Fitness-maximizing foragers can use information about patch quality to decide how to search for and within patches: optimal Lévy walk searching patterns from optimal foraging theory

A. M. Reynolds*

Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK

Optimal foraging theory shows how fitness-maximizing foragers can use information about patch quality to decide how to search within patches. It is amply supported by empirical studies. Nonetheless, the theory largely ignores the fact that foragers may need to search for patches as well as for the targets within them. Here, using an exact but simple mathematical argument, it is shown how foragers can use information about patch quality to facilitate the execution of Lévy walk movement patterns with $\mu = 2$ at inter-patch scales. These movement patterns are advantageous when searching for patches that are not depleted or rejected once visited but instead remain profitable. The analytical results are verified by the results of numerical simulations. The findings bring forth an innovative theoretical synthesis of searching for and within patches and, suggest that foragers’ memories may be adaptive under spatially heterogeneous reward schedules.

Keywords: optimal foraging theory; Lévy walks; heterogeneous landscapes; patchy distributions

1. INTRODUCTION

Foraging behaviour has been a central concern of ecology. Our current understanding of foraging behaviours owes much to the ground-breaking work of MacArthur & Pianka [1], who established the field of optimal foraging theory. The development of this theory and its empirical verifications have been charted by Pyke et al. [2] and Stephens & Krebs [3]. The theory posits that animals forage in a way that maximizes their rate of energy intake. Foraging for patchily distributed resources has been a focus of attention [4–12], as virtually every ecosystem will exhibit patchiness on a range of spatial, temporal and organizational scales [13]. When foraging for patchily distributed resources, it is initially advantageous to search within a patch because resources are present there in relatively high abundance. The benefit of remaining within a patch will, however, generally diminish with time because of the gradual depletion of resources due to foraging. Central to optimal foraging theory has been the marginal value theorem, which dictates that a forager should leave a particular patch when the expected energy gained from staying drops to the expected net gain from travelling to and foraging within a new patch. The durations of patch-foraging episodes can also be chosen to balance the conflicting demands for food and safety from predation [14]. Factors influencing this trade-off include the encounter rate with predators, predator lethality in the absence of vigilance, effectiveness of vigilance in reducing predator lethality, the marginal value of energy to the forager and the forager’s survivor’s fitness. Thus, patches are not necessarily depleted or rejected once visited but can instead remain profitable and so warrant repeated visits.

Most models of patch use assume that foragers assess spatial heterogeneity in feeding or fitness opportunities and respond accordingly. Prescient foragers use sensory cues or memory in predicting patch quality [9]. Bayesian foragers, on the other hand, estimate patch quality from: (i) an a priori knowledge of the distribution of patch qualities throughout the environment; (ii) time spent thus far in searching for food items within the patch; and (iii) the number of food items thus far encountered and harvested ([4–12] and see Valone [15], for a review of empirical data). Optimal foraging theory has identified how fitness-maximizing foragers should use this information to decide how thoroughly to use each food patch. The theory has, however, tended to ignore the large-scale search problem (i.e. patch detection). Most models of optimal foraging theory assume that foragers have information about patch locations so that time spent between patches is attributed to travelling rather than to searching. There has been a failure to recognize that foragers may need to search for patches as well as for the targets

*andy.reynolds@rothamsted.ac.uk
within them [16]. The need to search for patches will arise if the environment is new, highly dynamic or if foragers have limited perceptual/cognitive abilities [17].

To come closer to actuality, optimal foraging theory needs to improve the integration of search components at different scales of the foraging process [16]. This might be achieved by drawing on Lévy walk (LW) search theory, which has been concerned with identifying movement patterns that are advantageous when searching for randomly and sparsely distributed patches ([18–20], reviewed in Reynolds & Rhodes [21] and Viswanathan et al. [22]).

