Efficient search by optimized intermittent random walks

2009 J. Phys. A: Math. Theor. 42 434008

(http://iopscience.iop.org/1751-8121/42/43/434008)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **42** (2009) 434008 (14pp) doi:10.1088/1751-8113/42/43/434008

Efficient search by optimized intermittent random walks

Gleb Oshanin[1,2], **Katja Lindenberg**[3], **Horacio S Wio**[4] **and Sergei Burlatsky**[5]

[1] Laboratory J-V Poncelet (UMI CNRS 2615), Independent University of Moscow, Bolshoy Vlasyevskiy Pereulok 11, 119002 Moscow, Russia
[2] Laboratoire de Physique Thèorique de la Matière Condensée (UMR 7600), Université Pierre et Marie Curie/CNRS, 4 place Jussieu, 75252 Paris Cedex 5, France
[3] Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0340, USA
[4] Instituto de Física de Cantabria, Avda Los Castros s/n, E-39005 Santander, Spain
[5] United Technologies Research Center, United Technologies Corp., 411 Silver Lane, 129-21 East Hartford, CT 06108, USA

E-mail: oshanin@lptmc.jussieu.fr, klindenberg@ucsd.edu, wio@ifca.unican.es and burlatsf@utrc.utc.com

Received 28 January 2009, in final form 27 April 2009
Published 13 October 2009
Online at stacks.iop.org/JPhysA/42/434008

Abstract

We study the kinetics for the search of an immobile target by randomly moving searchers that detect it only upon encounter. The searchers perform intermittent random walks on a one-dimensional lattice. Each searcher can step on a nearest neighbor site with probability α or go off lattice with probability $1-\alpha$ to move in a random direction until it lands back on the lattice at a fixed distance L away from the departure point. Considering α and L as optimization parameters, we seek to enhance the chances of successful detection by minimizing the probability P_N that the target remains undetected up to the maximal search time N. We show that even in this simple model, a number of very efficient search strategies can lead to a decrease of P_N by orders of magnitude upon appropriate choices of α and L. We demonstrate that, in general, such optimal intermittent strategies are much more efficient than Brownian searches and are as efficient as search algorithms based on random walks with heavy-tailed Cauchy jump-length distributions. In addition, such intermittent strategies appear to be more advantageous than Lévy-based ones in that they lead to more thorough exploration of visited regions in space and thus lend themselves to parallelization of the search processes.

PACS numbers: 05.40.−a, 87.23.−n

1. Introduction

Search processes are ubiquitous in nature: predators search for prey, prey also hunt, molecules search for each other to recombine in order to produce required chemicals, proteins search for target sequences on DNAs. Human beings spend their lives searching for different things—better jobs, shelters, partners, lost keys; they also seek efficient search strategies to minimize the time to reach the desired target or at least to enhance their chances of eventually finding it.

The search for a desired target may depend on a variety of different conditions and may take place in different environments. Targets may be sparse, hidden, difficult to detect even when found. The targets may be mobile or immobile, they may try to avoid searchers, there may be one target or many. They may have a finite life-time and vanish before they are detected. Searchers, on the other hand, may search 'blindly', detecting the target only upon encounter or they may perceive distant targets and adjust their motion accordingly. They may have no memory of previously visited areas or they may avoid such areas. The searchers may act individually or in swarms, optimizing their search efficiency by exchanging information. Finally, the 'efficiency' of a search may be judged by a variety of measures, including the time to reach a target or targets, the number of encounters of searchers and targets per unit time or the exploration range of space per unit time. In general, for each specific situation different search strategies may be appropriate. The quest for optimal strategies has motivated a great deal of work in the past few years [1–12, 13–27].

Earlier work tended to focus on deterministic search algorithms (see, e.g., [1–4] and references therein) specific to human activities such as, say, rescue operations or the search for natural resources. More recent studies have focused on random search strategies. In this context, it has become quite clear that strategies based on Lévy flights (instantaneous) or Lévy walks (occurring with a finite velocity) are according to all measures more advantageous than strategies based on conventional diffusive motion or random walks with steps to nearest neighbors only (Brownian search) [5–12]. Lévy searchers perform excursions whose lengths l are random variables with heavy-tailed distributions $p(l)$ of the form

$$p(l) \sim \frac{B}{l^{\mu+1}}, \qquad 0 < \mu < 2, \tag{1}$$

where B is a normalization constant. In particular, it was demonstrated in [28] that using the Cauchy distribution of jump lengths as given in equation (1) with $\mu = 1$ instead of a Gaussian distribution (Brownian search) allows for a faster cooling scheme, and hence for a considerable reduction of computer time in the search for a global minimum in nonconvex (multiple extrema) energy landscapes by simulated annealing. Aside from this, extensive data have been presented allegedly supporting the idea that many of the species in the living world do indeed follow Lévy-type random motion in their search [5–12]. This point, however, has been questioned recently in [29, 30]. The main objections of [29, 30] have recently been re-examined in [12], where it was concluded that Lévy-based strategies may still be consistent with experimental observations if one takes into account a highly non-homogeneous spatial distribution of prey. More sophisticated models with adaptive behavior in which the foragers use their cognitive skills to develop more efficient foraging strategies have also started to appear in the literature (see, e.g., [31]).

Following the observation of trajectories of foraging animals such as lizards, fish or birds, in which active local search phases randomly alternate with relocation phases (see, e.g., [13, 14]), another type of random search—an intermittent search—has been proposed. In this algorithm, a search is characterized by two distinct types of motions, a ballistic relocation stage during which the searcher is non-receptive to the target and a relatively slow phase with random Brownian-type motion during which the target may be detected [15–18, 20] (see also

[19]). In this approach, one aims to minimize the time of the first passage to the target from a given location by varying, for example, the relative durations of the active and relocation stages.

