Dynamics of Agent-based Load Balancing on Grids

Yuanshi Wang
Department of Mathematics
Zhongshan University
Guangzhou, PRC
yuanshiwang@yahoo.com.cn

Jiming Liu
Department of Computer Science
Hong Kong Baptist University
Kowloon Tong, Hong Kong
jiming@comp.hkbu.edu.hk

ABSTRACT

We present a dynamic model of agent-based load balancing on grids. Our goal is to explore the effects of agents' strategies on the quality of load balancing. We show that under certain conditions, the model exhibits a known phenomenon that the load is perfectly balanced. This result is in agreement with recent experimental results on load balancing through Anthill. On the other hand, under general conditions, our model predicts unexpected phenomena that all the load balancing states tend to a unique steady state, where the load is not absolutely even balanced. The quality of load balancing, e.g., whether or not the load can be perfectly balanced, depends on agents' strategies. To measure the quality of load balancing, we define a total utility gain and further discuss the optimization of total utility gains.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence

General Terms

Design,Experimentations,Theory

Keywords

autonomous agents, load balancing, globally asymptotically stable, grid

1. INTRODUCTION

There is an explosion of interest in the study of agent-based grid computing in recent years. The interest is originally motivated from large-scale scientific computing, where the use of supercomputers is often impracticable. A number of less powerful computers on networks may be used to replace such supercomputers. Agent-based grid computing is the technology capable of exploiting and operating such distributed computers. After the workload is divided into a large number of independent tasks, the challenges in grid computing lie in load balancing.

The network of computers lacks fixed structures as its nodes come and go freely. Therefore the master-slave structure, where a centralized master disperses computing tasks to slaves, is not applicable to load balancing on grid. Complex adaptive systems (CAS) are considered to be suitable to the situation [4, 5, 1]. CAS are usually used to describe the dynamic and evolutionary behavior in biological and social systems. In CAS, single agents show simple strategies in isolation. However, as a large number of agents interact with each other, the system exhibits richer and more complex behavior than those single agents show in isolation. Ant colony is a kind of CAS [8]. Each ant's behavior is simple: wandering randomly or collecting objects. Based on the simple behavior, an ant colony can collect objects into piles. By using an artificial ant colony, Resnick [8] simulates single ant's strategies and exhibits three simple rules. The rules are: (i) ants wander around randomly, until it encounters an object; (ii) if it was carrying an object, it drops the object and continues to wander randomly; (iii) if it was not carrying an object, it picks the object up and continues to wander. While each agent follows the rules, Resnick shows that an ant colony can collect objects into piles.

The result of ant's strategies in [8] is collecting objects. Results of the inverse strategies is dispersing objects, which is considered in [5, 1]. In [5], a Peer-to-Peer (P2P) paradigm is designed to execute load balancing: ants which carry objects leave their objects behind only after they have wandered about randomly for a while without encountering object concentrations. As a result, objects are dispersed uniformly rather than collected into piles. In [5], three simple rules exhibited for load balancing are: (i) SearchMax: the ant wanders across the network, looking for overloaded nodes; (ii) SearchMin: the ant wanders across the network, looking for underloaded nodes; (iii) Transfer: the ant transfers tasks from the most overloaded node to the most underloaded one. Some experiments are done to show how the load balancing process is performed in [5]. In an experiment, a network of 100 idle nodes with 20 ants are given, and 10,000 jobs are generated in a single node. After 50 iterations, the load is perfectly balanced.

Microscopic simulation is an important method for understanding the interactions between agents. However, they offer only an empirical and indirect method for studying the global dynamic behavior of the system. Further studies, as is raised in [5], are needed to improve our understanding of ant colony's behavior on load balancing.

In the microscopic simulations of [5], the load is perfectly balanced. An open problem is followed that if a load can be perfectly balanced under general conditions, which factor is the most important to the load balancing. Macroscopic models can answer these kinds of questions. They can give a direct description of the collective behavior. They can describe the dynamic behavior in detail and have powerful predictions. Since there are only a few parameters in macroscopic models, properties can be analyzed quantitatively: some parameters can be proven to be important to the global characteristics of the system; some combination of parameters can result in an optimization of the system performance.

While we divide complex behavior into a few elements, a mathematical model of the behavior can be presented. Such kinds of macroscopic models are given in different fields: cyclic feedback [2], infection [3, 6], turning patterns [9] in chemical reaction-diffusion systems [7], etc. For coalition formation, a macroscopic model is given in [4]. In the model, the size and number of coalition are considered. Agents are supposed to wander on the networks randomly and know the relationship between the size of the coalition and the expected benefit of joining it. Agents' strategies lie in two elements: dispersion and aggregation. Based on the two simple elements and proper assumptions, a couple of differential equations are given to describe the coalition behavior. By numerical simulations of the differential equations, the dynamic behavior of coalition is shown. The steady state's stability and gain optimization are discussed. However, their results are not proven strictly.

