Reinforcement in Biology
Stochastic models of group formation and network construction

Qi Ma
Dissertation presented at Uppsala University to be publicly examined in Häggsalen, Ångström Laboratory, Lägerhyddsvägen 1, Uppsala, Thursday, January 10, 2013 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Empirical studies show that similar patterns emerge from a large number of different biological systems. For example, the group size distributions of several fish species and house sparrows all follow power law distributions with an exponential truncation. Networks built by ant colonies, slime mold and those are designed by engineers resemble each other in terms of structure and transportation efficiency. Based on the investigation of experimental data, we propose a variety of simple stochastic models to unravel the underlying mechanisms which lead to the collective phenomena in different systems. All the mechanisms employed in these models are rooted in the concept of selective reinforcement. In some systems the reinforcement can build optimal solutions for biological problem solving. This thesis consists of five papers. In the first three papers, I collaborate with biologists to look into group formation in house sparrows and the movement decisions of damsel fish. In the last two articles, I look at how shortest paths and networks are constructed by slime molds and pheromone laying ants, as well as studying speed-accuracy tradeoffs in slime molds’ decision making. The general goal of the study is to better understand how macro level patterns and behaviors emerges from micro level interactions in both spatial and non-spatial biological systems. With the combination of mathematical modeling and experimentation, we are able to reproduce the macro level patterns in the studied biological systems and predict behaviors of the systems using minimum number of parameters.

Keywords: reinforcement in biology, merge and split model, preferential attachment, reinforced random walk, network construction, shortest path problem, transport networks, ant algorithm, slime mould, physarum polycephalum, speed-accuracy tradeoff.

Qi Ma, Uppsala University, Department of Mathematics, Analysis and Applied Mathematics, Box 480, SE-751 06 Uppsala, Sweden.

© Qi Ma 2012

ISSN 1401-2049
urn:nbn:se:uu:diva-186989 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-186989)
List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


V Audrey Dussutour, Qi Ma, David Vogel, David J. T. Sumpter. (2012) Speed-accuracy tradeoffs and the construction of transport networks. (Manuscript)

Reprints were made with permission from the publishers.

As well as the papers presented in the thesis, I am an author of the following two papers completed during my PhD:

Contents

1 Introduction ................................................................. 7
2 Aggregation ................................................................. 8
   2.1 Preferential Attachment ............................................ 8
   2.2 Merge and split .................................................... 10
3 Network formation ....................................................... 13
   3.1 Biological and physical networks ................................. 13
      3.1.1 Markov Process and electric network analogy .......... 14
      3.1.2 Random walk on electric networks ....................... 15
   3.2 Trail formation .................................................... 17
      3.2.1 Active walker models in trail formation ................. 18
      3.2.2 Ant Colony Algorithm ...................................... 19
4 Paper summary ........................................................... 21
   4.1 Paper I ............................................................... 21
   4.2 Paper II ............................................................. 22
   4.3 Paper III ............................................................ 22
   4.4 Paper IV ............................................................. 23
   4.5 Paper V ............................................................. 24
5 Summary in Swedish ..................................................... 26
6 Acknowledgements ....................................................... 27
References ................................................................. 29
1. Introduction

Most biological systems show striking collective behaviors which cannot be predicted solely from the behavior of the individual units by which they are composed. For example, just looking around us we see people walking in a crowded street automatically form lanes to avoid bumping into each other; we see large bird flocks flying by in a spectacular manner; we see ant colonies gathering around food in a very short time; and we see lichens which produce interesting patterns on the surface of stones. These collective behaviors originate from interactions between individuals/parts within the systems and sometimes also with the interactions between the individuals and the environment. Linking the behaviors from different physical levels is therefore essential to understand collective phenomena. This thesis is devoted to give better insight of this problem through combining mathematical modeling with experiment.

One common property of biological systems is that they move in space. However, explicitly accounting for space is not always necessary in modeling the systems. For instance, when we model animal group size dynamics, for species that have relatively large moving speed such as birds and fish, movements are so rapid that space does not affect the interaction between groups. In other systems, for example those which involve physical networks, Euclidean space is crucial in their description and must be taken into account. The two specific problems I focus on in this thesis are (1) group formation dynamics and size distributions; (2) transport network construction inspired by ants and slime mold. The former is non-spatial and the later one has a spatial factor. I will address relative modeling background and applications of the two problems in separate chapters. Although there are many different aspects in these problems from a modeling perspective, there is a unifying theme in the mechanisms employed by the models. The models we used in this thesis are all based on the idea of reinforcement.
2. Aggregation

Aggregation is a very common scenario in biological systems, from pattern formation of bacteria, reaction-diffusion in metabolic processes, group decision making in nest selection by insects, to the motion of bird flocks, and evolution of human social communities. The whole branch of the study on collective behaviors won’t exist without the aggregation of individuals. Understanding the aggregation process helps us unravel the basic mechanisms employed by biological systems. One typical subject to the study of aggregation is how individuals come together to form groups.

Many animals live in groups. Individuals benefit from being a member of a group due to larger chance of getting food and higher survival probability if they encounter predators [38]. Individuals in groups also face disadvantages, such as the competition of sharing limited food resources [14]. Group formation is often a highly dynamical process and group sizes vary not only between species but also in species. Despite this, the group size distribution can be relatively stable. Empirical study shows some common character for the group size distribution from different species. For many mammals, fish and birds these are all highly skewed compared to the normal Gaussian distribution [32, 8, 31]. If similar group size distributions emerge from different biological systems, are there general rules for animal grouping? Previous studies gave a couple of related but different answers to this question, often using a class of models known as preferential attachment.