In this paper, we pursue the optimization of an intermittent strategy. A simpler version of our model was presented in [20], where we developed a search algorithm based on intermittent random walks that involve nearest neighbor steps and off-lattice relocations of a fixed length L. There, instead of minimizing the first passage time to the target, we sought to maximize the success of the search by minimizing the target non-detection probability *over a fixed finite maximum search time*. We note that the time derivative of this probability defines the distribution function of the first passage time to the detection event. Thus, contrary to previous work in which only the mean first passage time was optimized [15–18, 20], our goal was to optimize the full distribution function. Our optimization parameter was the intermittency parameter α that determines whether the next step will be on- or off-lattice. It was shown, both analytically and numerically, that the probability of failure to detect the target over a finite search time can be made smaller by many orders of magnitude upon an appropriate choice of this parameter. We note that in [21] a different intermittent search algorithm was proposed in which the length of the relocation stage is a random variable with a heavy-tailed distribution in equation (1). In that work, it was concluded that such a combined strategy is advantageous over intermittent searches with exponentially distributed [15–17] or fixed [20] relocation lengths since it allows a searcher to find the target more quickly in the critical case of rare targets, and since the search performance is much less dependent on adaptation to the target density. However, we argue that in some (albeit not all) situations, strategies involving Lévy-distributed relocations cannot be optimal. This occurs when there is some maximal time that the search is allowed to run. Allowing the length of the relocation stage (and consequently, the time spent on each relocation tour) to have a heavy-tailed distribution would lead to some portion of the search process that would involve unnecessarily long relocations divorced from the targets, thus not contributing to the overall finite time effort. Thus, when appropriate, an optimal search algorithm should be based on relocation length distributions that explicitly account for the fact that a search process is limited in time. It may well be that the optimal jump-length distribution should itself vary with time. Furthermore, trajectories of Lévy walks or flights are 'overstretched' in the sense that such walks explore space in a very irregular manner. The visited area consists of a patchy set of disconnected clusters, leaving large unexplored voids compared to the case of a Brownian search. Additionally, when many Lévy searchers are involved, the fact that a Lévy distribution does not have moments induces rapid mixing. This mixing might be advantageous if the detection probability is low (or the false alarm probability high), such that multiple rechecking of each site by other searchers is required in order to spot the target. Otherwise, this very efficient mixing might be a disadvantage since it does not favor the parallelization of the search process by dividing the searched area into subunits and assigning a separate domain to each searcher. In the 'living' world, animals are often constrained to their assigned territories, and even an occasional incursion into a neighbor's terrain while searching for prey may cause serious difficulties.

In this paper, we revisit the question of an optimal jump-length distribution underlying an efficient search algorithm. Focusing on the one-dimensional case (for which Lévy-based search strategies are said to most dramatically outperform intermittent ones), we study the search kinetics of a 'hidden' immobile target located at the origin of an infinite lattice by a concentration ρ of randomly moving searchers. The motion of the searchers is intermittent, consisting of two distinct, randomly alternating stages—ballistic, off-lattice relocations with finite velocity over a fixed distance L and random walks between nearest neighboring sites.

In other words, we consider a search by intermittent random walkers with a jump-length distribution of the form (very different from that in equation (1))

$$p(l) = \frac{\alpha}{2}\left[\delta(l-1) + \delta(l+1)\right] + \frac{(1-\alpha)}{2}\left[\delta(l-L) + \delta(l+L)\right], \qquad (2)$$

such that the searchers step on nearest neighbors with probability α and, with probability $1-\alpha$, perform long jumps over a distance L [20]. The process evolves in discrete time $n = 0, 1, 2, \ldots, N$, where N is the maximal time the search process may run. This time may depend, for instance, on our patience or on experimental constraints. Note that the constraint of the maximal search time N is the crucial aspect of our work which makes our analysis very different from other models. Steps to nearest neighbors take unit time, while off-lattice relocations over a distance L require time T. The term 'hidden' means that a searcher cannot perceive the target when off-lattice. A searcher only detects the target when it arrives at the site on which the target is located.

We pose the following question: is it possible to choose $\alpha = \alpha_N$ and $L = L_N$, dependent on the maximal search time N but independent of the running time n, which optimizes the search efficiency and leads to a performance that is better than Lévy-based strategies? In order to answer our question, our first goal is to calculate the probability P_N that the target remains undetected up to the maximal search time N and to determine its asymptotic behavior analytically in the large-N limit. Then, considering $\alpha = \alpha_N$ and $L = L_N$ as optimization parameters, we seek to enhance the searchers' chances of success by minimizing P_N. We will demonstrate that, depending on whether we are at liberty to tune α (as in [20]) or both α and L, different optimal strategies can be realized all of which can decrease the value of P_N by many orders of magnitude compared to a Brownian search. We also show that even the simple distribution (2) with optimal $\alpha = \alpha_N$ and $L = L_N$ yields very efficient search algorithms comparable to and in some cases better than Lévy-based strategies. Moreover, in striking contrast to the latter, optimal intermittent walks lead to much denser exploration of space. These results support our claim that Lévy-based searches are not the best algorithms when the search is limited in time.

2. Model and basic equations

Consider a one-dimensional regular lattice of unit spacing containing M sites labeled by s. The lattice is a circle, that is, we use periodic boundary conditions. At one of the lattice sites, say at the origin $s = 0$, we place an immobile target. Then we randomly place K searchers under the constraint that none is placed at the site of the hidden target. We focus on the behavior in the thermodynamic limit, $K \to \infty$, $M \to \infty$, with a finite mean density of searchers $\rho = K/M$. We note that the model under study can also be solved in the general case of finite K and M, but the calculations become more involved without adding significant new features to our conclusions.