Similar to the model of coalition formation in [4], we consider load balancing by using a macroscopic model. In the microscopic simulations of P2P, load balancing is performed by a large number of relatively simple autonomous agents. Each task is carried by a 'mobile' agent. Here, 'mobile' means the agent is active and free to choose nodes. From the viewpoint of relativity, we can suppose that each task is mobile. That is, each task is active and free to choose nodes. As a result, we'll consider each task as an agent in this paper. Therefore, load balancing is the opposite process of coalition formation. With this idea, we'll give a macroscopic model in this paper to describe the dynamic behavior of load balancing on grids. We show that under certain conditions, the model exhibits a known phenomenon that the load is perfectly balanced. This result is in agreement with recent experimental results on load balancing through Anthill. On the other hand, under general conditions, our model predicts unexpected phenomena that all the load balancing states tend to a unique steady state, where the load is not absolutely even balanced. The steady state is proven to be unique. A strict proof is given for the global stability of steady state if no more than three tasks are permitted at each node of grids. A total utility gain is defined to evaluate the load balancing performance and the optimization of gains is discussed. We'll prove that the average gain increases as the load size increases.

2. AGENT-BASED LOAD BALANCING

Consider a system of multiple mobile agents. Each agent has to obtain a service at a node on grids and each agent's service occupies the same time . Agents has to exploit nodes and queue for a service. We assume that agents don't know the global information about teams at nodes.

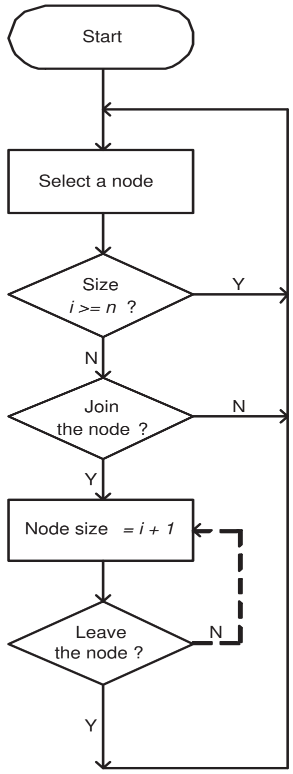

Figure 1: Single agent behavior: wandering, queuing and leaving. The broken line towards "Node size $i = i + 1$" means the agent stays in the node and the node size remains unchanged.

As small teams can save time, each agent moves among nodes to search for proper ones. When it encounters a service node, it leaves its trail behind and reads the trails left behind by other agents. It can leave or join the team in a service node, or form a new team if there is no other agents there. After joining a team, a single agent can also leave the team and moves to other nodes. Each agent makes decisions directly from its own experience and indirectly from the trails by other agents. Patience and other types of constraints limit the team length to a maximum size, that is, agents will not join the team whose size is greater than this maximum. Therefore, agents' behaviors can be decomposed into two elements: leaving and queuing. The primitive strategies for the two elements will determine the global dynamic behaviors of the system. Fig. 1 shows single agents' behavior. The load balancing mechanism outlined in the following procedure requires no global knowledge and minimal communication between agents. The mechanism will exhibit the formation of robust load balancing.

```
load balancing mechanism;
begin
initialize Single Agent (SA) position: SA queues at a
      team of size i0;
```

```
    if i0 > n,
        SA leaves the team of size i0;
        the size of the team becomes i0 − 1;
        goto START;
    endif
    if SA doesn't leave the team of size i0,
        goto EXIT;
    endif
the size of the team becomes i0 − 1;
START: SA wanders randomly on the network and meets
    a team of size i;
    if i > n or SA leaves the team of size i
        goto START;
    endif
    if SA joins the team of size i,
        the size of the team becomes i + 1;
    endif
    if SA leaves the team of size i + 1,
        the size of the team becomes i;
        goto START;
    endif
EXIT: SA queues in a team and doesn't leave;
end
```

To construct the initial mathematical model, we list below the primitive strategies and additional proper assumptions:

(i) Agents are homogeneous, in a sense that each agent occupies the same service time and follows the same strategies of leaving and queuing.

(ii) Initially all the agents queue in the same node.

(iii) Each agent is free to choose teams in nodes; teams are not.

(iv) It is beneficial for agents to join short teams; an agent can't join a team already of the maximum size.

(v) Agents encounter other agents and service nodes randomly.

