* Observe that sojourn 6 generated no reports, and so sojourns 5 and 7 appear to give nine consecutive reports at the same zone. Only six distinct times are represented among the nine reports during sojourns 5 - 7.

5. Sporadic Report Model given a Motion Sequence

The remaining sections of this paper do not contain mathematical results, but specify a reporting model which is implemented in FRACTALRABBIT [3].

5.1. Structure of the Report Model

Suppose $(x_0, x_1, \ldots, x_n)$ is a sequence of points in a finite metric space (V, ρ), where $\rho(x_{i-1}, x_i) > 0$ for each i. The time taken to travel from x_{i-1} to x_i is assigned a lower bound of $\rho(x_{i-1}, x_i)/\sigma_*$, where σ_* is a **maximum speed**.

In applications, $(x_0, x_1, \ldots, x_n)$ could form the trajectory of a Retro-preferential Process on a realization of an Agoraphobic Point Process. However the report model we shall now develop is conceptually separate from the model which generates the trajectory.

In Section 4 each x_i was regarded as a point. Once we shift to the reporting model in continuous time, one may think of x_i as the label of a zone. The trajectory may be viewed as a sequence of $n+1$ sojourns by a target, where sojourn j is spent in zone x_j. During a sojourn, some random number (possibly zero) of reports will be created, at randomly spaced times, indicating the label of the zone containing the target. The random length of a sojourn will be generated as a sum of times between reports.

The zones, labelled by the elements of V, are not necessarily disjoint, nor need they cover the space of target motion. If we study the target during the real time interval $[0, T]$, the sojourns are disjoint subintervals of $[0, T]$, but need not cover the whole of $[0, T]$. Position of the target between consecutive sojourns may be unspecified and unreported. As mentioned already, some sojourns may generate no reports.

5.2. Event Times in the Report Model

We generate two sequences $(Y_i)_{0\leq i\leq n}$ and $(Z_i)_{1\leq i\leq n}$ of non-negative, mutually independent random variables. Here Y_i denotes the duration of a sojourn at x_i, and Z_i denotes the

duration of travel from x_{i-1} to x_i via some unspecified route. Assign to Z_i a distribution whose minimum is $\rho(x_{i-1}, x_i)/\sigma_*$.

Suppose $T_0 := 0$ is the time of arrival at x_0. The unobserved time of arrival at x_j is

$$T_j := Y_0 + (Z_1 + Y_1) + (Z_2 + Y_2) + \cdots + (Z_{j-1} + Y_{j-1}) + Z_j \tag{20}$$

and the unobserved time of departure from x_j is

$$T'_j := T_j + Y_j. \tag{21}$$

Suppose $R \geq 0$ is an integer random variable, and $R(0), R(1), \ldots, R(n)$ are independent copies of R. Interpret $R(j)$ as the number of reports which occur at point x_j of the trajectory, that is during the interval $[T_j, T'_j]$. If $\mathbf{P}[R = 0] > 0$, there may be values of j for which the sojourn at x_j goes unreported, as in Table 1. Also suppose $U \geq 0$ is a real random variable, and $(U_j^k)_{k\geq 0}$ are independent copies of U. If $R(j) = 0$, then $Y_j = U_j^0$. If $R(j) \geq 1$, then

$$Y_j = U_j^0 + U_j^1 + U_j^2 + \cdots + U_j^{R(j)}. \tag{22}$$

If $R(j) \geq 1$, reports indicating zone x_j occur at times

$$T_j^k := T_j + U_j^0 + U_j^1 + \cdots + U_j^k, \quad k = 0, 1, 2, \ldots, R(j) - 1. \tag{23}$$

In other words U models the gaps

1. Between arrival and the first report (or departure, if there are no reports),
2. Between successive reports at the same zone of the trajectory, if there are two or more reports, and
3. Between the last report and departure, if there is at least one report.

It could happen that $U_j^k = 0$ for all $k = 0, 1, 2, \ldots, R(j) - 1$, giving $R(j)$ simultaneous reports, and $Y_j = 0$.

5.3. Partition of Time into (Reported) Sojourns and (Unreported) Motions

To summarize: the total elapsed time $[T_0, T'_n]$ may be partitioned into two sets.