2.1 Preferential Attachment

One mechanism for describing aggregation is preferential attachment. Preferential attachment is a positive feedback mechanism in which, during a dynamical process, some quantity changes proportionally to the quantity itself. Take the size of commercial firms as example, if we consider the number of employees of the firm as the measure of firm size, big companies are more likely to hire more people to get bigger. There are many other names for the same "rich-get-richer" concept, such as "Yule process", "Gibrat’s principle", "cumulative advantage" and "Matthew effect".

Although the term "preferential attachment" was proposed by Barabási and Albert in their study of scale-free networks in 1999 [3], Yule first studied this process to explain the size distribution of genera in 1925 [41]. In his study, the sizes of the genera are measured by the number of species contained in
each genus. Yule modeled the growth of the number of species in each genus through time, using the following general idea. New species appear at every time step but species do not go extinct. Since new species are generated by mutation from old species, by assuming that mutation happens with the same probability for every species, genera which contain more species will generate more new species. When some new species appear, they belong to their parent’s genus. Therefore, larger genera produce more new species and become larger. Moreover, once there are certain number of new species, the mutation will be strong enough to generate one new species that does not belong to any old genera so that the new species forms a new genus. The model also assumes that at every time step, the number of newly generated species is constant, say $m + 1$, of which $m$ new species belong to old genera and the other one establishes a new genus.

With this simple model, Yule deduced that the size distribution of genera follows Power law distribution. However, due to the limitation of the theory of stochastic processes back at Yule’s time, the derivation he gave was rather complicated. Later, Simon used the master equation to analyze stochastic processes, similar to but more generalized than Yule’s model, which lead to skew distributions [36]. Newman also employed master equation to study Yule process and gave a clear analysis of a simplified model [30]. Here we briefly introduce Newman’s analysis.

As in Yule’s model, we assume that on each time step a new genus appears, so that at time $n$, there are $n$ genera in the system. Meanwhile, $m$ other new species are generated each step, so the total number of species at step $n$ is $n(m + 1)$. Denoting $p_{k,n}$ as the fraction of genera which have $k$ species when the total number of genera is $n$, then the number of genera that have $k$ species is $np_{k,n}$. Denote $k_i$ as the $i$th genus with $k$ species in it, the probability of a newly generated species lies in the $i$th genus is proportional to $k_i$. With normalization the probability is then $k_i / \sum_i k_i$, where $\sum_i k_i = n(m + 1)$ is the total number of species. Since every step there are $m$ new species joining the different genera, the average number of species gained by genus $i$ is $mk_i / n(m + 1)$. Therefore, the total expected number of new species gained by genera with size $k$ during step $n$ is

\[ \frac{mk}{n(m + 1)} np_{k,n} = \frac{m}{m+1} kp_{k,n} \tag{2.1} \]

This is also the decrease of $k$-sized genera at step $n$ since $k$-sized genera become $k + 1$-sized by gaining one new species. Meanwhile, the increase of $k$-sized genera is calculated the same way by $k - 1$-sized genera gaining one new species. Now we can write the master equation for the number of $k$-sized genera at step $n + 1$ as

\[ (n + 1)p_{k,n+1} = np_{k,n} + \left[ \frac{m}{m+1} [(k - 1)p_{k-1,n} - kp_{k,n}] \right] \tag{2.2} \]
For the special case where \( k = 1 \), the corresponding equation is

\[
(n + 1)p_{1,n+1} = np_{1,n} + 1 - \frac{m}{m+1} p_{1,n}
\]

(2.3)
since the assumption is that every time step there is a new genus appearing with one species. In the limit of long time when \( n \to \infty \), assuming the fraction of genera with size \( k \) is stabilized, i.e \( p_k = \lim_{n \to \infty} p_{k,n} \), then equation (2.3) has solution

\[
p_1 = \frac{m + 1}{2m + 1}
\]

(2.4)
and equation (2.2) becomes

\[
p_k = \frac{k - 1}{k + 1 + 1/m} p_{k-1}
\]

(2.5)
Iterating from \( p_k \) to \( p_1 \) and substitute equation (2.3) in, one can get

\[
p_k = (1 + 1/m) \frac{\Gamma(k) \Gamma(2 + 1/m)}{\Gamma(k + 2 + 1/m)} = (1 + 1/m) B(k, 2 + 1/m)
\]

(2.6)
according to the property of \( \Gamma \)-function \( \Gamma(a) = (a - 1) \Gamma(a - 1) \) and \( \Gamma(1) = 1 \).

Many biological systems can be described as the process of preferential attachment. For example, in *Saccharomyces cerevisiae* protein-protein interaction system, older proteins are likely to react with more proteins than the young ones and will gain more interactions when new proteins are introduced to the system [13], similar phenomena are also found in the evolution of metabolic systems of *Escherichia coli* for older enzymes [25]. Moreover, a scientific paper that has been cited many times has a higher chance to be cited again [21], famous actors who have already casted in many movies are more likely to get a new contract to play in more movies [2], web pages linked by large number of other pages gain another linkage with higher probability [4]. In human social interactions, the chance of a scientist acquiring a new collaborator increases with the number of his/her past collaborators [28] and the probability of a person having a new sex partner is proportional to the number of partners he/she had in the past 2-4 years [15]. Like in Yule’s model, all these systems having preferential attachment dynamics also exhibit a scale-free property, namely power law distribution.