(vi) There is no net change in the number of agents in the system.

3. MATHEMATICAL MODEL

We now present the mathematical model that describes the time evolution of load balancing. The quantities we are interested in studying load balancing are the size and number of teams. Let n denote the maximum team size, $x_i(t)$ denote the number of teams of size i at time t, $1 \leq i \leq n$. Initially the system consists of N agents which queue in the same node, $N \geq n$. If $N > n$, load balancing will tend to make the maximum team size not greater than n and a detailed explanation for the dynamic process is given below in Theorem 4 in this paper. Similar to the model in [4], we have the following model to describe how the number of teams of different sizes changes in time:

$$x_1' = 2b_2x_2 + \sum_{k=3}^{n} b_k x_k - 2a_1x_1^2 - x_1\sum_{k=2}^{n-1} a_k x_k,$$

$$x_i' = -b_ix_i + b_{i+1}x_{i+1} + a_{i-1}x_1x_{i-1} - a_ix_1x_i, 2 \leq i \leq n-1, \tag{1}$$

$$x_n' = -b_nx_n + a_{n-1}x_1x_{n-1},$$

where $\sum_{i=1}^{n} ix_i = N, a_j > 0, b_j > 0, x_j \geq 0, 1 \leq j \leq n$.

Here x_i' denotes $dx_i(t)/dt$, which is the rate of change of this number, $1 \leq i \leq n$. Parameters a_i, the attachment rate, denotes the rate at which an unaffiliated agent joins teams of size i. This parameter includes two factors: the rate at which an unaffiliated agent encounters a team of size i and the probability of joining the team. Parameters b_i, the detachment rate, gives the rate at which agents leave teams of size i. This parameter includes the probability that an agent in a team of size i leaves the team. As we have discussed in section 2, parameters a_i, b_i come from agents' experience after they do load balancing on the same network for a few times.

System (1) is in agreement with the nature of load balancing. In the first equation of (1), "$2b_2x_2$" shows that one team of size 2 becomes two teams of size 1 after an agent's leaving. "$-2a_1x_1^2$" shows that two teams of size 1 become one team of size 2 after an agent's queuing. For $3 \leq k \leq n$, "b_kx_k" shows that one team of size k becomes one team of size 1 and one team of size $k-1$ after an agent's leaving. For $2 \leq k \leq n-1$, "$-a_1x_1x_k$" shows that one team of size 1 and one team of size k become one team of size $k+1$ after an agent's queuing. Inductively, each expression in (1) can be explained in agreement with the nature of load balancing.

Since $\sum_{i=1}^{n} ix_i' = 0$, we have $\sum_{i=1}^{n} ix_i(t) = N$ as $t > 0$, which is what assumption (vi) means.

4. UNIQUENESS

Experiments in [5] show that a load can be perfectly balanced by ants after dozens of iterations. That is, the dynamic states of load balancing tend to a steady state and the steady state is perfect. Here, a strict proof is given to show that the steady state is unique and the perfectness of steady states depends on an agent's strategies of leaving and queuing.

Let $R_+^n = \{x = (x_1, x_2, ..., x_n) : x_i \geq 0, 1 \leq i \leq n\}$, $\Sigma_n = R_+^n \cap \{x : \sum_{i=1}^{n} ix_i = N\}$, and let $intR_+^n$ be the inner points of R_+^n, $int\Sigma_n = intR_+^n \cap \Sigma_n$. Let $x(t) = (x_1(t), x_2(t), ..., x_n(t))$ be a solution of (1) with $x_i(0) \geq 0, 1 \leq i \leq n$. It follows from $\sum_{i=1}^{n} ix_i(0) = N$ that there is $k, 1 \leq k \leq n$ such that $x_k(0) > 0$.

Consider the equilibrium (steady state) of (1) and let

$$c_1 = 1, c_i = \frac{a_1a_2...a_{i-1}}{b_2b_3...b_i}, 2 \leq i \leq n,$$

we have:

THEOREM 1. *There exists a unique equilibrium of (1) on Σ_n.*

PROOF. Let the right hand of equations in (1) be zero, we have:

$$x_i = \frac{a_{i-1}}{b_i}x_1x_{i-1}, 2 \leq i \leq n,$$

$$x_i = c_i x_1^i, 1 \le i \le n.$$

Since $\sum_{i=1}^{n} i x_i = N$, we have:

$$f(x_1) = \sum_{i=1}^{n} i c_i x_1^i - N = 0.$$

It follows from $f(0) = -N < 0, f(N) > 0, f^{'}(x_1) > 0$, as $x_1 > 0$ that there exists a unique solution of $f(x_1) = 0$ in $[0, N]$. $\square$

5. GLOBAL STABILITY

Under general conditions, whether or not the dynamic behavior of load balancing tends to a steady state is one of the important problems. Here, a strict proof of the above problems is given if no more than three agents are permitted to queue in the same node. For general conditions, a load is shown to be perfectly balanced by numerical integrations if the strategies of leaving and queuing (a_i, b_i) is well designed, which is similar to the perfect load balancing in the experiments of [5].