1. The $(n + 1)$ intervals $\{[T_j, T'_j], 0 \leq j \leq n\}$, during which reports are received about a sojourn at some x_j, have total length

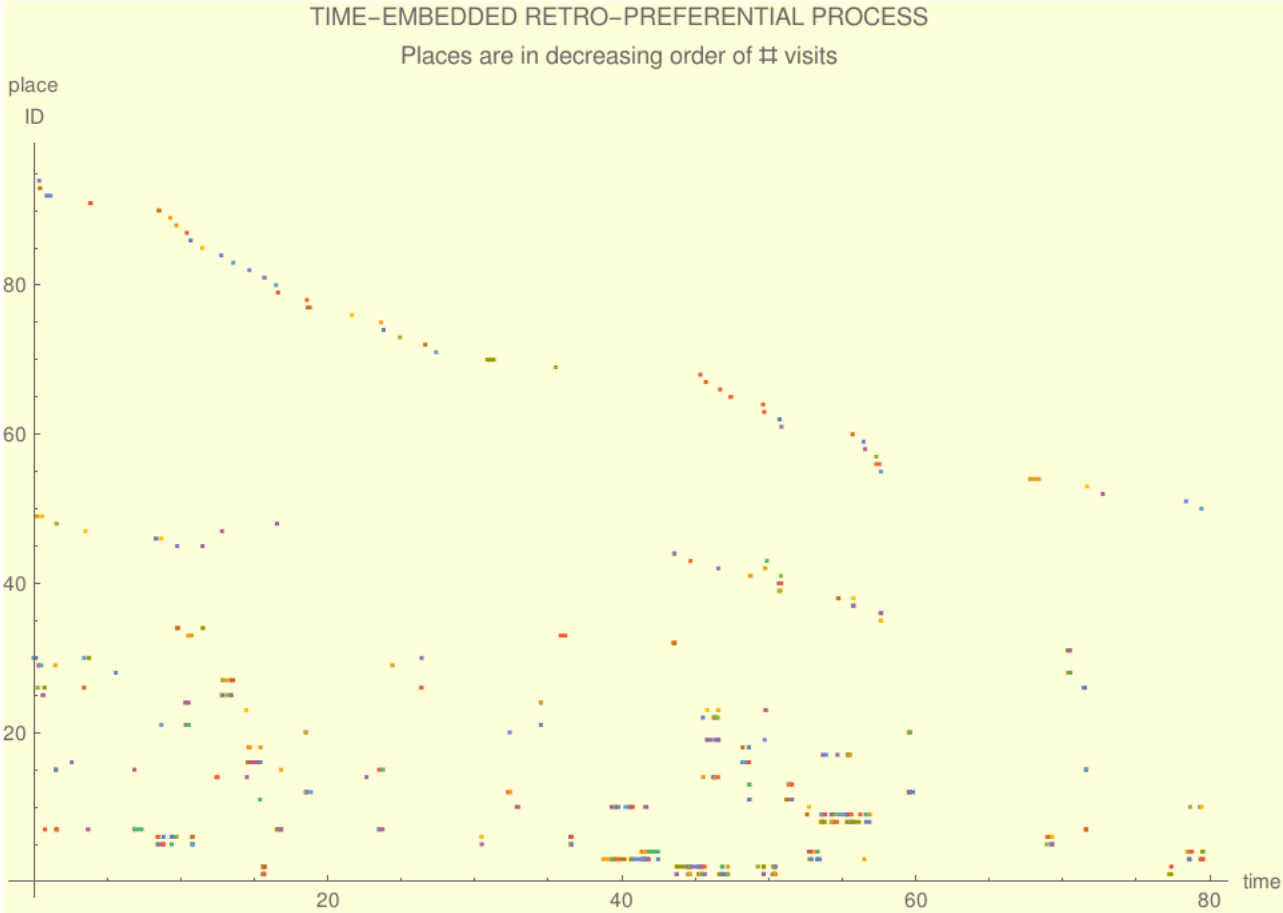
$$\sum_0^n Y_j = \sum_{j=0}^n (U_j^0 + U_j^1 + \cdots + U_j^{R(j)}).$$

   For j such that $R(j) \geq 1$, report times are

$$\{T_j^k : 0 \leq k \leq R(j) - 1\}$$

   and the reports take the form: "target is in zone x_j at time T_j^k". Such a continuous time embedding of the Retro-preferential Process is illustrated in Figures 6 and 7.

Figure 6: **Continuous Time Embedding of the Retro-preferential Process** *The 431 point visits shown in Figure 4, are embedded into eighty time units after sporadic report generation. The vertical axis lists the 94 places visited in decreasing order of frequency. Times when these places were visited are shown on the horizonatal axis.*



2. The n intervals $(T'_{j-1}, T_j)_{1\le j\le n}$, during which motion occurs, but no reports are received, have total length

$$\sum_1^n Z_j.$$

The distribution of the ratio

$$\frac{\sum_0^n Y_j}{\sum_1^n Z_j} \tag{24}$$

provides a check on whether model parameters are suitably chosen. Bear in mind that neither numerator nor denominator need have finite mean.