### 2.2 Merge and split

In a classic review, Okubo argued that in a conserved system when the mean group size is well-defined, the distribution of animal groups should be exponential [32]. This result follows by a principle of maximum entropy. One
particular example summarized by Okubo is that if the probability of each individual joining a group increases linearly with the size of the group, namely larger group is preferred, with a corresponding linear splitting probability the final group size distribution will be truncated negative binomial or exponential. These distributions have shorter tails than power laws.

Merging and splitting does not rule out power laws, however. Bonabeau et al. showed that on the assumption that two groups merge into one larger group when they meet, an increasing system will stabilize at power law group size distribution with power exponent -3/2 [7, 8]. Here the individuals are not exactly preferentially attached to larger groups but the system exhibits a power law. Moreover, when splitting is introduced to the model and the total population is conserved, the power law distribution is still valid but with some cutoff function (usually exponential). The truncated power laws with power exponent between -0.7 and -1.8 give better fit of the data than the exponential distribution suggested by Okubo [8, 31]. However, in this model the biological explanation of the power exponent and the cutoff size were not properly investigated.

Niwa proposed a merge and split model which predicts that group size distributions are a combination of power law and exponential function. Specifically, he predicts a power law truncated by an exponential tail. The power exponent is $-1$ and the only parameter in the distribution is the cutoff size. This special truncated power law is known as the logarithm distribution. Distinguished from other models, in Niwa’s model, the cutoff size has a biological meaning and can be calculated from data, namely the mean group size experienced by an individual. The assumptions in Niwa’s model is simply as this: every group has equal probability to split to two uniformly distributed groups and then all groups move randomly, if two groups meet they merge to a larger group.

Gueron & Levin proposed a general framework for merge and split model with the merging and splitting rate related to the size of the groups. The merging probability of a group of size $x$ and a group of size $y$ is $\psi(x, y) = \alpha a(x)a(y)$, the splitting probability of a group of size $x$ is $\phi(x) = 2\beta a(x)$. They discussed three cases: (1) $a(x) = 1$; (2) $a(x) = x$; (3) $a(x) = 1/x$; and found that the general stationary solution for the group size distribution is

$$f(x) = 2\frac{\beta}{\alpha} \left( \frac{1}{a(x)} \right) e^{-\lambda x}$$

under the condition of $\lim_{x \to 0} a(x)/x^2 > 0$. $\lambda$ is a normalization factor. Note that when $a(x) = x$, larger groups will merge with higher probability, but they will split with higher probability too, the balance of merging and splitting gives approximately the same truncated power law distribution as in Bonabeau and Niwa’s papers [8, 31].

In paper II, I study the group formation dynamics and group size distributions by this type of merge and split model. The derivation of our model
allows us to link splitting and joining behavior to the exponent and truncation parameters in power laws.
3. Network formation

Networks are a hot and broad research topic in the area of complex systems. Lots of research has focused on the structures of networks including aspects such as mean distance, degree distribution, clustering effect, community identification, route factor etc [29, 5, 16]. These structures affect the function of networks. Modeling the formation of networks with different structures is a very practical and interesting subject, not only because it helps us explore the connection between structure and function of the network, but also it provides inspiration of how to construct efficient networks in diverse systems.

3.1 Biological and physical networks

Networks exist everywhere close to our daily life and in the nature. The formation of networks in assorted biological and physical systems has a fundamental influence on the efficiency of the system which they make up.

Neural networks are made up of neurons as vertices and synapses connecting neurons as edges. They transmit chemical or electrical signals between different parts of a body in order that certain physiological function can be performed. Blood vessels in animals and vascular systems in plants are another network example of great importance. The formation of the vascular networks determines the transportation efficiency of materials that are crucial to the survival of the organism [40, 34]. Angiogenesis (growth of new blood vessels) is a network formation problem that concerns cancer treatment. Blood vessels growing toward the tumor are a threat to the patient since once the vessel is connected to the tumor, the blood it provides will support tumor growth. Mycelial fungi and plasmodial slime moulds (myxomycetes) form network in their foraging space [17]. They spread hyphae at the edges, and the network is highly dynamical so the hyphae grow and move to forage. Amazingly, when they find multiple food sources, the slime mold (Physarum polycephalum) is able to form very efficient network to connect all the food sources [26].

Public transportation systems such as metro, railways and airline routes are human designed spacial networks to transport commodity, and travelers among vertices (stations/cities). The internet is also a widely studied physical network. It connects enormous numbers of computers and sends information from one computer to another. The electric power grid represents a network connected by power transmission lines and transport electric power with different voltage on different functional edges. Within these networks, the structure needs to be carefully arranged with regard to space, geography and also...
with respect to the transportation capacity and efficiency. Besides optimization, another subject interesting to study for these networks is the robustness or attack tolerance, which means how the failure of some vertices or edges will affect the function of the system [1, 37, 27]. The robustness of the network is of course also determined by the structure of the network.

Ant and human pedestrian trails are common examples of networks one can observe easily. These networks are built by large number of individual ants or pedestrians without central control, yet sometimes they are surprisingly efficient in terms of the average travel time by each ants or pedestrian [10, 18]. We will discuss the formation of these trails in more detail later in this section. Before this we define Markov processes and discuss their relation to electric networks and random walks.

3.1.1 Markov Process and electric network analogy

A stochastic process where the future state of the system only depends on the current state and does not depend on the previous state is a Markov process. Markov processes are memoryless since previous states have no effect on the future state. For example, short term weather prediction can be described as a Markov process. The weather tomorrow can be thought of as a stochastic state based solely on the weather condition today, but not the weather yesterday nor earlier. Queueing systems can also be modeled as Markov processes. Customers arrive in the queue randomly, the expected number of customers waiting in the queue at next time step is dependent on the number of customers in the queue now, not dependent on the number at precedent time steps. Other typical Markov processes includes Brownian motion, birth-death process, and the merge and split models we mentioned in the previous section.