Let $n = 2$, system (1) becomes:

$$x_1^{'} = 2b_2x_2 - 2a_1x_1^2,$$

$$x_2^{'} = -b_2x_2 + a_1x_1^2, \tag{2}$$

where $x_1 + 2x_2 = N$.

Let $x_1 = N - 2x_2$, the second equation of (2) is:

$$x_2^{'} = g(x_2), \tag{3}$$

where $g(x_2) = a_1N^2 - (b_2 + 4Na_1)x_2 + 4a_1{x_2}^2, 0 \le x_2 \le N/2$.

Since $g(x_2) = 0$ has a unique solution on $[0, N/2]$, $g(0) > 0, g(N/2) < 0$, the unique equilibrium of (3) is globally asymptotically stable on $[0, N/2]$. That is:

THEOREM 2. *The equilibrium of (2) is globally asymptotically stable on Σ_2.*

The global portrait of (2) is shown in Fig. 2.

Let $n = 3$, system (1) becomes:

$$x_1^{'} = 2b_2x_2 + b_3x_3 - 2a_1x_1^2 - a_2x_1x_2,$$

$$x_2^{'} = -b_2x_2 + b_3x_3 + a_1x_1^2 - a_2x_1x_2, \tag{4}$$

$$x_3^{'} = -b_3x_3 + a_2x_1x_2,$$

where $x_1 + 2x_2 + 3x_3 = N$. Let $x_3 = (N - x_1 - 2x_2)/3$, the first two equation of (4) are:

$$x_1^{'} = h_1(x_1, x_2),$$

$$x_2^{'} = h_2(x_1, x_2), \tag{5}$$

where $0 \le x_1 + 2x_2 \le N$,

$$h_1(x_1, x_2) = 2b_2x_2 + b_3(N - x_1 - 2x_2)/3 - 2a_1x_1^2 - a_2x_1x_2,$$

$$h_2(x_1, x_2) = -b_2x_2 + b_3(N - x_1 - 2x_2)/3 + a_1x_1^2 - a_2x_1x_2.$$

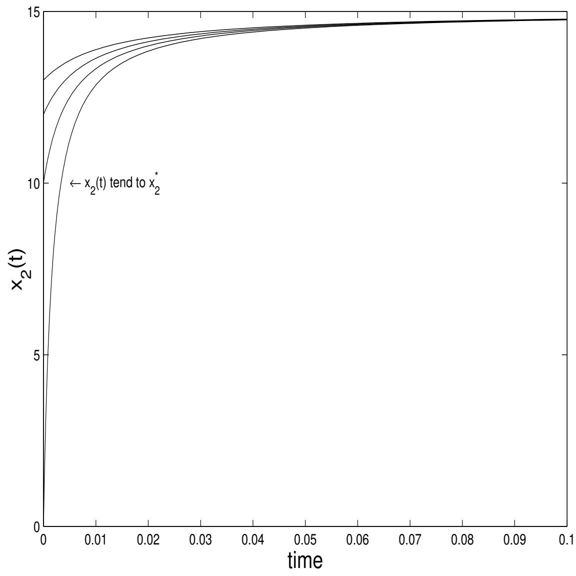

Figure 2: Let $n = 2, a_1 = 10, b_2 = 0.001$. The solutions of (1) tend to the steady state.

Since

$$\frac{\partial h_1}{\partial x_1} = -4a_1x_1 - a_2x_2 - b_3/3, \frac{\partial h_2}{\partial x_2} = -a_2x_1 - b_2 - 2b_3/3,$$

then $\frac{\partial h_1}{\partial x_1} + \frac{\partial h_2}{\partial x_2} < 0$. It follows from Dulac's criteria [10] that there is no periodic orbit of (5), we have:

LEMMA 1. *There is no periodic orbit of (4).*

LEMMA 2. *Suppose $x(t)$ is a solution of (1) such that there is $m, 1 \le m \le 3$, $x_m(0) = 0$. Then there is $\delta > 0$ such that*

$$x_i(t) > 0, t \in (0, \delta), 1 \le i \le 3.$$

PROOF. Suppose $x_1(0) = x_2(0) = 0, x_3(0) > 0$, then $x_1^{'}(0) > 0, x_2^{'}(0) > 0$, that is, there is $\delta > 0$ such that $x_i(t) > 0, t \in (0, \delta), 1 \le i \le 3$. Similar proof can be given for other cases. $\square$