Next, we allow the searchers to move according to the following rule. At each tick of the clock, $n = 1, 2, 3, \ldots, N$, each searcher selects randomly between two possibilities: with probability α, it moves with equal likelihood to one of its nearest neighboring sites and with probability $(1-\alpha)$ it leaves the lattice and flies off-lattice with a given velocity V until it lands L sites away from the departure site. The distance L is fixed, but either direction of the flight is chosen at random with equal probabilities. The time it spends off-lattice during the flight is $T = L/V$. This parameter can take integer values, $T = 1, 2, \ldots, L$. Note that this condition defines the velocity V. There is no restriction on multiple searcher occupancy of the sites. Note as well that in the two 'pure' cases, $\alpha = 1$ and $\alpha = 0$, the model reduces to standard random walks. Here we take both α and L to be independent of n but possibly dependent on the maximal search time N. An interesting situation with time-dependent $\alpha = \alpha_n$ and $L = L_n$ will be discussed elsewhere [33]. We focus on perfect detection, that is, a searcher recognizes the target immediately upon first contact. The case of imperfect recognition can be solved using the same techniques and will also be discussed elsewhere [33].

The probability that the target has not been detected by step N is related in a simple way to S_N, the *number of distinct sites* visited up to that time, as

$$P_N = \exp\left(-\rho S_N\right) \tag{3}$$

(see, e.g., [35]). This result, which holds for independent searchers and was explicitly shown to be valid for our model in [20], is a crucial equation since we will arrive at results for P_N via calculations of S_N. A larger S_N leads to a smaller probability that the target remains undetected and thus to a better search algorithm. In general, $S_N \sim AN^\gamma$ with $0 < \gamma \leqslant 1$. One thus expects that larger γ leads to a more efficient search and explains why, intuitively, it was believed that the most efficient search algorithm should be based on Lévy walks with a broad distribution of jump lengths, for which S_N grows more rapidly than in the case of simple Brownian motion. As an aside, however, we note that even for standard Brownian motion $S_N \sim AN/\ln(N)$ in two dimensions and $S_N \sim AN$ in three dimensions; consequently, choosing Lévy walks as a search mechanism will not lead to any significant gain compared to a Brownian search in these higher dimensions (except perhaps through the prefactor A). In one dimension, however, there are significant differences in the growth rates of S_N between Lévy and Brownian motions, and our task is to explore the place of intermittent random walks in this panorama.

3. Expected number of distinct visited sites

The expected number of distinct sites visited can be calculated once we determine the probability $P(s|s_0; n)$ that a given searcher, starting its intermittent random walk at site s_0 at time moment $n = 0$, appears at site s at time moment n. More specifically, the generating function $S(z)$ of S_N, defined as

$$S(z) = \sum_{N=0}^{\infty} S_N z^N, \tag{4}$$

and the lattice Green function (or site occupation generating function)

$$P(s|s_0; z) = \sum_{n=0}^{\infty} P(s|s_0; n) z^n \tag{5}$$

of the intermittent random walk are related to each other through [20]

$$S(z) = \frac{1}{1-z} \frac{\sum_s P(s|0; z)}{P(0|0; z)}. \tag{6}$$

Hence, given $P(s|0; z)$, we obtain $S(z)$ by virtue of equation (6). We then determine the N-dependence of S_N by inverting the discrete Laplace transform in equation (4).

The probability $P(s|s_0; n)$ obeys the recurrence relation

$$\begin{aligned} P(s|s_0; n) = {} & \frac{\alpha}{2}[P(s-1|s_0; n-1) + P(s+1|s_0; n-1)] \\ & + \frac{1-\alpha}{2}[P(s-L|s_0; n-T) + P(s+L|s_0; n-T)], \end{aligned} \tag{7}$$

which explicitly takes into account that jumps between nearest neighboring sites proceed in unit time, while long-range jumps over distance L require an integer time T. Equation (7) thus defines a non-Markovian process with a memory. Note also that since the intermittent random walk defined by equation (7) is homogeneous so that $P(s|s_0; n) = P(s - s_0|0; n)$, without loss of generality we henceforth set $s_0 = 0$.

Multiplying both sides of equation (7) by z^n and performing the summation, we find that $P(s|0; z)$ obeys

$$P(s|0; z) = \frac{1}{\pi} \int_0^{\pi} \frac{\cos(ks)\, \mathrm{d}k}{1 - \alpha z \cos(k) - (1-\alpha) z^T \cos(kL)}, \tag{8}$$

and consequently the generating function of the expected number of distinct sites visited is given by

$$S(z) = \frac{\pi}{(1-z)(1-\alpha z - (1-\alpha) z^T)} \left[\int_0^{\pi} \frac{\mathrm{d}k}{1 - \alpha z \cos(k) - (1-\alpha) z^T \cos(kL)} \right]^{-1}. \tag{9}$$

Before we proceed further, the following remarks are in order. Note that $P(s|0; z)$ in equation (8), and consequently $S(z)$ in equation (9), can be calculated explicitly in the two 'pure' random walk cases, $\alpha = 1$ and $\alpha = 0$. In these two limits, one finds for S_N the large-N asymptotic behavior

$$S_N(\alpha = 1) = \left(\frac{8N}{\pi} \right)^{1/2} + O\left(\frac{1}{N^{1/2}} \right) \tag{10}$$

and

$$S_N(\alpha = 0) = \left(\frac{8N}{\pi T} \right)^{1/2} + O\left(\frac{1}{N^{1/2}} \right). \tag{11}$$

The result in equation (10) is well known (see, for example, [35]).

- Sublinear growth of S_N with time N signifies that each site visited by such a walk is most probably visited many times. This oversampling is precisely the reason why searching a target in a one-dimensional system by a Brownian search is not very efficient, since the walker wastes a great deal of time revisiting sites that do not contain the target. That is why, in fact, recourse has been made to Lévy-based searches, since they reduce oversampling and lead to stronger growth of S_N with N.
- The result in equation (11) is the same as that in equation (10) with the replacement $N \to N/T$ and hence does not represent a good search strategy either—in fact, it is worse. When relocations over a distance L take unit time as do nearest neighbor steps, the results in equations (10) and (11) coincide, as they should.
- Note as well that in one dimension, equation (3) with S_N in equation (10) defines the asymptotically exact behavior of the non-detection probability of a target which *diffuses* in the presence of a concentration of *diffusive* searchers [36]. Thus, the asymptotic behavior of P_N is *independent* of the target diffusion coefficient (see also [37] and [22–26] for more details).
- The result in equation (3) can be generalized to the case of imperfect recognition of the target, that is, when target recognition upon encounter occurs with probability $p < 1$ [34]. In one-dimensional systems, the leading asymptotic behavior of P_N is *independent* of p provided that $p > 0$, and thus is also described by equation (3).