LW comprise clusters of relatively short step lengths with longer steps between them. This pattern is repeated across all scales, with the resultant scale-free clusters creating fractal patterns. Step lengths, \( l \), are drawn from distributions with inverse power-law tails: \( p(l) \sim l^{-\mu} \), where \( 1 < \mu \leq 3 \). Distributions of total displacements (i.e. sums of many individual step lengths) tend to Lévy stable distributions by virtue of a generalized central limit theorem due to Gnedenko & Kolmogorov [23]. Directions of movement are uniformly distributed between \( 0^\circ \) and \( 360^\circ \). As \( \mu \rightarrow 1 \), the fraction of steps with length less than the typical spacing between patches becomes negligible, so that an LW forager effectively moves along a straight line until it detects a patch [18]. Movement patterns with \( \mu > 3 \) are effectively Brownian at large spatio-temporal scales because total displacements (i.e. sums of many individual step lengths) eventually become Gaussian distributed.

Over much iteration an LW will be distributed much further from its starting position than a Brownian walk of the same length (same sum of step lengths) because there are relatively more long steps and so less meandering. The fractal and ‘superdiffusive’ properties of LW can be advantageous when searching for randomly and sparsely distributed patches that are not depleted or rejected once visited but instead remain profitable and so targets for later searches. Foragers may be unable to return directly to the last patch visited if they have become satiated or leave the area after feeding or are taken outside the patch by an extended but unsuccessful bout of intensive (within patch) searching. Theoretical studies [18–20] have shown that LW with \( \mu = 2 \) can increase the patch encounter probability compared with simple Brownian, straight line and LW with \( \mu \neq 2 \) moments.

LW first entered the literature on animal movement patterns when it was proposed that their characteristics may be observed in foraging ants [24]. Escalating empirical support for LW movement patterns (reviewed in Reynolds & Rhodes [21]) subsequently faulted when methodological shortcomings in some approaches to determining the goodness of fits of power-laws to movement pattern data were identified [25]. More recent studies have, however, provided compelling evidence that basking sharks along with bigeye tuna, Atlantic cod, leatherback turtles, and Magellanic penguin use a movement strategy approximated by LW with \( \mu = 2 \) when they are searching for resources [26,27]. These marine predators feed on patchily distributed prey. Spider monkeys (Ateles geoffroyi), fruit flies (Drosophila), microzooplankton (Oxyrrhis marina), and bacteria (Escherichia coli) appear to use the same movement strategy when foraging [28–32].

Here, using an exact but simple mathematical argument, it is shown that how foragers can use information about patch quality to facilitate the execution of LW movement patterns with \( \mu = 2 \) and thereby enhance their net energy intake by reducing the average time spent searching for patches. In this picture, the lower quality (smaller) patches act as ‘signposts’ that help guide foragers to the most profitable patches: the lower quality patches indicate that the forager should ‘keep going forward’ whilst the high-quality patches indicate that the forager should ‘turn’. The results of this theoretical analysis reveals that the mechanism could operate whenever patch encounters occur with unchanging likelihood and are independent of one another, i.e. whenever patch encounters are described by a Poisson process. The analysis, which draws upon the previous work of Eliazar & Klafter [33] for the temporal generation of power-law distributions, is described in §2. The findings are subsequently supported by the results of numerical simulations presented in §3. This is followed by a discussion in §4.

2. LÉVY WALK INTER-PATCH MOVEMENT PATTERNS FROM THE BASIC ATTRIBUTES OF CLASSIC OPTIMAL FORAGING THEORY

Here it is shown that foragers will execute LW movement patterns with \( \mu = 2 \) at inter-patch scales if patch quality is continuously rather than discretely distributed and if the foragers adopt the following movement strategy: (i) randomly select a direction of travel, and travel in a straight line along that direction for some time, \( T \) (exclusive of any time spent foraging at patches encountered on the way). (ii) Take note of the quality, \( Q_{\text{max}} \), of the best patch encountered. This is set to zero if no patches are encountered. (iii) Continue to move in a straight line until coming across a patch whose quality is perceived to be greater than or the same as \( Q_{\text{max}} \) (which is always possible when patch quality is continuously distributed). Foraging can take place at patches encountered on the way. After leaving a patch with quality \( Q \geq Q_{\text{max}} \), a new direction of travel is chosen at random and the cycle is repeated. The results of numerical simulations presented in §3 reveal that the LW movement patterns with \( \mu = 2 \) arise if instead of (i) the forager keeps moving until encountering a patch and then uses the quality of that patch when subsequently assessing patch quality along that direction of travel. An example of such a movement pattern is illustrated in figure 1.