It follows from Theorem 1 that there is a unique equilibrium of (5), the Jacobian matrix of (5) at the equilibrium is:

$$H = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} \end{pmatrix}.$$

Since $\frac{\partial h_1}{\partial x_1} = -b_3 - 4a_1x_1 - 2a_2x_2 < 0$, $\det H = b_2(b_3 + 3a_2x_2) + a_1x_1(6a_2x_1 + 4b_3) > 0$, the eigenvalues of H have negative real parts, then the equilibrium of (5) is asymptotically stable. Then we have:

LEMMA 3. *The equilibrium of (4) is asymptotically stable.*

It follows from Poincare-Bendixon Theorem [10] and Lemma 1-3 that

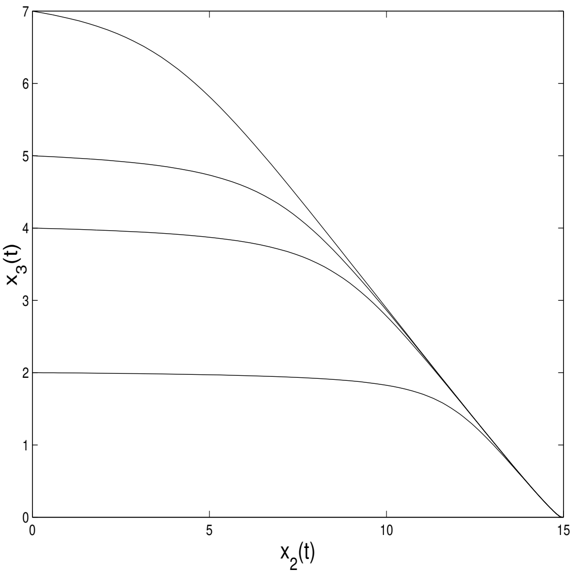

Figure 3: Let $n = 3, a_1 = b_3 = 10, a_2 = b_2 = 0.001$. The solutions of (1) tend to the steady state.

THEOREM 3. *The equilibrium of (4) is globally asymptotically stable on Σ_3.*

The global phase portrait of (4) is shown in Fig. 3.

For $n > 3$, the unique equilibrium is shown to be globally asymptotically stable on Σ_n by numerical integration. Let $n = 10, N = 100, a_i = b_{n-i} = 10, 0 \leq i \leq 4$; $a_j = b_{n-j} = 0.1, 6 \leq i \leq 9$; $a_5 = b_5 = 0.001$. In Fig. 4, solution $x(t)$ of (1) with initial values $x(0) = (0,0,0,0,0,0,0,0,0,10)$ tends to the steady state. In Fig. 5, the components $x_5(t)$ of $x(t)$ with five different initial values, tend to the same value and other components of $x(t)$ tend to different components of the steady state.

Consider the case that an agent wouldn't queue at a large team, we have:

THEOREM 4. *Suppose there is $k, 1 \leq k \leq n-1$ such that $a_i = 0, k \leq i \leq n-1$. Let $b_j > 0, 2 \leq j \leq n$, then each solution $x(t)$ of (1) satisfies*

$$\lim_{t\to+\infty} x_i(t) = 0, k+1 \leq i \leq n.$$

PROOF. It follows that the last $n-k$ equations of (1) are linear, they can be expressed as: $y^{'} = Ay$, where $y = (x_{k+1}, x_{k+2}, ..., x_n)^T$. The eigenvalues of A are

$$-b_{k+1}, -b_{k+2}, ..., -b_n,$$

which are negative. $\square$

By Theorem 4, we can explain that assumption (ii) in section 2 is applicable to system (1). In fact, we can give a model to describe the initial situation as all the $N(N > n)$ agents

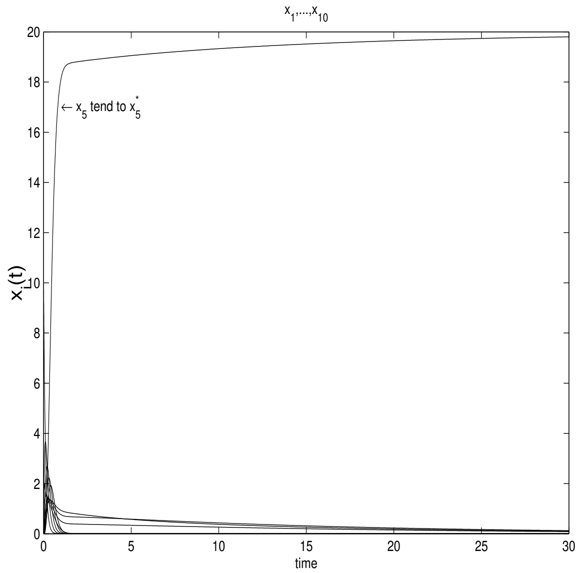

Figure 4: The phase portrait of x(t) of (1) tends to the steady state as $n = 10$, where the load is perfectly balanced.