We seek a large-N expansion of S_N, in which (to arrive at the correct optimization) it is essential to retain not only the leading divergent contribution as $N \to \infty$ but also, if present, a constant N-independent correction term. Turning to the limit $z \to 1^-$ ($N \gg 1$) and inverting equation (9) we find, after some rather tedious but straightforward calculations, that S_N obeys

$$S_N = f_1 N^{1/2} + f_2 + O\left(\frac{1}{N^{1/2}}\right), \tag{12}$$

where the coefficient f_1 is given by

$$f_1 = \left(\frac{8}{\pi}\frac{\tau + L^2}{\tau + T}\right)^{1/2}, \qquad \alpha > 0, \tag{13}$$

$$= \left(\frac{8}{\pi T}\right)^{1/2}, \qquad \alpha \equiv 0. \tag{14}$$

The parameter

$$\tau \equiv \frac{\alpha}{1-\alpha} \tag{15}$$

is an important physical parameter which defines a characteristic time for a 'continuous tour of diffusion', that is, the typical time spent by a searcher on the substrate between two consecutive off-lattice flights. Note that the leading term in equation (12) grows as $N^{1/2}$, which means that the leading behavior is that of a random walk, albeit intermittent, unless there is an additional dependence of f_1 on N via the optimization of equation (12) with respect to α and L. Note also that the coefficients $f_1, f_2, \ldots$ are discontinuous functions of α and that $\alpha \equiv 0$ is a singular point since it is not possible for a random walker that skips over L sites at each step to ever visit all sites, cf [32]. This discontinuity should be viewed with appropriate caution since for any fixed finite N, S_N is a smooth function of α; the discontinuity arises because we are describing an asymptotic behavior that is strictly valid only in the $N \to \infty$ limit. Additional details and explanations can be found in [20] and [32].

Returning to the asymptotic $N \to \infty$ limit, for $\alpha \equiv 0$ the constant term $f_2 \equiv 0$, while for $0 < \alpha < 1$ (note that this double-sided inequality is strict) it is determined by

$$f_2 = -\frac{2(\alpha + (1-\alpha)L^2)}{\pi\sqrt{\alpha(1-\alpha)}} \int_0^{U_m} \frac{\mathrm{d}u}{\mathrm{sh}(u)\sqrt{1 - \tau\,\mathrm{sh}^2(u)}} \times \left(\frac{1}{2}\frac{\mathrm{sh}^2(2Lu)}{\mathrm{sh}^2(Lu) + \tau\,\mathrm{sh}^2(u)} - \frac{L}{\tau + L^2}\,\mathrm{cth}(u)\right), \tag{16}$$

where

$$U_m = \frac{1}{2}\ln\left(\frac{2-\alpha}{\alpha} + \sqrt{\left(\frac{2-\alpha}{\alpha}\right)^2 - 1}\right). \tag{17}$$

The integral in equation (16) cannot be performed in closed form, but its important contribution to the problem can be estimated. Anticipating that effective search strategies take place when $2U_m L \gg 1$ (see below), we find that in this limit the leading behavior of f_2 is given by

$$f_2 \sim -\frac{2}{\pi}\left(\tau^{1/2} + \frac{L^2}{\tau^{1/2}}\right) g(L), \qquad 0 < \alpha < 1, \tag{18}$$

where $g(L)$ is a slowly varying function of L,

$$g(L) = \ln\left(\frac{L}{(1+\tau)(1-\alpha)^{1/2}}\right) + 0.126 + O\left(\frac{1}{L}\right). \tag{19}$$

Equations (12)–(19) *constitute our main general result* and provide the basis for the design of an optimal strategy through the choice of α and L. Below we discuss such a design and show that, indeed, the optimal search strategies fulfill the assumption $2U_m L \gg 1$.

3.1. Optimization: tuning α at fixed L

To highlight the optimization procedure, we start with the case studied analytically and numerically in [20], shown again here for completeness, namely, we tune α but hold the relocation length L fixed.

Note that S_N defined by equations (12)–(19) is a non-monotonic function of the characteristic diffusion time τ. Differentiating S_N with respect to τ (discarding a weak dependence of $g(L)$ on τ), we find that the maximum of S_N with respect to τ is given implicitly as the solution of the equation

$$\frac{\partial f_1}{\mathrm{d}\tau}\left(\frac{8N}{\pi}\right)^{1/2} = \frac{1}{\pi\tau^{1/2}}\left(1-\frac{L^2}{\tau}\right)g(L), \tag{20}$$

where

$$\frac{\partial f_1}{\mathrm{d}\tau} = \frac{1}{2}\left(\frac{1}{[(\tau+T)(\tau+L^2)]^{1/2}} - \frac{(\tau+L^2)^{1/2}}{(\tau+T)^{3/2}}\right). \tag{21}$$

The left-hand side of equation (20) diverges when $N \to \infty$, which indicates that in this case with fixed L the optimal time τ of continuous tours of diffusion should tend to zero. We find that to leading order in N the optimal $\tau = \tau_{\mathrm{opt}}$, and hence the optimal value α_{opt} of the intermittency parameter α, are given by

$$\alpha_{\mathrm{opt}} \sim \tau_{\mathrm{opt}} \sim T\frac{L^{2/3}\ln^{2/3}(L)}{(2\pi N)^{1/3}}. \tag{22}$$

This is consistent with the condition $2U_mL \gg 1$ since $U_mL \sim L\ln(1/\alpha) \sim L\ln(N) \gg 1$. The symbol $\sim$ here and henceforth signifies the exact behavior to leading order in N. The expected number of distinct sites visited by an intermittent random walk with an optimal α and fixed L then is

$$S_N \sim \frac{L}{T^{1/2}}\left(\frac{8N}{\pi}\right)^{1/2}, \tag{23}$$

i.e., it differs by a factor of $L/\sqrt{T}$ from the corresponding result for a standard nearest neighbor random walk with $\alpha = 1$ (Brownian search), equation (10).