5.4. Examples of Heavy-tailed Distributions

Heavy-tailed distributions abound in the data we seek to model, as described in [1], [8]. Without truncation, these distributions would have infinite first moment. In practice these distributions are truncated to a suitable range of validity, depending on the application. The truncation parameters will be denoted c, c', c'', respectively. Typical choices for the laws of Z_i, R and U include:

- Assign to Z_i a truncated Pareto Distribution:

$$\mathbf{P}[Z_i \le t] = \frac{1-(t_i/t)^\alpha}{1-(1/c')^\alpha}, \quad t_i < t \le c' t_i; \quad t_i := \frac{\rho(x_{i-1}, x_i)}{\sigma_*}.$$

  If a reported value $z_i \ge t_i$ is available for $1 \le i \le n$, then in the limit as $c' \to \infty$ the maximum likelihood estimate of α is

$$\hat{\alpha} := \frac{n}{\sum_1^n \log(z_i/t_i)}.$$

- Assign to R a mixture of a mass $1-\eta \in [0,1)$ at zero, and a truncated Zipf law

$$\mathbf{P}[R = m] = \eta \frac{m^{-s}}{\zeta_c(s)}, \quad m = 1, 2, 3, \ldots, c,$$

  where ζ_c denotes the first c terms of the series for the Riemann Zeta Function. Given a sample $\{r_1, r_2, \ldots, r_n\}$ of reported values of R conditioned on $R \ge 1$, the maximum likelihood estimate of s satisfies

$$\frac{\zeta_c'(s)}{\zeta_c(s)} = -\frac{1}{n}\sum_1^n \log r_i.$$

- Assign to U a mixture of a mass $1-\delta \in [0,1)$ at zero, and a truncated Pareto Distribution:

$$\mathbf{P}[U \le u] = 1 - \delta + \delta \frac{1-(u_o/u)^\beta}{1-(1/c'')^\beta}, \quad u_o < u \le c'' u_o$$

where u_o is a minimal positive time increment. This implies that the probability of a cluster of $k \geq 1$ simultaneous reports is $(1-\delta)^{k-1}\delta$, i.e. the Geometric distribution which describes the number of Bernoulli(δ) trials up the first success. Given successive clusters of size $k_1, k_2, \ldots, k_r$, the maximum likelihood estimate of δ is

$$\frac{r}{\sum_{i=1}^{r} k_i}.$$

Given a sample $\{u_1, u_2, \ldots, u_m\}$ of reported values of U conditioned on $U > 0$, in the limit as $c'' \to \infty$ the maximum likelihood estimate of β is

$$\hat{\beta} := \frac{m}{\sum_{1}^{m} \log(u_i/u_o)}.$$

6. Simulating a Retro-preferential Process with Sporadic Reports

FRACTALRABBIT [3] combines constructions from Sections 2.2, 3.2, and 5 to present a candidate for a motion and reporting model. For simplicity take $d = 2$.

6.1. Context-Specific Choices for the Motion Model

The user must decide at the outset on approximate values of integers k, n, h as follows.

A. The number k of spatial points which the target *could* possibly visit. This means $V = \{v_0, v_1, \ldots, v_{k-1}\}$ will be a subset of the unit ball of $\mathbf{R}^2$.

B. The number n of points along the trajectory $(x_0, x_1, \ldots, x_n)$ through V, with the convention $x_0 = v_0$. Note that n could be larger than k because the same point could be visited multiple times. In our model, $x_i \neq x_{i+1}$.

C. The number h of *distinct* points of V which the trajectory visits. In other words, there are $k - h$ unvisited points in V.

6.2. Motion Model Implementation

1. Construct a set $V = \{v_0, v_1, \ldots, v_{k-1}\}$ in the unit ball of $\mathbf{R}^2$ using, for example, the Hard Threshold Agoraphobic Point Process with innovation parameter $\theta > 0$ and exponent $\alpha < 2$, as illustrated in Figure 2.

2. Define the Retro-preferential Process exploration parameter φ to be $h^2/(2(n-h))$, where h estimates distinct places visited during the n steps of the trajectory, to match the mean behavior (15)
$$h = \sqrt{(\varphi(2n+\varphi))} - \varphi.$$

3. Use the matrix (18) and the probabilities (13) to create a Retro-preferential Process on V, with trajectory $(x_0, x_1, \ldots, x_n)$, with $x_0 = v_0$.