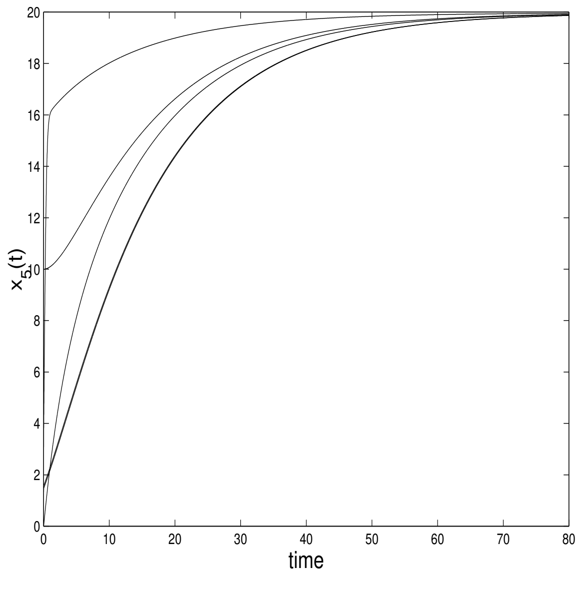

Figure 5: The components $x_5(t)$ of $x(t)$ with different initial values tend to the same value as $n = 10$.

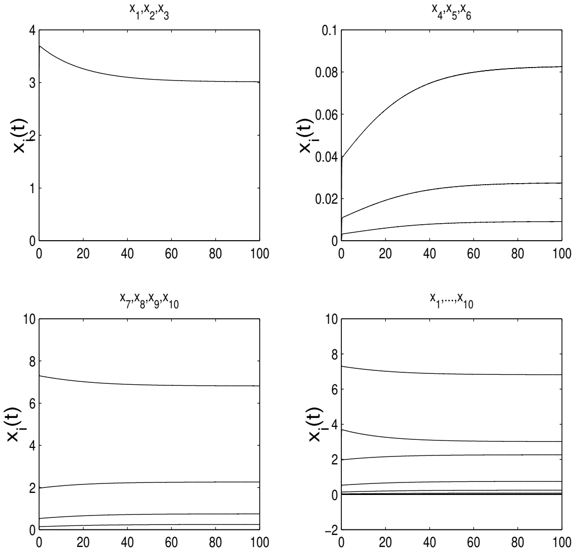

Figure 6: Let $a_1 = 0.0001, a_i = b_i = 9, 1 < i \le 10$, the load is not absolutely even balanced. x_2, x_3 is too small to be seen.

queue in the same service node:

$$x_1' = 2b_2x_2 + \sum_{k=3}^{N} b_k x_k - 2a_1x_1^2 - x_1\sum_{k=2}^{N-1} a_k x_k,$$

$$x_i' = -b_ix_i + b_{i+1}x_{i+1} + a_{i-1}x_1x_{i-1} - a_ix_1x_i, 2 \le i \le N-1, \quad (6)$$

$$x_N' = -b_Nx_N + a_{N-1}x_1x_{N-1},$$

where $x_i(0) = 0, 1 \le i \le N-1, x_N(0) = 1$.

By assumption (iv) in section 2, $a_i = 0, n \le i \le N-1$. It follows from Theorem 4 that

$$\lim_{t\to+\infty} x_i(t) = 0, n+1 \le i \le N,$$

that is, the dynamic behavior of system (1) determines that of system (6).

6. TOTAL UTILITY GAINS

It follows from Theorem 1 that different strategies lead to different steady states. Suppose agents have little experience. They only know that the team of size 1 is good and can't make any other decision. Then we may let $a_1 = 0.0001, a_i = b_i = 9, 1 < i \le 10$. Fig. 6 shows that there is a stable equilibrium of (1) and the distribution at the equilibrium is not perfect. Therefore, the quality of load balancing should be discussed.