The essential result of this subsection is an enhancement by a factor $L/\sqrt{T}$ of the expected number of distinct sites visited by an intermittent random walk with the distribution in equation (2) and an appropriate N-dependent choice of the intermittency parameter compared to the outcome of an ordinary random walk. Note that this effect appears in an *exponent* in the non-detection probability, so it can become dramatically apparent. For example, for $L = 5$, $T = 1$ and $N = 10^4$, with a density of searchers as low as $\rho = 0.01$, the non-detection probability for a Brownian search is $P_N \approx 0.2$ while that of the optimal intermittent search ($\alpha_{\mathrm{opt}} \approx 0.07$ for these parameters) we find $P_N \approx 0.0003$, a reduction of three orders of magnitude! Note finally that in one dimension for fixed L the optimal strategy involves a progressively smaller fraction of nearest neighbor steps as N is increased.

3.2. Optimization: tuning α and L for $T = 1$

Next, we consider both α and L in equation (2) to be tunable, but we fix $T = 1$, that is, relocation to a nearest neighbor and to a neighbor a distance L away both take one unit of time. This causes the relocation velocity $V = L/T$ to become dependent on N via L. Setting $T = 1$ in equations (12)–(19), we have

$$S_N = (\alpha + (1-\alpha)L^2)^{1/2}\left(\frac{8N}{\pi}\right)^{1/2} - \frac{2}{\pi}\left(\left(\frac{\alpha}{1-\alpha}\right)^{1/2} + \left(\frac{1-\alpha}{\alpha}\right)^{1/2}L^2\right)g(L). \tag{24}$$

Note that the first term on the right-hand side of equation (24) grows with L and thus favors high values of L, but that the second term is negative and contains a higher power of L. This 'competition' suggests that there exists an optimal N-dependent value of L which leads to a maximum in the number of distinct sites visited.

To determine the optimal value of L, we differentiate S_N with respect to L. Again discarding a logarithmically weak dependence of $g(L)$ on L and anticipating that the optimal value of L is large, such that $L \gg (\alpha/(1-\alpha))^{1/2}$ (to be checked later for consistency), we find that the optimal value $L = L_{\rm opt}$ obeys

$$L_{\rm opt} \ln(L_{\rm opt}) \sim \left(\frac{\pi\alpha N}{2}\right)^{1/2}, \tag{25}$$

and hence,

$$L_{\rm opt} \sim 2\frac{(\pi\alpha N/2)^{1/2}}{\ln(\pi\alpha N/2)}. \tag{26}$$

Substituting this result into equation (24) then leads to the expected number of distinct sites visited in an intermittent random walk with an optimal length $L_{\rm opt}$ of the relocation stage,

$$S_N \sim 2\,(\alpha(1-\alpha))^{1/2}\,\frac{N}{\ln(\pi\alpha N/2)}. \tag{27}$$

Note the parabolic form of the prefactor as a function of α. The prefactor vanishes in the pure limits $\alpha = 0$ and $\alpha = 1$, indicating that optimization in these pure cases is not possible. This is a reflection of the fact that the leading behavior of S_N in these limits is determined by terms proportional to $N^{1/2}$ and constant terms independent of N are absent.

Further optimizing the prefactor in equation (27) with respect to α, we find that $\alpha_{\rm opt} \to 1/2$ as $N \to \infty$, and hence

$$S_N \sim \frac{N}{\ln(\pi N/4)}. \tag{28}$$

This result is consistent with the earlier assumptions $2U_m L \gg 1$ and $L \gg [\alpha(1-\alpha)]^{1/2}$ since here U_m and $\alpha/(1-\alpha)$ are constants while $L_{\rm opt}$ diverges as $N \to \infty$.

The result (28) shows that when $T = 1$, optimization of the intermittent search with respect to both α and L leads to an additional factor $N^{1/2}/\ln(N)$ in S_N via the coefficient f_1 in equation (12), which results in a much stronger dependence of the expected number of distinct sites visited on the maximal time N. In fact, we have obtained a behavior close to that of a two-dimensional Brownian motion, which signifies that optimal intermittent random walks lead to only marginal oversampling. The optimal strategy here consists of taking $\alpha_{\rm opt} = 1/2(\tau_{\rm opt} = T = 1)$, and relocation length $L_{\rm opt}$ as given in equation (26). Note that a similar result, i.e., that the maximum S_N is attained when the time spent on relocations is equal to the time spent in the diffusive stage, has been obtained for a model describing a stochastic search of a target site on a DNA by a protein [18, 38].

At this point, one might be tempted to say that exactly the same temporal behavior of S_N as in equation (28) can be found without resorting to any optimization procedure but by merely taking a Lévy walk with $\mu = 1$ in equation (1) (Cauchy distribution). Indeed, in this case one obtains (see equation (2.20) in [39])

$$S_N^{\rm Cauchy} \sim \frac{3}{2\pi^2}\frac{N}{\ln(N)}, \tag{29}$$

where the normalization constant B in equation (1) with $\mu = 1$ has been set to $3/\pi^2$. Remarkably, comparing the prefactors in equations (28) and (29) one notes that the search based on the intermittent strategy with fine-tuning of the optimization parameters outperforms the one based on a Cauchy distribution due to a numerical factor which is more than six times larger in the intermittent walk.