Example: Figures 2 and 4 illustrate $\theta = 2.0$, $1/\alpha = 0.8$, $k = 301$, $n = 430$, $\varphi = 13.75$.

6.3. Context-Specific Choices for the Reporting Model

Suppose that probability distributions for the reporting model are chosen as in Section 5.4, where we also indicate how to fit some model parameters by maximum likelihood.

D. For the motion intervals, choose a maximum speed σ_*. Set a value for the Pareto parameter for motion intervals; possible defaults are $\alpha = 0.55$, and truncation parameter $c' = 10$.

E. Set the probability $1 - \eta$ that a visit to a point provokes no reports, and the Zipf parameter s for the number of reports, if greater than 1. Possible default values are $s = 1.5$, and truncation parameter $c = 50$.

F. Decide on a minimum discrete time unit u_o (for example, one second). Set the probability $1 - \delta$ that a report coincides in time with the previous report, and the Pareto parameter β for the intervals between reports at the same location. Possible default values are $\beta = 0.287$, and truncation parameter c'' as the number of seconds in six hours.

6.4. Reporting Model Implementation

Assume that the steps in Section 6.2 have been carried out. In particular, a trajectory $(x_0, x_1, \ldots, x_n)$ has been created.

4. Compute the distances $\|x_i - x_{i-1}\|$, and choose a maximum speed σ_*, to establish the minimum t_i for Z_i. Generate independent truncated Pareto[2] random variables $(Z_i)_{1 \le i \le n}$ as in Section 5.4, using the parameters σ_*, α, and c'.

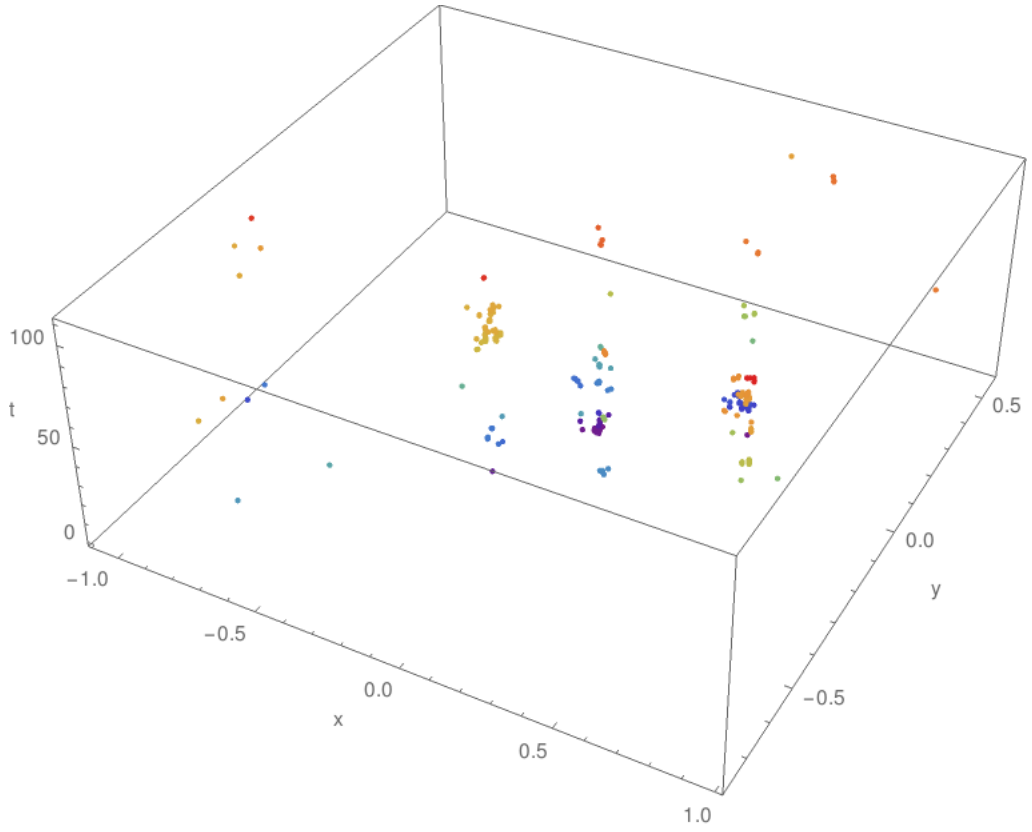
5. To generate i.i.d. integer random variables $R(0), R(1), \ldots, R(n)$, perform a Bernoulli(η) trial to decide whether $R(i) \ge 1$, and if so generate a Zipf[3] random variable with exponent s, for each i.