Consider a Markov process \( X(t) \) with finite state space \( \phi \). Denote \( u_i(t) = P(X(t) = i) \) as the probability of the variable equal to \( i \) at time \( t \), \( q(i, j) \) as the joint transition probability from state \( i \) to \( j \). Then for \( i \in \phi \), the evolution of \( u_j(t) \) is

\[
\frac{d}{dt} u_j(t) = \sum_{i \in \phi} (u_i(t)q(i, j) - u_j(t)q(j, i)).
\]  

When the Markov process is reversible, namely the detailed balance condition is satisfied, in this case

\[
\pi(i)q(i, j) = \pi(j)q(j, i)
\]  

where \( \pi(i) \) represents the probability of \( \{X = i\} \) in equilibrium. The detailed balance condition means that in equilibrium the probability of the system in state \( i \) changing to state \( j \) equals to the probability of the system in state \( j \) changing to \( i \). Using the detailed balance condition, equation (3.1) can be
rewritten as
\[
\frac{d}{dt} u_j(t) = \sum_{i \in \Phi} \frac{1}{r(i, j)} \left( \frac{u_i(t)}{\pi(i)} - \frac{u_j(t)}{\pi(j)} \right)
\]  
(3.3)
where \( r(i, j) = (\pi(i) q(i, j))^{-1} = (\pi(j) q(j, i))^{-1} = r(j, i). \)

Kelly [22] made the analogy between this Markov process and electric networks. The system can be considered as charge which spreads itself over an electric network with a set of nodes \( \Phi \). Node \( i \) and \( j \) are connected by a wire of resistance \( r(i, j) \). Every node is attached to a capacitor connected to earth, the capacitance of the capacitor attached to node \( i \) is \( \pi(i) \). \( u_i(t) \) is the probability that the charge is on node \( i \) at time \( t \). Equation (3.3) then is the Kirchhoff’s equation stating that the change of expected charge at node \( i \) is the probability of the charge coming into the node subtract the probability of the charge leaving the node, namely the expected net flow of charge going to the node. Kelly further discussed that when the process is not just reversible but also symmetric, i.e. \( q(i, j) = q(j, i) \), then equation (3.1) becomes
\[
\frac{d}{dt} p_j(t) = \sum_{i \in \Phi} \frac{1}{r(i, j)} (p_i(t) - p_j(t))
\]  
(3.4)
where \( p_j(t) = u_j(t)/\pi(j) \) now represents the potential (charge/capacitance) on node \( j \). Note that in this case, the difference from equation (3.3) is \( \pi(i) = \pi(j) = \pi \), which means the charge will spread out evenly in \( \Phi \) in equilibrium. While in equation (3.3), the final expected charge on each node is determined by the distribution of \( \pi(i) \). The symmetric condition implies the resistance between two connected nodes are the same in the network, and the capacitance attached to each node has the same capacitance.

3.1.2 Random walk on electric networks

A random walk is an obvious example of Markov process. Earlier than Kelly, Kingman [23] also proposed the electric network analogy from a birth-death Markov process, but without the limitation of the reversibility. In this model, particles do a random walk among different sites. Assume that each site can contain no more than one particle, so the number of particles \( n_i \) on site \( i \) is either 0 or 1. A particle enters site \( i \) with rate \( \alpha_i \) if \( n_i = 0 \) and leaves site \( i \) with rate \( \beta_i \) if \( n_i = 1 \), site \( i \) and \( j \) swap their state with rate \( \gamma_{ij} \). Swapping state means that if site \( i \) is occupied \( (n_i = 1) \) and site \( j \) is empty \( (n_j = 0) \), then after the swapping, site \( j \) becomes occupied \( (n_j = 1) \) and \( i \) becomes empty \( (n_i = 0) \). If both of the sites are occupied or empty, their state do not change after the swapping. Suppose there are \( k \) sites and the state space \( \Omega \) consists of all \( k \)-vectors \( \mathbf{n} = (n_1, n_2, ..., n_k) \), \( n_i \in \{0, 1\} \). If \( \Omega \) is an irreducible class, which means for all \( \mathbf{n} \in \Omega \) there is equilibrium that the probability of state \( \mathbf{n} \) changing to other state equals to the sum of probabilities of other states changing to state
Figure 3.1. An example of the analogy. Figure taken from [12].

\[ n. \] According to irreducibility, after some calculation we get

\[ \alpha_i (1 - f_i) - \beta_i f_i - \sum_j \gamma_{ij} (f_i - f_j) = 0 \]  \hspace{1cm} (3.5)

where \( f_i \) denotes the probability that site \( i \) is occupied in equilibrium. The technical assumption required in the calculation is the symmetric swapping rate, i.e., \( \gamma_{ij} = \gamma_{ji} \). Kingman stated that we can then build up an electric network with nodes 1, 2, ..., \( k \) and a battery with voltage 1 based on the random walk process. Connect node \( i \) and \( j \) with a wire of resistance \( \gamma_{ij}^{-1} \) when \( \gamma_{ij} > 0 \), connect the positive terminal of the battery to node \( i \in A \) with a wire of resistance \( \alpha_i^{-1} \) (with \( \alpha_i > 0 \)) and connect the negative terminal with a wire of resistance \( \beta_i^{-1} \) (with \( \beta_i > 0 \)). Then \( f_i \) is the potential of node \( i \) determined by equation (3.5) which acts as the Kirchhoff equation here.