Moreover, we emphasize that these algorithms are very different in their quality of exploration of space. Consider, for example, the 'density of visited sites' defined by

$$\Omega_N = \frac{S_N}{2M_N}, \tag{30}$$

where M_N is the expected maximal displacement in, say, the positive direction so that $2M_N$ is a measure of the range of the walk. The parameter Ω_N is thus a measure of how many sites have been visited within the range of the walk. By definition, $0 \leqslant \Omega_N \leqslant 1$. For a Lévy walk with a Cauchy distribution of the relocation length, the distribution of the maximal displacement is well known [40]; the tail of this distribution exactly follows the behavior of the parent Cauchy variables and hence M_N^{Cauchy} is infinite. This implies that the density of visited sites vanishes, $\Omega_N = 0$, and thus the exploration quality is very poor. The expected maximal displacement of the intermittent random walk can be found from the general result of [41],

$$2M_N = (\alpha + (1-\alpha)L^2)^{1/2}\left(\frac{8N}{\pi}\right)^{1/2} + \gamma_N + O\left(\frac{1}{N^{1/2}}\right), \tag{31}$$

where

$$\gamma_N = \frac{2}{\pi}\int_0^\infty \frac{\mathrm{d}k}{k^2}\ln\left(\frac{2(1-\alpha\cos(k)-(1-\alpha)\cos(kL))}{(\alpha+(1-\alpha)L^2)k^2}\right). \tag{32}$$

We thus find that $\Omega_N \to 1/2$ as $N \to \infty$, which signifies that intermittent random walks have a very good exploration quality in that they visit half of the sites within their range. This is not a result expected *a priori* since we are dealing with random walks that involve steps not only to nearest neighbor sites but also to distant sites. We note that the exploration quality may be further enhanced by optimizing both S_N and Ω_N.

3.3. Optimization: tuning α and L for fixed V

We finally turn to the most difficult case, when relocation over distance L proceeds with a finite fixed velocity V. Note that this differs from the previous case, where $T = L/V$ was fixed. We take note of two points:

- We expect that $\alpha_{\text{opt}} \sim 1$ since flights over distance L are less favorable now that each relocation costs time $T = L/V$ during which no new sites are visited. If it turns out that the optimal relocation distance again grows with N (as it does, see below), then the time T grows with N as well. This in turn implies that it might become more advantageous to remain on the lattice, which means that the optimal τ might be larger than that in the previous case.
- On the other hand, the expression in equation (18) is only valid when $2U_mL \gg 1$. If $\alpha_{\text{opt}} \sim 1$, then $U_m \sim \sqrt{(1-\alpha)/\alpha} = 1/\sqrt{\tau} \ll 1$. Consequently, approximation (18) will be valid if the optimal characteristic time τ is not too large, i.e., $\tau \ll L^2$.

Differentiating equation (12) with respect to τ (again discarding the logarithmically slow variation of $g(L)$ with τ), we have

$$\left(\frac{1}{\sqrt{(\tau+L/V)(\tau+L^2)}} - \frac{\sqrt{\tau+L^2}}{(\tau+L/V)^{3/2}}\right)\left(\frac{8N}{\pi}\right)^{1/2} = \frac{2}{\pi}\left(\frac{1}{\sqrt{\tau}} - \frac{L^2}{\tau^{3/2}}\right)g(L). \tag{33}$$

Since equation (18) is only valid when $\tau \ll L^2$, equation (33) can be simplified to yield

$$\frac{L}{(\tau + L/V)^{3/2}} \left(\frac{8N}{\pi}\right)^{1/2} = \frac{2}{\pi} \frac{L^2}{\tau^{3/2}} g(L). \tag{34}$$

Now we proceed as follows. We first assume that $\tau \ll L/V$, find τ_{opt} and L_{opt}, and check whether the assumption is valid. Then we will follow with a more accurate calculation.

If $\tau \ll L/V$, then

$$\tau_{\text{opt}} = \frac{L^{5/3} g^{2/3}(L)}{V\,(2\pi N)^{1/3}}. \tag{35}$$

Substituting this expression into equation (12), we have

$$S_N = V^{1/2} L^{1/2} \left(\frac{8N}{\pi}\right)^{1/2} - \frac{2}{\pi} V^{1/2} L^{7/6} g^{2/3}(L)\,(2\pi N)^{1/6}. \tag{36}$$

Differentiating the latter expression with respect to L, we find that L_{opt} is defined implicitly by

$$L_{\text{opt}}^{2/3} g^{2/3}(L_{\text{opt}}) = \tfrac{3}{7}(2\pi N)^{1/3}, \tag{37}$$

which implies that

$$\tau_{\text{opt}} = \frac{L_{\text{opt}}}{V} \frac{L_{\text{opt}} g^{2/3}(L_{\text{opt}})}{(2\pi N)^{1/3}} = \frac{3}{7} \frac{L_{\text{opt}}}{V}. \tag{38}$$

Hence, τ_{opt} is *not* much smaller than L_{opt}/V. It is only smaller by a numerical factor and scales as $L_{\text{opt}} \sim N^{1/2}/\ln(N)$ (note that nonetheless $\tau_{\text{opt}} \ll L_{\text{opt}}^2$). Note also that $\tau_{\text{opt}} \to \infty$ as $N \to \infty$, which implies that $\alpha_{\text{opt}} \to 1$ and $U_m \sim (1-\alpha)^{-1/2} \sim \tau^{1/2}$.

We next try to search for optimal L and τ from equation (34) supposing that $U_m \sim \tau^{1/2}$ and $\tau = CL/V$, where C is a constant to be determined. We note that in this case

$$\begin{aligned} g(L) &= \ln(L) - \ln((1+\tau)\sqrt{1-\alpha}) + 0.126 + O\left(\frac{1}{L}\right) \\ &\approx \tfrac{1}{2}\ln(L) - \tfrac{1}{2}\ln(C/V) \\ &\approx \tfrac{1}{2}\ln(L). \end{aligned} \tag{39}$$

Hence, equation (34) becomes

$$\frac{V^{3/2} L}{(1+C)^{3/2} L^{3/2}} \left(\frac{8N}{\pi}\right)^{1/2} = \frac{2}{\pi} \frac{L^2}{\tau^{3/2}} \frac{1}{2} \ln(L). \tag{40}$$