6. When $R(j) = k$, generate independent Pareto(u_o, β) random variables $U_j^0, U_j^1, \ldots, U_j^k$, and define sojourn time Y_j as in (22). All the $(Y_i)_{0 \le i \le n}$ and $(Z_i)_{1 \le i \le n}$ are available, so we may compute the arrival times (T_j) as in (20), and report times (T_j^k) as in (23).

At the conclusion, a total of $\sum_{j=0}^{n} R(j)$ reports have been generated, each of the form: "target is in zone x_j at time T_j^k".

[2]Take independent Uniform$(0,1)$ random variables (ξ_i), and set $Z_i := t_i(1 - q\xi_i)^{-1/\alpha}$, where $q = 1 - c'^{-\alpha}$.
[3]Multinomial sampling with a suitable upper bound will suffice.

Figure 7: **Three-Dimensional Embedding of the Retro-preferential Process** *The 431 point visits shown in Figure 4, are shown with spatial* (x, y) *coordinates for places visited, and a vertical time coordinate for the time of the visit.*



6.5. Continuous Motion Around Finite Range Sensors

We may interpret the report "target is in zone x_j at time T_j^k" as follows: suppose point x_j is the location of a sensor with finite range r_j, and the report means "the target is within distance r_j of point x_j at time T_j^k". In the sensor interpretation, the possible trajectories $t \mapsto y(t)$ of the target, consistent with the reports, satisfy

$$\|y(T_j^k) - x_j\| \leq r_j, \quad \forall j, k.$$

One option is to fix $r_* > 0$, and take $r_j := r_*$ for all j.

6.6. Seasonality Adjustment

Introduce further parameters

G. D is the number of days over which the series of reports should extend.

H. Pick shape parameters $\alpha' > 0, \beta' > 0$ for a Beta(α', β') distribution whose cumulative distribution function $F : [0,1] \to [0,1]$ can be used in daily seasonality adjustments. For example, $\alpha' = 4.17$, $\beta' = 2.98$ are typical of daylight animals.

Recall that the final report time (21) occurs at time T'_n. To introduce daily seasonality into the reports (T_j^k):

7. Rescale to span a duration of D days, by taking

$$\tilde{T}_j^k := \frac{DT_j^k}{T'_n}.$$

Thus the day on which afore-mentioned report occurs is the floor $\lfloor \tilde{T}_j^k \rfloor$.

8. Introduce daily seasonality by changing the report time (in days) to

$$\hat{T}_j^k := \lfloor \tilde{T}_j^k \rfloor + F^{-1}(\tilde{T}_j^k - \lfloor \tilde{T}_j^k \rfloor).$$

Acknowledgments: The author thanks Robin Pemantle (U. Penn.) for several key insights, especially into Euclidean embedding of random trees; Dylan Molho (Michigan State U.) for data explorations, and for coding a Python version of FRACTALRABBIT; and Cliff Joslyn (Pacific Northwest NL) for motivating this work.

References

[1] Albert-László Barabási, The origin of bursts and heavy tails in human dynamics, Nature, Vol. 435, 207-211, 12 May 2005

[2] D. Boyer, C. Solis-Salas, Random walks with preferential relocations to places visited in the past and their application to biology, Phy. Rev. Letters, 112, 240601, 2014

[3] R. W. R. Darling, FRACTALRABBIT (Mathematica™, Java™, Python™), *National Security Agency:* Stochastic Models and Simulation Tools, `https:github.com`, 2018

[4] R. W. R. Darling & J. R. Norris, Differential equation approximation for Markov chains, Probability Surveys 5, 37 - 70, 2008

[5] R. W. R. Darling & Robin Pemantle, Euclidean embedding of the Poisson weighted infinite tree and application to mobility models, In Preparation, 2018

[6] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*, Springer, New York, 1995

[7] A. Fujihara & H. Miwa, Homesick Lévy walk: a mobility model having Ichi-go Ichi-e and scale-free properties of human encounters, IEEE 38th Annual Computer Software & Applications Conf., 2014.

[8] Marta C. González, Cesar A. Hidalgo & Albert-László Barabási, Understanding individual human mobility patterns, Nature Vol 453, 779-782 5 June 2008

[9] National Land Cover Database `https://www.mrlc.gov/nlcd2011.php`, *embedded graphic*, 2011

[10] Jim Pitman, *Combinatorial Stochastic Processes*, École d'Été de Probabilités de Saint-Flour XXXII, Springer Verlag, Berin, 2002

[11] Robin Pemantle, Vertex-reinforced random walk, Probab. Theory Relat. Fields 92, 117-136, 1992