Doyle and Snell [12] discuss the analogy between random walk and electric network from a more individual/particle based point of view. First consider a two-dimensional lattice with some boundary nodes (see figure 3.1a). Boundary nodes marked with \( E \) are escaping points and boundary nodes marked with \( P \) are absorbing points. A particle does a random walk on the lattice. If it hits an escaping point, it will escape successfully from the lattice. If it hits an absorbing point it will stay there. Assume that the particle moves to each of the neighboring edges with equal probability. The question is then, “what is the probability function \( f_i \) that the particle will escape successfully if it starts from a random node \( i \)?” Now consider instead an electric network with the same set of nodes and edges (see figure 3.1b), all the escaping points are directly connected to the positive terminal of the one-voltage battery and all the absorbing points are directly connected to the negative terminal. The wires connecting each node are with identical resistance, then by Ohm’s law and Kirchhoff’s law we can identify the voltage function \( v_i \) for node \( i \). Doyle and Snell pointed out that \( f_i \) and \( v_i \) are both Harmonic functions and have the same boundary condition, and therefore they are the solutions for the same Dirichlet problem. Based on Uniqueness principle, \( f_i \) and \( v_i \) are equivalent to each other and if we calculate \( v_i \) we can answer our question about \( f_i \). If the whole structure
of the network is known, we can find \( v_i \) by solving a set of linear Kirchhoff’s equations. Conversely, by the equivalence of \( f_i \) and \( v_i \), we are able to obtain \( v_i \) from Monte Carlo simulations of random walk without knowing the whole structure of the network, i.e. only local information is needed for the calculation. This is extremely useful when the network is large or global information is unavailable. In paper IV and V, we applied Doyle and Snell’s random walk model for the migration of particles, and extended it to include reinforcement.

### 3.2 Trail formation

Trail formation in ant colonies and human pedestrians started to draw research attention in the late 1980s. Individual ants are tiny in size compared to their colonies’ foraging and nest migration area. How those large number of ants in a colony cooperate to explore and exploit food without centralized control is an interesting subject for collective decision making. Especially, since lots of ant species are able to build efficient (short average length to all food sources), yet energy saving (short total length, easy maintenance) trail systems between their nests and food sources [24]. For human pedestrians, the problem is slightly different since individuals are aware of the destinations of their motions. However, there are similar efficient and energy saving trail patterns observed in urban lawn areas [19]. Furthermore, both ants and human pedestrians form self-organized lane separation when individuals walk in two directions [18, 9]. Understanding individual based rules to form trails will not only help us explain the emergence of collective decision making, provide insight on public transport system design but also help us predict crowd behavior and prevent inefficient evacuation in urgent situations such as fires and earthquakes.

Lots of ant species are known to release and sense pheromones while walking. Individual ants sense the pheromone with their antennae and incline to walk in areas with higher pheromone concentration [20]. Pheromone dense areas attract more ants, therefore more pheromone will be released in those areas so they become even more attractive to other ants. Pheromone deposit is the way by which ants interact with others indirectly through the environment and produce positive feedback in the trail formation process. Pheromone concentration can be treated as a potential field in the same two-dimensional space as the ants spread out. For human pedestrians, stepping on the grass makes the vegetation grow slower and more comfortable for people to walk on, therefore pedestrians prefer to walk on balder area of the lawn. So the positive feedback or the interaction through the environment is done by walking on balder areas of the lawn, and in this case the baldness of the grass is the potential field. Naturally, most of the individual based study on trail formation used random walk model with bias towards higher potential area and individuals are able to interact with the potential field.
3.2.1 Active walker models in trail formation

A random walk where individual walkers can modify the environment is sometimes called an active walker model. Active walker models consist of two main parts: (1) how individuals move; (2) how the potential field is regulated by the motion of individuals.

In lots of active walker models, the individual rules of motion are determined by a combination of internal persistence of directions and external influence composed by the potential field. Usually the speed of each individual is set as a constant for the simplicity, while the direction of the velocity is

\[
e_i(r, t + \Delta t) = \frac{e_i(r, t) + f_i(Q, t)}{\|e_i(r, t) + f_i(Q, t)\|}
\]

where \(e_i(r, t)\) represents the direction of ant \(i\) at position \(r\) at time \(t\), \(f_i(Q, t)\) represents a function of pheromone concentration distribution \(Q(r, t)\) around \(r\) sensed by ant \(i\) at time \(t\). Although different models have differences on the definition of \(f\), generally all definitions of \(f\) are proportional to the gradient of the pheromone concentration, i.e. it will lead ant \(i\) to move towards a trail where the pheromone concentration is higher [10, 39, 19, 35, 9, 33].

Experimental results show that ants use their antennae to detect pheromone [20], but how exact individual ants respond to pheromones is not fully investigated by experiments yet. Denuebourg and co-workers proposed that when the ants face a bifurcation in their trail, their choice is decided by

\[
P_L = \frac{(L + h)^a}{(L + h)^a + (R + h)^a}
\]

where \(P_L\) is the probability that the ant will choose the left branch, \(L\) is the total pheromone concentration sensed by the left antenna, \(R\) is the total pheromone concentration sensed by the right antenna, \(h\) is a constant that acts like a threshold for the response to pheromone, \(a\) is a parameter indicating the nonlinearity of the response. The probability of choosing the right branch is then \(P_R = 1 - P_L\). With larger nonlinearity index \(a\), a slight difference between the concentration on left and right branch will lead to the ants to have strong bias. Several experiments using double bridge (Y-shape bridge) confirmed that some ants species have nonlinear response, i.e. \(a > 1\), to pheromones [10]. However, Denuebourg’s and other models [39, 35, 9, 19] all have the limitation that the possible turning angles are preset in the model assumptions with some random noise, while real ants move more freely in their directions.