The total utility gain and the average utility gain reflect the quality of system (1). Analysis about the gains can give the optimization of strategies. Suppose the unique equilibrium can be reached after some iterations. Let $x^* = (x_1^*, x_2^*, ..., x_n^*)$ be the unique equilibrium of (1) and each agent occupies the same service time $T, T > 0$. While there is no load balancing, the total waiting time of system (1) is $\sum_{i=1}^{N-1} iT = N(N-1)T/2$. While there exists load balancing and the equilibrium x^* is reached, the total waiting time of system (1) at x^* is $\sum_{i=2}^{n} i(i-1)Tx_i^*/2$. The total utility gain of system (1) is the saving time, that is, the difference between the two waiting times:

$$E = \frac{1}{2}N(N-1)T - \frac{1}{2}\sum_{i=2}^{n} i(i-1)Tx_i^*$$

$$= \frac{1}{2}N^2T - \frac{T}{2}\sum_{i=1}^{n} i^2 x_i^*.$$

By Theorem 1, we have:

$$x_i^* = c_i x_1^{*i}, 1 \le i \le n.$$

$$\sum_{i=1}^{n} i c_i x_1^{*i} = N.$$

It follows from the Implicit Function Theorem that x_1^* is a function of N and

$$\frac{dx_1^*}{dN} = \left(\sum_{i=1}^{n} i^2 c_i x_1^{*(i-1)}\right)^{-1} > 0.$$

Then

$$\frac{2}{T}\frac{dE}{dN} = \left(\sum_{i=1}^{n} i^2 c_i x_1^{*i}\right)^{-1} \sum_{i=1}^{n} i^2(2N-i)c_i x_1^{*i} > 0.$$

Let $F(E)$ be the average utility gain,

$$F(E) = \frac{E}{N} = \frac{NT}{2} - \frac{T}{2N}\sum_{i=1}^{n} i^2 c_i x_1^{*i},$$

then

$$\frac{2}{T}\frac{dF(E)}{dN} = 1 - N^{-2}\left(N\sum_{i=1}^{n} i^3 c_i x_1^{*(i-1)}\frac{dx_1^*}{dN} - \sum_{i=1}^{n} i^2 c_i x_1^{*i}\right)$$

$$= N^{-2}\sum_{i=1}^{n} i^2 c_i x_1^{*i} + \left(N\sum_{i=1}^{n} i^2 c_i x_1^{*i}\right)^{-1} \sum_{i=1}^{n} i^2(N-i)c_i x_1^{*i} > 0.$$

Therefore, we have:

THEOREM 5. *The total utility gain $E = E(N)$ and the average utility gain $F = F(E)$ of system (1) satisfy:*

$$\frac{dE}{dN} > 0, \frac{dF(E)}{dN} > 0.$$

Consider the uniform attachment-uniform detachment: $a_i = a, b_i = b, 1 \le i \le n$. Let $s = at, B = b/a$, then system (1) becomes (the variables are still the original variables):

$$x_1' = 2bx_2 + \sum_{k=3}^{n} bx_k - 2x_1^2 - x_1\sum_{k=2}^{n-1} x_k,$$

$$x_i' = -bx_i + bx_{i+1} + x_1x_{i-1} - x_1x_i, 2 \le i \le n-1, \quad (7)$$

$$x_n' = -bx_n + x_1x_{n-1}.$$

Let $y = \frac{x_1^*}{b}$, then the equilibrium $x^* = (x_1^*, x_2^*, ..., x_n^*)$ of (7) satisfies

$$x_i^* = by^i, i = 1, ..., n.$$

$$\sum_{i=1}^{n} iy^i = Nb^{-1}. \quad (8)$$

It follows from Theorem 1 and the Implicit Function Theorem that equation (8) defines a smooth function $y = y(b, N)$ and

$$\frac{\partial y}{\partial b} = -\frac{N}{b^2 \sum_{i=1}^{n} i^2 y^{i-1}} < 0,$$

$$\frac{\partial y}{\partial N} = \frac{1}{b \sum_{i=1}^{n} i^2 y^{i-1}} > 0.$$

By (8), $x_i^* = x_1^*, 2 \le i \le n$ if and only if $y = 1$, that is, $b = \frac{2N}{n(n+1)}$. It follows from $\frac{\partial y}{\partial b} < 0$ that the distribution of equilibrium x^* is monotone:

LEMMA 4. *The equilibrium x^* of (7) satisfies:*

(i) If $b < \frac{2N}{n(n+1)}$, then $x_1^* < x_2^* < ... < x_n^*$;

(ii) If $b = \frac{2N}{n(n+1)}$, then $x_1^* = x_2^* = ... = x_n^*$;

(iii) If $b > \frac{2N}{n(n+1)}$, then $x_1^* > x_2^* > ... > x_n^*$.