The movement rules have resonance with the solution of the ‘secretary problem’ (also known as the marriage or best choice problem). In this problem, there is a single secretarial position to fill and \( n \) applicants that can be ranked from best to worst. The applicants are interviewed sequentially in a random order, with each order being equally likely. After each interview, the applicant is accepted or rejected. The decision to accept or reject an applicant can be
probability that the forager travels for a time \( t > \tau \) before encountering a better patch with quality \( Q > Q_{\text{max}} \) is

\[
P_t(t > \tau | Q_{\text{max}} = Q) = \exp(-\Lambda(Q_{\text{max}})\tau). \tag{2.3}
\]

The unconditional probability is obtained by combining equations (2.2) and (2.3) and then using the change of variables \( x = TA(Q) \):

\[
P_t(t > \tau) = \int_0^{Q_{\text{out}}} P_t(t > \tau | Q_{\text{max}} = Q) dP(Q) \]
\[
= \int_0^{Q_{\text{out}}} T \exp(-\Lambda(Q)T) \Lambda(Q) \exp(-\Lambda(Q)\tau)dQ \]
\[
= T \int_0^{Q_{\text{out}}} \exp(-\Lambda(Q)(T+\tau))d\Lambda(Q) \]
\[
= \int_0^\infty \exp \left( - \frac{T + \tau}{T} x \right) dx \]
\[
= \frac{T}{T + \tau}. \tag{2.4}
\]

If the patch-encounter rate is independent on patch quality then, equation (2.4), can be deduced using a simple, intuitive argument: if at least one patch has been found at time \( T + \tau \) then the probability that the best (or the only) patch was found before time \( T \) is just \( T/(T + \tau) \).

It follows from equation (2.4) that the probability density function of travel times to path-changing patches is

\[
p(\tau) = \frac{T}{(T + \tau)^\nu}. \tag{2.5}
\]

Notice that this probability distribution function is independent of the Poissonian rate function, \( \lambda(Q) \), and so independent of the average target density. The probability distribution function of travel times, equation (2.5), is therefore applicable whenever patch encounters can be described by a Poisson rate function. A probability density function directly analogous to that given in equation (2.5) is also obtained when it is assumed that patches are encountered at discrete times \( (t = 1, 2, 3, \ldots) \) in arbitrary time units) rather than continuously in time.

The probability density function of travel times between consecutive turns (i.e. the probability density function of distances travelled between consecutive turns when moving at constant speed) therefore has an inverse-square power-law tail (irrespective of the duration, \( T \), of initial patch quality assessment periods). This is indicative of LW movement patterns with \( \mu = 2 \). This is a general result that holds true for all time-homogeneous Poisson patch-encounter processes possessing a rate function. Selection pressures for LW with \( \mu = 2 \) will, however, decrease as the target density increases, because the advantageous that these movement patterns have over other movement patterns when randomly searching decreases as the target density increases [18]. The reorientation mechanism that gives rise to LW with \( \mu = 2 \) is therefore not expected
3. EMERGENT LÉVY WALK MOVEMENT PATTERNS ARE ADVANTAGEOUS

The emergence of $\mu = 2$ LW movement patterns from the aforementioned movement rules is here examined with the aid of numerical simulations. Also examined is the effectiveness of these Lévy foragers in locating randomly distributed patches compared with the effectiveness of foragers that maintained their direction of travel or randomly change their direction of travel only after leaving a patch.