Recently, Perna et al. [33] found that Weber’s law can be applied to explain the individual motions for Argentine ants. Weber’s law is a psychological law quantifying the perception of different stimuli, it holds if the difference between two stimuli is proportional to the ratio of "signal difference"/"average signal". In the case of ants, it can be understood as that individual ant’s response to pheromone, reflected by the turning angle \(\alpha\), i.e. the anti-clockwise
angle change from the previous direction, follows

$$\alpha = A \frac{L - R}{L + R}$$  \hspace{1cm} (3.8)$$

where $A$ is a constant, $L$ and $R$ is the same as in equation (3.7). It means that with the same concentration difference sensed by the left and right antenna, the ants are more sensitive to the difference when the total pheromone level is not very high. This conclusion is consistent with Denuebourg’s model when $h = 0$ and $a = 1$ in equation (3.7).

As to the regulation of the potential field/pheromone concentration, the previous studies all had similar assumptions. These assumptions are that individual ant leaves pheromone when they move and that the pheromone has constant evaporation rate. As a result only frequently used trails where pheromone is often renewed by passing-by ants can be maintained. The general evolution form for pheromone concentration $Q(r,t)$ is

$$\frac{dQ(r,t)}{dt} = \sum_{i=1}^{N} q(i,t) \delta(r - r_i(t)) - \gamma Q(r,t)$$  \hspace{1cm} (3.9)$$

where $q(i,t)$ is the amount of pheromone left by ant $i$ at time $t$, $\delta(\Delta r)$ is the Delta function. So the first term on the right hand side of above equation represents the total pheromone dropped by the ants passing by position $r$ during a small interval after time $t$. $\gamma$ is the evaporation rate of the pheromone. For simplicity, some models set $q(i,t)$ as a constant, which means each ant drop the same amount of pheromone all the time [10, 35, 33, 39]. Other models used a more realistic assumption that the amount of pheromone dropped by each individual per time step should decrease as they continue to walk in the space. However, this variation of $q$ does not change the main results of the models [19, 9]. Note that in the updating rule equation (3.9), the reinforcement of pheromone is proportional to the density of ants on the location. All the density reinforced active walker models produce trails that contain lots of loops which are not consistent with that observed in real ant trails. This problem can be solved by multi-pheromone deposit and treating outbound and inbound ant differently [19, 35], but this solution is not biologically realistic for many species. Another possible solution for this problem is to use current-reinforcement to update the pheromone concentration instead of the density-reinforcement as in equation (3.9). Details about the current-reinforced random walk model can be found in paper IV.

### 3.2.2 Ant Colony Algorithm

Ant Colony Optimization (ACO) algorithm is a classic application of active walker models. It was proposed by Marco Dorigo in his PhD thesis to solve computationally difficult problems. The basic idea of the ACO is positive
feedback between the random walk and pheromone concentration, as in the above models.

The original simple ACO is an iteration of the following process [11]. An individual ant is released from the nest. The probability that the ant will walk from current location $i$ to $j$ is proportional to

$$\frac{D_{ij}(t)^{\alpha} / l_{ij}^{\beta}}{\sum_{k \in E_i} D_{ik}(t)^{\alpha} / l_{ik}^{\beta}}$$

where $D_{ij}(t)$ is the pheromone concentration between $i$ and $j$ at step $t$, $l_{ij}$ is the distance between $i$ and $j$, $E_i$ is the set of possible locations one can move to from $i$. If ant $n$ reaches the food source, then the ant finishes the tour and the total length of the tour will be recorded as $L_n$. After all ants reached the food source by the random walk, the pheromone concentration on edge $ij$ will be updated as

$$D_{ij}(t) \leftarrow (1 - \lambda)D_{ij}(t) + \sum_k \Delta D_{ij}^n(t)$$

where $\lambda$ represents the evaporation rate of the pheromone and

$$\Delta D_{ij}^n(t) = \begin{cases} Q/L_n & \text{if ant } n \text{ uses edge } ij \text{ in its tour}, \\ 0 & \text{otherwise.} \end{cases}$$

where $Q$ is a constant and $L_n$ is the total distance traveled by ant $n$ in the tour. Hence only edges that had ants walked by will be reinforced and edges in shorter tours will have stronger reinforcement than edges in long tours.

The above process is iterated until all the ants take the same tour and the length of the tour does not change. Then we have the final tour as the shortest path between the nest and the food source. Lots of variations were made from the original ACO to solve optimization problems such as the traveling salesman problem, scheduling problems, assignment problems and set problems [6]. As such, ACO has been widely applied to different systems.
4. Paper summary

4.1 Paper I

*Understanding animal group-size distributions*

Michael Griesser, Qi Ma, Simone Webber, Katharine Bowgen, David J. T. Sumpter. (2011) *PloS one* 6(8), e23438.

One of the most striking aspects of animal groups is their remarkable variation in size, both within and between species. While a number of mechanistic models have been proposed to explain this variation, there are few comprehensive datasets against which these models have been tested. In particular, we only vaguely understand how environmental factors and behavioral activities affect group-size distributions. Here we use observations of House sparrows (*Passer domesticus*) to investigate the factors determining group-size distribution. Over a wide range of conditions, we observed that animal group sizes followed a single parameter distribution known as the logarithmic distribution. This single parameter is the mean group size experienced by a randomly chosen individual (including the individual itself). For sparrows, the experienced mean group size, and hence the distribution, was affected by four factors: morning temperature, place, behavior and the degree of food spillage. Our results further indicate that the sparrows regulate the mean group size they experience, either by groups splitting more or merging less when local densities are high. We suggest that the mean experienced group size provides a simple but general tool for assessing the ecology and evolution of grouping.