From equation (40) we find that the optimal value of τ obeys

$$\tau_{\text{opt}} = \frac{(1+C)}{V} \frac{L^{5/3} \ln^{2/3}(L)}{(2\pi N)^{1/3}}, \tag{41}$$

and S_N, optimized with respect to τ, then follows

$$S_N = \left(\frac{LV}{1+C}\right)^{1/2} \left(\frac{8N}{\pi}\right)^{1/2} - \frac{1}{\pi} \left(\frac{V}{1+C}\right)^{1/2} L^{7/6} \ln^{2/3}(L)(8\pi N)^{1/6}. \tag{42}$$

Differentiating this equation with respect to L, we find that the optimal flight length L obeys

$$L_{\text{opt}} \ln(L_{\text{opt}}) = \left(\tfrac{3}{7}\right)^{3/2} (8\pi N)^{1/2}, \tag{43}$$

so that

$$L_{\rm opt} \sim \frac{\left(\frac{3}{7}\right)^{3/2}(8\pi N)^{1/2}}{\ln\left(\left(\frac{3}{7}\right)^{3/2}(8\pi N)^{1/2}\right)}. \tag{44}$$

Lastly, we obtain C from the definition $\tau_{\rm opt} = CL_{\rm opt}/V$. This gives

$$(1+C)\frac{L_{\rm opt}}{V}\frac{L_{\rm opt}^{2/3}\ln^{2/3}(L)}{(8\pi N)^{1/3}} = C\frac{L_{\rm opt}}{V} \tag{45}$$

which is solved by

$$C = \tfrac{3}{4}. \tag{46}$$

Thus the optimal strategy is realized when we choose $L_{\rm opt}$ as in equation (44), and $\tau = (3/4)T$, i.e., for the optimal strategy the characteristic time of a tour of diffusion between two consecutive long jumps is three-fourths of the time a searcher spends on jumps over a distance L.

Finally, we combine these results to find that the expected number of distinct sites visited optimized with respect to both α and L at fixed V is

$$S_N = \left(\frac{4}{7}\right)^{1/2} V^{1/2} L_{\rm opt}^{1/2}\left(\frac{8N}{\pi}\right)^{1/2} \sim V^{1/2}\frac{N^{3/4}}{\ln^{1/2}(N)}. \tag{47}$$

Note that this more intricate optimization procedure results in a faster growth law of S_N with N than in the pure random walk cases. Consequently, the search efficiency has again been enhanced by orders of magnitude, albeit not as much as in the fixed-T case of the previous subsection.

4. Conclusions

We have considered the design of an optimal search strategy of a hidden target by a given density ρ of random walkers who have a limited maximal time N to find the target. The analysis is restricted to one-dimensional systems. Our measure for the quality of a strategy is *the minimization of the probability P_N that the target is undetected within the given maximal search time*. In particular, we considered strategies that consist of a combination of nearest neighbor walks and jumps of fixed length L, both involving steps in either direction with equal probability. The motion of each searcher is thus intermittent. The probability that the target is undetected is just the survival probability for the target and is related to the distinct number of sites S_N visited by a walker up to time N by the well-known relation $P_N = \exp(-\rho S_N)$. Our goal was thus to maximize S_N. We stress that the time derivative of P_N defines the distribution function of the first passage time to the detection event. This means that in contrast to previous work [15–18, 20], our aim has been to optimize the full distribution function and not only its first moment.

Our model has three parameters: α, the probability that the next step of the walker is a nearest neighbor step; L, the length of a long step; and $T = L/V$, where T is the time it takes to cover a long step and V is the velocity of a long step. The parameters α and L are optimized as functions of the maximal time N under different constraints. We compared our results for S_N with those of a nearest neighbor random walk, $S_N \sim (8N/\pi)^{1/2}$, and of a Lévy walk with a Cauchy distribution of step lengths, $S_N \sim (3/2\pi^2)N/\ln(N)$.

If L and T are fixed and only α is picked for optimal strategy, the best choice is to take it to be very small, $\alpha \sim N^{-1/3}$ [20]. Most of the random motion then consists of steps of length L.

The resulting S_N is larger than that obtained with a nearest neighbor random walk by the numerical coefficient $L/\sqrt{T}$, $S_n \sim (L/T^{1/2})(8N/\pi)^{1/2}$, so that the non-detection probability, which involves this factor in the exponent, can be decreased by orders of magnitude even for a low density of searchers.

If we optimize both α and L, the best choice of these parameters depends on the constraint we place on the third parameter. It the time for a long step is the same as the time for a nearest neighbor step, then the optimal choices are $\alpha = 1/2$ and $L \sim N^{1/2}/\ln(N)$. Thus, short and long steps should occur with equal likelihood, and the optimal distance covered by the long steps grows (slowly) with increasing observation time. In this case, the distinct number of sites visited is even larger (again by a numerical factor) than that of a Lévy walk, $S_N \sim N/\ln(\pi N/4)$. These N dependences are the same for any fixed value of T, although the specific numerical coefficients depend on this value. We find an important difference between the Lévy walk and our optimized intermittent walk in the coverage of space, which may be an important feature if one wishes to parallelize the searches of different walkers. The density of visited sites vanishes in the Lévy case, while that of the intermittent walk approaches a constant with increasing N.

Finally, if we again optimize both α and L but now keeping the velocity V of the long steps fixed (which means that the time for a long step grows with the length of the step), we find the optimal choice α to be close to unity, $\alpha \sim 1 - 4V\ln(N)/3N^{1/2}$, and the optimal length step to grow with N as $L \sim N^{1/2}/\ln(N)$. Now the walkers rarely jump over long distances, but the time spent diffusing and the time it takes to make a jump of length L both grow with increasing observation time N, so that $\tau = 3T/4$. The distinct number of sites visited here shows a growth intermediate between the other two intermittent strategies, $S_N \sim V^{1/2}N^{3/4}/\ln^{1/2}(N)$.

We stress that in all the strategies considered here, we have implemented a maximum observation time N as part of the optimization process, and we have used a particular measure of the quality of a strategy, namely, that the probability of survival of the target be minimized within this time. We have shown that in one dimension even a simple intermittent step distribution consisting of nearest neighbor steps and long steps of fixed N-dependent length can be optimized so as to yield far better search outcomes than a nearest neighbor random walk and even a Lévy walk with a Cauchy distribution of step lengths. The model can be generalized and further optimized in a number of ways, some of which we have noted in the course of our analysis.