The search arena in these numerical simulations had size $5000 \times 5000$ in arbitrary space units and contained 250 000 randomly placed patches. All patches were circular with radius one and patch quality was exponentially distributed with mean $Q = 10$ arbitrary quality units. Simulation data were pooled from 500 independent simulations for the movement patterns of individual foragers moving with unit speed for $10^7$ time steps. The initial directions of travel are chosen at random. Each forager then moves continuously without interruption in a straight line until encountering a patch that is better in quality than the first patch encountered along that direction of travel. The first patch is typically encountered after 50 time steps. A new direction of travel is then chosen at random and the cycle repeats. Periodic boundary conditions were applied at the edges of the search arena to keep the foragers within the search arena. The target distribution was randomized after each application of this condition to prevent the foragers from encountering the same set of patches.

The Akaike information criterion (AIC) was used to test whether the simulation data provided more evidence for distances, $l$, travelled between successive turns coming from power-law distributions

$$P_1(l) = (\mu - 1) a^{\mu-1} l^{\mu} \quad (l > a)$$

exponential distributions

$$P_2(l) = \lambda e^{-\lambda l} \quad (l > a)$$

or gamma distributions

$$P_3(l) = \frac{\Gamma(s)}{I(s)} l^{s-1} e^{-\lambda l},$$

where $I(s)$ is a gamma function. Power-law distributions are consistent with the presence of LW patterns. Exponential and gamma distributions are not. The Akaike weight for a power-law distribution can be considered as the weight of evidence in favour of a power-law distribution being the better model of the simulation data, i.e. the Akaike weight for a power-law can vary from 0 (no support) to 1 (complete support) [25]. A power-law probability distribution function of move-lengths is indicative of LW movement patterns. The exponential probability distribution function of move-lengths corresponds to a simple uncorrelated Poisson random process. Plots of the survival function (the complement of the cumulative distribution function) were also used to further examine the form of the tails and to determine the extent of any power-law scaling: an approach that is more reliable than probability density function plots [35]. To construct the survival function, the simulation data $\{l\}$ are first ranked from largest to smallest ($i = 1 \ldots n$). The probability that a length is greater than or equal to $l_i$ (the survival function) is then estimated as $i/n$. The fractal dimension, $D$, of the turning points in the simulated movement patterns was also used to test for the presence of LW. The fractal dimension, $D$, of the turning points in an LW approaches $1$ asymptotically for sufficiently long sequences [36]. Here the fractal dimension, $D$, was computed by calculating the average number, $n_i$, of boxes of size $l$ arbitrary space units required to enclose all of the turning points in a simulated movement pattern. A power-law relationship of the form $n_i \propto l_i^{-D}$ would be indicative of fractal scaling with fractal dimension $D$.

The Akaike weight for power-law distribution is 1 and the maximum-likelihood estimate for the Lévy exponent $\mu = 2$ ($a = 1000$ arbitrary space units). The Akaike weights for the exponential and gamma distributions are 0. The power-law scaling extends from about the mean-free path length (i.e. the average straight line distance between targets, $\lambda = 100$ arbitrary space units) to about the total length of the search path (figure 2). These estimates for the Lévy exponents, $\mu$, are consistent with the estimate, $D \approx 1$, for the fractal dimension shown in figure 3. The maximum-likelihood estimates for the parameters, $\lambda$, $r$ and $S$, in the exponential and gamma distributions are $6.9 \times 10^{-3}$ reciprocal arbitrary space units, $1.3 \times 10^{-4}$

![Figure 2](http://rsif.royalsocietypublishing.org/Downloaded from)}
A power-law relationship of the form \( n_l \) to decide how to search maximising foragers should use information about exclusively concerned with understanding how fitness-maximizing foragers will move in a straight line when searching for patches as well as for the targets within them. Searching for patches will be of importance when the environment is new, dynamic or cannot be mapped in its entirety because the foragers have limited cogitative/perceptual capabilities. LW theory, on the other hand, has been principally concerned with identifying advantageous large-scale movement patterns for the location of patches [18–20]. This theory has revealed that LW with \( \mu = 2 \) can increase the patch‐encounter probability compared with simple Brownian, straight line and LW with \( \mu \neq 2 \) when patches are randomly and sparsely distributed, prompting the suggestion that LW with \( \mu = 2 \) represent an evolutionary optimal searching strategy [37,38].