Sparrow groups are an example of a system which may be modeled by the merge and split dynamics discussed in the introduction. This paper gave us insight in linking parameters estimated from individual level behaviors to group size distribution patterns. The parameters needed to predict the distribution do not involve spatial factors, meaning they can be omitted the follow up modeling paper. The consistency of the data fitting to highly skewed (logarithmic) distribution implies that preferential attachment might be a possible mechanism in the group formation process. This leads us to the more detailed analysis in paper II.

Contribution: I analyzed the data, did the data fitting and model comparison, and participated in writing the paper.
4.2 Paper II

A first principles derivation of animal group size distributions


Several empirical studies have shown that the animal group size distribution of many species can be well fit by power laws with exponential truncation. A striking empirical result due to H-S Niwa is that the exponent in these power laws is one and the truncation is determined by the average group size experienced by an individual. In this paper we provide first principles derivations of such truncated power laws using a site-based merge and split framework. In particular, we investigate two such models. Firstly, we look at a model in which groups merge whenever they meet but split with a constant probability per time step. This generates a distribution similar, but not identical to that proposed by Niwa. Secondly, we propose a model, based on preferential attachment, that produces Niwa’s distribution exactly. Given the wide variety of merge and split mechanisms which generate Niwa’s distribution, our derivation helps explain why such distributions are so widely observed in nature. Our derivation also allows us to link splitting and joining behavior to the exponent and truncation parameters in power laws.

Based on the conclusion of paper I, we looked in more theoretical detail at the distributions found in the sparrows. We used the ‘merge and split’ model scheme to incorporate preferential attachment. From two different mechanisms, we obtained the same logarithmic distribution. This result confirms that preferential attachment is a mechanism that leads to a highly skewed distribution. The linkage we built up between splitting and joining behavior and the parameters in truncated power laws can in return help empirical study on prediction and parameter estimation.

Contribution: I participated in creating the model, performed model simulations and analytical calculations, and was responsible for writing the paper.

4.3 Paper III

Initiators, leaders and recruitment mechanisms in the collective movements of damselfish


Collective movements, where multiple individuals travel synchronously and cohesively, provide a crucial context for studying social behaviour. Here,
we examined the movements of groups of humbug damselfish between habitat patches. In doing so, we considered a range of individual- and group-level characteristics in relation to the collective behaviour of the group as a whole. Following a period of increasing activity during the pre-departure phase, groups moved between patches. The success or failure of any attempt to instigate a group movement was not solely dependent on the initiator’s behaviour, but on the behaviour of the group as a whole. Specifically, groups were more active and less cohesive prior to a successful initiation attempt in comparison to before a failed attempt, and where the behaviour of group members was similar to that of the initiator. Leadership was not related to dominance, but was consistent between trials. The probability of fish recruiting to the group movement initiative was an approximately linear function of the number of fish already recruited. Overall, these results are consistent with non-selective local mimetism, with the decision to leave based on a group’s, rather than any particular individual’s, readiness to leave.

In this paper, the agreement between data and model prediction on the leaving probability of fish indicated that the probability of individuals joining a group is proportional to the size of the group. This is essentially preferential attachment mechanism again, but this time for small groups of individuals. My role in this paper was to model this ‘mimetism’ mechanism using a Markov chain. Combining the three papers, we draw the conclusion that reinforcement through preferential attachment is possibly a very important rule applied by biological systems during group formation when spatial factor is not essential.

Contribution: I assisted in analysis of experimental data, created the Markov chain model and did the data fitting, and participated in writing the paper.

4.4 Paper IV

*Current-reinforced random walks for constructing transport networks*


Biological systems that build transport networks, such as trail-laying ants and the slime mould *Physarum*, can be described in terms of reinforced random walks. In a reinforced random walk the route taken by ‘walking’ particles depends on the previous routes of other particles. Here we present a novel form of random walk in which the flow of particles provides this reinforcement. Starting from an analogy between electrical networks and random walks, we show how to include current reinforcement. We demonstrate that current-reinforcement results in particles converging on the optimal solution of shortest path transport problems, and avoids the self-reinforcing loops seen in standard density-based reinforcement models. We further develop a variant
of the model that is biologically realistic, in the sense that the particles can be identified as ants and their measured density corresponds to those observed in maze-solving experiments on Argentine ants. For network formation we identify the importance of non-linear current reinforcement in producing networks which optimize both network maintenance and travel times. Other than ant trail formation, these random walks are also closely related to other biological systems, such as blood vessels and neuronal networks, which involve the transport of materials or information. We argue that current reinforcement is likely to be a common mechanism in a range of systems where network construction is observed.

In this and the next paper, our study systems are placed in two-dimensional space. The current-reinforced random walk model we proposed in the paper is an active walker model. There are many ways to implement reinforcement in active walker models. However, almost all previous models on trail/network formation problems (as we listed in the introduction) used density reinforcement. We used current-reinforcement instead. We combined the current-reinforcement with Doyle and Snell’s analogy between random walk and electric networks (see introduction) to build our model. This combination enables our model to reproduce dendritic network patterns from only local information. We then applied it to investigate the connection of network structures to its function in paper V.