Acknowledgments

The authors gratefully acknowledge helpful discussions with O Bénichou, J Klafter and A Grosberg. The research of GO is partially supported by Agence Nationale de la Recherche (ANR) under grant 'DYOPRI - Dynamique et Optimisation des Processus de Transport Intermittents'. HSW acknowledges partial support from MEC, Spain, through Grant CGL2007-64387/CLI, and from AECID, Spain, project A/018685/08.

References

[1] Stephens D W and Krebs J R 1986 *Foraging Theory* (Princeton, NJ: Princeton University Press)

[2] Stone L D 1989 *Theory of Optimal Search* (Arlington, VA: Operations Research of America)

[3] Bell J W 1991 *Searching Behavior, The Behavioural Ecology of Finding Resources (Chapman and Hall Animal Behaviour Series)* (London: Chapman and Hall)

[4] Frost J R and Stone L D 2001 *Review of Search Theory: Advances and Applications to Search and Rescue Decision Support* (US Coast Guard Research and Development Center) (Springfield, VA: Natl Tech. Inf. Serv.)
[5] Klafter J, Shlesinger M and Zumofen G 1996 *Phys. Today* **49** 33
Shlesinger M and Klafter J 1986 Lévy walks versus Lévy flights *On Growth and Form* ed H E Stanley and N Ostrowski (Amsterdam: Martinus Nijhof)
[6] Viswanathan G M, Afanasyev V, Buldyrev S V, Murphy E J, Prince P A and Stanley H E 1996 *Nature* **381** 413
[7] Viswanathan G M, Buldyrev S V, Havlin S, da Luz M G E, Raposo E P and Stanley H E 1999 *Nature* **401** 911
[8] Levandowsky M, Klafter J and White B S 1988 *J. Protozool.* **35** 65
[9] Bartumeus F, Catalan J, Fulco U L, Lyra M L and Viswanathan G M 2002 *Phys. Rev. Lett.* **88** 097901
[10] Boyer D, Ramos-Fernandez G, Miramontes O, Mateos J L, Cocho G, Larralde H, Ramos H and Rojas F 2006 *Proc. Biol. Sci.* **273** 1743
[11] Bartumeus F 2007 *Fractals* **15** 151
[12] Boyer D, Miramontes O and Ramos-Fernandez G 2008 arXiv:0802.1762
[13] O'Brien W J, Browman H I and Evans B I 1990 *Am. Sci.* **78** 152
[14] Kramer D L and MacLaughlin R L 2001 *Am. Zool.* **41** 137
[15] Bénichou O, Coppey M, Moreau M, Suet P-H and Voituriez R 2005 *Phys. Rev. Lett.* **94** 198101
[16] Bénichou O, Loverdo C, Moreau M and Voituriez R 2006 *Phys. Rev.* E **74** 020102 Reviewed in:
Shlesinger M F 2006 *Nature* **443** 281
[17] Bénichou O, Coppey M, Moreau M and Voituriez R 2006 *Europhys. Lett.* **75** 349
[18] Coppey M, Bénichou O, Voituriez R and Moreau M 2004 *Biophys. J.* **87** 1640
[19] Reynolds A M 2006 *Europhys. Lett.* **75** 517
[20] Oshanin G, Wio H S, Lindenberg K and Burlatsky S F 2007 *J. Phys.: Condens. Matter* **19** 065142
[21] Lomholt M A, Koren T, Metzler R and Klafter J 2008 *Proc. Natl Acad. Sci. USA* **105** 11055
[22] Yuste S B and Lindenberg K 2005 *Phys. Rev.* E **72** 061103
Yuste S B and Lindenberg K 2006 *Phys. Rev.* E **73** 039909
[23] Yuste S B, Ruiz-Lorenzo J J and Lindenberg K 2006 *Phys. Rev.* E **74** 046119
[24] Ruiz-Lorenzo J J, Yuste S B and Lindenberg K 2007 *J. Phys.: Condens. Matter* **19** 065120
[25] Yuste S B and Lindenberg K 2007 *Phys. Rev.* E **76** 051114
[26] Yuste S B, Oshanin G, Lindenberg K, Bénichou O and Klafter J 2008 *Phys. Rev.* E **78** 021105
[27] James A, Plank M J and Brown R 2008 *Phys. Rev.* E **78** 051128
[28] Szu H and Hartley R 1987 *Phys. Lett.* A **122** 157
[29] Edwards A M *et al* 2007 *Nature* **449** 1044
[30] Travis J 2007 *Science* **318** 742
[31] Boyer D and López-Corona O 2009 *J. Phys. A: Math. Theor.* **42** 434014 (this special issue)
[32] Lakatos-Lindenberg K and Shuler K 1971 *J. Math. Phys.* **12** 633
[33] Oshanin G, Wio H S, Lindenberg K and Burlatsky S F (in preparation)
[34] Bénichou O, Moreau M and Oshanin G 2000 *Phys. Rev.* E **61** 3388
[35] Hughes B D 1995 *Random Walks and Random Environments* (Oxford: Clarendon)
[36] Bray A J and Blythe R A 2002 *Phys. Rev. Lett.* **89** 150601
Blythe R A and Bray A J 2003 *Phys. Rev.* E **67** 041101
[37] Oshanin G, Bénichou O, Coppey M and Moreau M 2002 *Phys. Rev.* E **66** 060101
[38] Berg O G, Winter R B and Von Hippel P H 1981 *Biochemistry* **20** 6929
[39] Gillis J E and Weiss G H 1970 *J. Math. Phys.* **11** 1307
[40] Darling D A 1956 *Trans. Am. Math. Soc.* **83** 164
[41] Comtet A and Majumdar S N 2005 *J. Stat. Mech. Theor. Exp.* **06** 06013