In this article, it was shown how foragers can use information about patch quality to facilitate the execution of LW movement patterns with \( \mu = 2 \) and thereby search effectively for patches. In this novel picture, the low-quality (smaller) patches act as signposts that guide foragers to the most profitable patches.

It has become apparent that other mechanisms (e.g. conspecific odour trail avoidance and re-orientating at odour trails left by correlated random walkers) can also lead to the emergence of LW with \( \mu = 2 \) movement patterns [39–43]. However, when competing hypotheses are equal in other respects, the law of parsimony recommends selection of the hypothesis that introduces the fewest assumptions and postulates the fewest entities while still sufficiently answering the question. Therefore, searching for and within patches, should, so far as possible, be attributed to similar causes and cogitative abilities.

Whether or not foragers actually use patch quality when searching for patches is an open question that has not been explored in the literature. Nonetheless, a candidate exemplar may be found in the classic observational study made by Heinrich [44] of bumblebee (Bombus terricola) movements (distances and turning angles) at clover (Trifolium repens) patches. Heinrich [44] observed that bumblebees tend to approach the nearest flower but then often depart without landing or probing it, if it has been visited previously; unvisited flowers (high-quality resources) are not rejected in this manner. Patch quality is assessed through the detection of scent marks [45,46] and learnt by experiencing that recently visited flowers are unrewarding [47]. Heinrich [44] also observed that on successive flower-head visits in depleted areas, bumblebees moved more forward than backward (82% versus 18%). These observations are consistent with the theoretical expectation that fitness-maximizing foragers will move in a straight line until encountering a high-quality patch, neglecting any poor patches encountered on route. The theoretical prediction that such foragers will have LW movement patterns with \( \mu = 2 \) is consistent with Edwards et al. [25] re-examination of Heinrich’s [44] data. Edwards et al. [25] found tentative evidence of LW flight patterns with \( \mu \approx 2 \), reporting that these LW were marginally favoured over an alternative model in which step lengths were exponentially distributed.

The emphasis has been on prey choice and patch residence time. The theory has largely ignored the fact that foragers may need to search for patches as well as for the targets within them. Searching for patches will be of importance when the environment is new, dynamic or cannot be mapped in its entirety because the foragers have limited cogitative/perceptual capabilities. LW theory, on the other hand, has been principally concerned with identifying advantageous large-scale movement patterns for the location of patches [18–20]. This theory has revealed that LW with \( \mu = 2 \) can increase the patch‐encounter probability compared with simple Brownian, straight line and LW with \( \mu \neq 2 \) when patches are randomly and sparsely distributed, prompting the suggestion that LW with \( \mu = 2 \) represent an evolutionary optimal searching strategy [37,38].

In this article, it was shown how foragers can use information about patch quality to facilitate the execution of LW movement patterns with \( \mu = 2 \) and thereby search effectively for patches. In this novel picture, the low-quality (smaller) patches act as signposts that guide foragers to the most profitable patches.

It has become apparent that other mechanisms (e.g. conspecific odour trail avoidance and re-orientating at odour trails left by correlated random walkers) can also lead to the emergence of LW with \( \mu = 2 \) movement patterns [39–43]. However, when competing hypotheses are equal in other respects, the law of parsimony recommends selection of the hypothesis that introduces the fewest assumptions and postulates the fewest entities while still sufficiently answering the question. Therefore, searching for and within patches, should, so far as possible, be attributed to similar causes and cognitative abilities.