Contribution: I participated in creating the model, performed model simulations and analysis, and was responsible for writing the paper.

4.5 Paper V

Speed-accuracy tradeoffs and the construction of transport networks

Audrey Dussutour, Qi Ma, David Vogel, David J. T. Sumpter, Manuscript.

One of the key challenges in the study of networks is linking structure to function. For example, how does the speed and accuracy with which information is transferred through a network determine the form of the network? We show that different strains of the slime mould Physarum polycephalum form different network structures, ranging from a diffuse network of thin links to a tree-like branching structure. Using a current-reinforced random walk model, we explain these different structures in terms of two model parameters: the strength and the degree of non-linearity in the reinforcement. These parameters are further shown to tune the speed and accuracy with which the network can detect resource gradients. We use a battery of experimental tests to show that Physarum strains with diffuse networks make more accurate but slower decisions and those with thick, trunk branches make faster less accurate decisions. Intermediate structures can also be found which are relatively fast and
accurate. The current reinforced random walk employed by the slime mould provides a tunable algorithm for decision-making, which may also apply in other systems where transport networks are constructed.

We examined the function of evolving networks based on the speed and accuracy of their decision making processes. The current-reinforcement random walk model established that the reinforcement intensity controls the speed of the diffusion of particles, and the degree of non-linearity in the reinforcement affects the accuracy of the decisions in a gradient field formed by resources. Together with paper IV, we can summarize that during the process of network formation, the form of reinforcement (current-based or density-based) determines if the networks are dendritic or not, while the intensity and nonlinearity of the current reinforcement specify the structure of the branching.

Contribution: I participated in the data analysis, creating the model, and performed model simulations plus analysis, and was responsible for writing the paper.
The first and most important person to thank is of course my advisor David Sumpter. Thank you, David, for all the help and advice you have given, also for the great opportunities you gave me to see the world and work with amazing people. During the past few years, I’ve learnt a lot from you, not only about mathematical modeling, but also about how to collaborate with others, how to manage time efficiently. You set up a role model, from who I see that it is actually possible for a human being to balance work and family perfectly. I really appreciate the encouragement you provide whenever it is needed. The positive energy I got from you, I believe, will support me for a long way.

I would also like to thank my second advisor Anders Johansson, for all the inspiring discussions and very helpful guidance on my work. Anders, your enthusiasm for math affects me, I am always extra awake when we discuss or when I am listening to your lectures. Hope I can keep learning from you.

I feel very lucky being a member of the collective behavior group. I should thank David here again for recruiting such a wonderful team. Boris, Daniel, Stam, Richard, Andrea, Natasha, Shyam, Arianna, Anna, Teddy, Mario, Betty, Rashid and Andreas, everyone of you has so colorful and unique character, I enjoy the days having you around, providing knowledge and friendship, intelligence sparkles and laughter.

Boris, you are the first one of us who accomplished the ‘mission (seems) impossible’, I am proud of you for your passion towards things you love, and thanks for the hands-on PhD checklist. Learning the do’s and (mostly) don’ts of surviving a PhD together with you and Daniel has been a great fun. Daniel, firstly, thanks for the effort you made on my ’Summary in Swedish’, with your input, I think it’s now Swedish enough. More than that, you were here from the beginning of my PhD, we’ve been through a lot together, such as visiting the best city in the world, and updating the name of the city constantly. You can always make people laugh without any effort, besides given you this gift, God clearly has chosen you as the envoy of the figure of 8. I am waiting to see that one day you convince the rest of the world that figure of 8 rules the universe. Richard, thank you for always being there when I showed up with a long face asking for help, the ’don’t panic’-s you said really helped sooth my nerve. And I have to say that I am jealous of your working efficiency. Stam and Natasha, thank you for the laughters and company, you guys made all my good memories of the trips to Japan and Brussels. An extra ’thank you’ to Stam for the countless times of very patient help, as well as the book and movie recommendations. Shyam, thanks for letting me occupy your computer.
in the last two weeks. Andrea, thank you for providing the code for graph reading and plotting, and the help on Xcode. Arianna, thanks for the help on hours of dumpling making, good luck to your exploration trip to Australia! Also, thank Anna, Teddy, Mario, and other members of the Rullan lunch gang, with your conversations, the lunches taste better, even on some of the Fridays!

Without the excellent collaborators, the papers in this thesis wouldn’t have existed. I would like to thank Michael Griesser, Toshiyuki Nakagaki, Atsushi Tero, Audrey Dussutour, Ashley Ward and Sara Bergek for their great job, which gave me different perspectives and inspirations on my research.

I would also like to thank Zengru Di for the mentoring during the three years of my master program, and being supportive and having faith in me for all these years. Also, thank you prof. Fan and my fellows: Liang Gao, Peng Zhang, Daqing Li, Liu Tian, Lida Xu.

Nono and Viga, thanks for the company in the past four years, and for indulging me as a freeloader of your Chinese food. It always feels warm when we share food and thoughts together. Carina, thanks for inviting me over for Xmas, and for sharing fun stories and politically incorrect thoughts, I approve you can ’fly’ now and then when we walk together.

Thank you, the reason that made me interested in Sweden in the first place. Thanks for once showing up in my life.

Mom and Dad, I am happy that I finally get through my PhD, I couldn’t have made it without your love and support, this thesis is dedicated to you, I love you.

Finally, I am thankful to the staff and faculty of the Department of Mathematics in Uppsala University for all your help and support during my PhD. I would also like to thank Chinese Scholarship Council and Lilewalchs Stipendium for financial support of my research.
References


January 2010.