Since $y = y(b, N)$ is a smooth function, $E = E(b, N)$ is a smooth function of the parameters b, N. It follows from (8) that

$$E = E(b, N) = \frac{1}{2}N^2 T - \frac{bT}{2}\sum_{i=1}^{n} i^2 y^i$$

$$= \frac{1}{2}N^2 T - \frac{bT}{2}(nNb^{-1} - \sum_{i=1}^{n-1} iny^i + \sum_{i=1}^{n-1} i^2 y^i)$$

$$= \frac{N^2 T}{2} - \frac{nNT}{2} + \frac{bT}{2}\sum_{i=1}^{n-1}(n-i)iy^i. \quad (9)$$

Similar to the proof of Theorem 5, we have:

THEOREM 6. *The total utility gain $E = E(b, N)$ of system (7) satisfies:*

$$\frac{\partial F(E)}{\partial N} > 0, \frac{\partial E}{\partial N} > 0, \frac{\partial E}{\partial b} > 0.$$

An unexpected phenomenon is shown by Theorem 5 that both the average and total utility gain of (1) increases as the number of agents (tasks) increases. That is, the more the tasks, the better the efficiency of load balancing for both single agent and the global system.

An interesting distribution is given by Lemma 4 that the monotonicity of equilibrium x^* of (7) varies as b moves through a critical value: as b is less than the value, the smaller the size, the larger the number of teams; as b is larger than the value, the smaller the size, the smaller the number of teams. That is, as the attachment rate is small and the detachment rate is large, the number of short teams is greater than that of long teams; as the attachment rate is large and the detachment rate is small, the number of short teams is less than that of long teams.

It follows from Theorem 6 that both the average and total utility gains of (1) increase as the detachment rate increases and the attachment rate decreases in case of uniform attachment-uniform detachment. That is, the larger the detachment, the better the efficiency; the smaller the attachment, the better the efficiency.

7. CONCLUSIONS

We have shown that our model, which is based on strategies of leaving and queuing, predicts new characteristics of load balancing on grids which can't be given by microscopic simulations in [5, 1]: under general conditions, the load is not absolutely even balanced, see Theorem 1 and numerical integration in Fig. 6 for discussions. We prove that the steady state is unique. For low dimensions, we prove the steady state is globally asymptotically stable. The stable steady state determines the distribution of load balancing and is important in evaluating the strategies of leaving and queuing. By steady states, we show that both the average and the total utility gain increase as the number of agents (tasks) increases.

There are still some different elements that the model does not concern. For example, we do not introduce randomness into the model. Our proof about global stability is strict for low dimensions. For more general conditions, the global stability of steady state is still open although it seems right by computing simulations. Another focus on applying our results to a real situation is using microscopic simulations to determine the parameters in our model.

As a summary, our results directly describe the dynamic behavior of load balancing on grids and offer a new means of modeling and predicting load balancing behavior. Our predictions for load balancing are in agreement with recent experiments on the nature of load balancing in [5, 1]. More complex models about the load balancing, which consider a large number of elements, should be given in order to examine the nature of load balancing even more closely.

8. REFERENCES

[1] O. Babaoglu, H. Meling, and A. Montresor. Anthill: Aframework for the development of agent-based peer-to-peer systems. *Proc. of the 22th Int. Conf. on Distributed Computing Systems*, July 2002.

[2] J. Hofbauer and K. Sigmund. *Evolutionary games and replicator dynamics.* Cambridge University Press, 1998.

[3] F. C. Hoppensteadt and C. S. Peskin. *Mathematics in Medicine and the Life Sciences.* Springer-Verlag, 1992.

[4] K. Lerman and O. Shehory. Coalition formation for large-scale electronic markets. *Proceedings of the*

International Conference on Multi-Agent Systems, (ICMAS'2000), 2000.

[5] A. Montresor and H. Meling. Messor:load-balancing through a swarm of autonomous agents. *Technical Report , Dept. of Computer Science, University of Bologna*, UBLCS-02-08, May 2002.

[6] W. P. Nelson and A. S. Perelson. Mathematical analysis of delay differential equation model of hiv-1 infection. *Mathematical Biosciences*, 179:73–94, 2002.

[7] Z. Noszticzius, W. Horsthemke, W. D. McCormick, H. L. Swinney, and W. Y. Tam. Sustained chemical waves in an annular gel reactor: a chemical pinwheel. *Nature*, 329, 1987.

[8] M. Resnick. *Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds.* MIT Press, 1994.

[9] A. M. Turing. The chemical basis of morphogenesis. *Phil. Trans. Royal Soc. Lond.*, B327, 1995.

[10] Zhang and et al. *Qualitative Theory of Differential Equations*, volume 101. AMS, Providence, RI, translations of mathematical monographs edition, 1992.