Whether or not foragers actually use patch quality when searching for patches is an open question that has not been explored in the literature. Nonetheless, a candidate exemplar may be found in the classic observational study made by Heinrich [44] of bumblebee (Bombus terricola) movements (distances and turning angles) at clover (Trifolium repens) patches. Heinrich [44] observed that bumblebees tend to approach the nearest flower but then often depart without landing or probing it, if it has been visited previously; unvisited flowers (high-quality resources) are not rejected in this manner. Patch quality is assessed through the detection of scent marks [45,46] and learnt by experiencing that recently visited flowers are unrewarding [47]. Heinrich [44] also observed that on successive flower-head visits in depleted areas, bumblebees moved more forward than backward (82% versus 18%). These observations are consistent with the theoretical expectation that fitness-maximizing foragers will move in a straight line until encountering a high-quality patch, neglecting any poor patches encountered on route. The theoretical prediction that such foragers will have LW movement patterns with \( \mu = 2 \) is consistent with Edwards et al. [25] re-examination of Heinrich’s [44] data. Edwards et al. [25] found tentative evidence of LW flight patterns with \( \mu \approx 2 \), reporting that these LW were marginally favoured over an alternative model in which step lengths were exponentially distributed.
Theoretical expectation that foragers adapt to local variations in patch quality is also consistent with observations of the nectar-gathering foraging behaviours of bumblebees (Bombus flavifrons and B. bifarius) on control and nectar-enriched inflorescences of three Rocky Mountain herbs [48]. Cibula & Zimmerman [48] observed that bumblebees encountering a series of enriched plants initially increased the number of flowers per inflorescence that they visited relative to the control inflorescences. After the blossoms from a small number of enriched racemes were probed, the average number of flowers visited per plant then returned to control levels.

Spider monkeys (Ateles geoffroyi) may be another candidate organism. The travel paths of spider monkeys are composed of highly linear ‘beeline’ segments, each typically ending at a place where some resource (fruiting trees) is exploited [49]. The monkeys can travel up to 1.5 km in a straight line before stopping at a resource [49]. Even when food is scarce, a number of fruiting trees would typically be found along a straight transect of this length (G. Ramos-Fernández, personal communication). The monkeys do not generally continue in the same ranging direction after exploiting a resource: travel paths are likely to deviate at the site of resource exploitation rather than between such sites [49]. Ramos-Fernández et al. [29] reported that a power-law with $\mu \approx 2$ provides a good fit to the distribution of distances travelled by spider monkeys between successive foraging bouts. Each of these observations is consistent with model expectations and with the emergence of LW with $\mu = 2$.

The simulations of Boyer et al. [39] and Santos et al. [50] provide alternative explanations of the power-law scaling reported for spider monkeys. Boyer et al. [39] suggested that the movement patterns are, in fact, an emergent property of the deterministic way in which spider monkeys interact with their environment. The suggestion of Boyer et al. [39] stems from simulations of foragers that move so that the quantity $l_i/Q_i$ is minimal among previously unvisited patches $j$, where $l_i$ is the distance between patches $i$ (the current location of the forager) and patch $j$, and $Q_j$ is the quality of patch $j$. The quantity $l_i/Q_j$ roughly represents a cost/gain ratio for a movement segment. Boyer et al. [39] found that a power-law with $\mu = 2$ provides a good fit to the distribution of simulated movement lengths when resources are randomly distributed and when resource-qualities are power-law distributed with exponent $b$. The latter is consistent with observations of Ramos-Fernández et al. [29] because a power-law with $b = 2.6 \pm 0.2$ provides a good fit to the distribution of the tree sizes at the observation site [39] and because tree size is commonly regarded as one of the most accurate methods for estimating fruit abundance of tropical tress species [51]. The fitting $b = 2.6 \pm 0.2$ is consistent with previous studies of tree size distributions [50,51]. The simulations of Santos et al. [50] show that power-law distributions of path length will arise in animals that forage within narrow strips and always move to the nearest previously unvisited resource. This is pertinent because the spider monkeys studied by Ramos-Fernández et al. [29] resided within a strip-like region of land constrained by a lake.

It remains to be seen whether spider monkeys assess patch quality as in the current model, or minimize $l_i/Q_j$ as in the model of Boyer et al. [39], or whether some other mechanism [50] is operating. The findings reported here show how the basic ingredients of optimum foraging theory could enable foragers to execute LW searches with $\mu = 2$ for patches and thereby integrates search components at two different scales. It is hoped that this potential synthesis of searching for and within patches will be examined experimentally. It would also be interesting to examine
REFERENCES


J. R. Soc. Interface (2012)


