Resource allocation on sparse graphs

K.Y. Michael Wong[1], David Saad[2] and Zhuo Gao[1,3]

[1]The Hong Kong University of Science & Technology, Clear Water Bay, Kowloon Hong Kong
[2]The Neural Computing Research Group, Aston University, Birmingham B4 7ET, UK
[3] Department of Physics, Beijing Normal University, Beijing 100875, China.

Key words: Statistical physics, inference, message passing, resource allocation, sparse graphs

Optimal resource allocation is a well known problem in the area of distributed computing [1, 2] to which significant effort has been dedicated within the computer science community. The problem itself is quite general and is applicable to other areas as well where a large number of nodes are required to balance loads/resources, such as reducing internet traffic congestion [3]. The problem has many flavours and usually refers, in the computer science literature, to finding practical algorithmic solutions to the distribution of computational load between computers connected in a predetermined manner. Many of the solutions are heuristic and focus on practical aspects (e.g., communication protocols).

The problem we are addressing is more generic and is represented by nodes of some computational power that should carry out some tasks. Both computational powers and tasks will be chosen at random from some arbitrary distribution. The nodes are located on a randomly chosen sparse graph of some connectivity. The goal is to migrate tasks on the graph such that demands will be satisfied while minimising the migration of (sub-)tasks. Decisions on messages to be passed are carried out locally. We focus here on the satisfiable case where the total computing power is greater than the demand, and where the number of nodes involved is very large. We analyse the problem using both the Bethe approximation and the replica method [4] of statistical physics. The latter will not be discussed in this extended abstract but can be found in [5]. Insights gained from the analysis give rise efficient message passing distributed algorithms for solving the problem with a modest computational cost. The approach is based on passing local information between nodes to facilitate decisions about the movement of tasks.

The Bethe approximation: We consider a typical resource allocation task on a sparse graph of N nodes, labelled $i = 1, .., N$. Each node i is randomly connected to c other nodes[1] and has a capacity Λ_i randomly drawn from a distribution $\rho(\Lambda_i)$. The objective is to migrate tasks between nodes such that each node will be capable of carrying out its tasks. The *current* $y_{ij} \equiv -y_{ji}$ drawn from node j to i is aimed at satisfying the constraint $\sum_j \mathcal{A}_{ij} y_{ij} + \Lambda_i \geq 0$, representing the 'revised' assignment for node i, where $\mathcal{A}_{ij} = 1/0$ for connected/unconnected node pairs i and j, respectively. To illustrate the statistical mechanics approach to resource allocation, we consider the load balancing task of minimising the energy function (cost) $E = \sum_{(ij)} \mathcal{A}_{ij}\phi(y_{ij})$, where the summation (ij) runs over all pairs of nodes, subject to the above constraints; $\phi(y)$ is a general even function of the current y.

When the connectivity c is low, the probability of finding a loop of finite length on the graph is low, and the Bethe approximation describes well the local environment of a node. In the approximation, a node is connected to c branches in a tree structure, and the correlations among the branches of the tree are neglected. In each branch, nodes are arranged in generations. A node is connected to an ancestor node of the previous generation, and another $c - 1$ descendent nodes of the next generation.

We derive a recursion relation for calculating the free energy of the system, the average asymptotic cost and the current distribution. Similar results have been obtained from the replica approach [5]. Although both derivations have been formulated for general cost functions, we concentrate on the particularly simple case of $\phi(y) = y^2/2$, where one can compare the obtained solutions with known results.

[1]Although we focus here on graphs of fixed connectivity, one can easily accommodate any connectivity profile within the same framework; the algorithms presented later are completely general.

Distributed algorithms: The local nature of the recursion relation we obtained points to the possibility that the network optimisation can be solved by message passing approaches, which have been successful in problems such as error-correcting codes [6] and probabilistic inference [7]. The major advantage of message passing is its potential to solve a global optimisation problem via local updates, thereby reducing the computational complexity. For example, the computational complexity of quadratic programming for the load balancing task typically scales as N^3, whereas capitalising on the network topology underlying the connectivity of the variables, message passing scales as N. An even more important advantage, relevant to practical implementation, is its distributive nature; it does not require a global optimiser, and is particularly suitable for distributive control in evolving networks.

However, in contrast to other message passing algorithms which pass conditional probability estimates of *discrete variables* to neighbouring nodes, the messages in the present context are more complex, since they are *functions* of the current y. We simplify the message to 2 parameters, namely, the first and second derivatives of these functions. For the quadratic load balancing task, it can be shown that the message functions are piecewise quadratic with continuous slopes. This makes the 2-parameter message a very precise approximation. The message passed from node j to i, (A_{ij}, B_{ij}) becomes

$$A_{ij} \leftarrow -\mu_{ij}, \quad B_{ij} \leftarrow \Theta(-\mu_{ij}) \left[\sum_{k \neq i} \mathcal{A}_{jk} (\phi''_{jk} + B_{jk})^{-1} \right]^{-1}, \tag{1}$$

$$\text{where} \quad \mu_{ij} = \min \left[\frac{\sum_{k \neq i} \mathcal{A}_{jk} [y_{jk} - (\phi'_{jk} + A_{jk})(\phi''_{jk} + B_{jk})^{-1}] + \Lambda_j - y_{ij}}{\sum_{k \neq i} \mathcal{A}_{jk} (\phi''_{jk} + B_{jk})^{-1}}, 0 \right], \tag{2}$$

with ϕ'_{jk} and ϕ''_{jk} representing the first and second derivatives of $\phi(y)$ at $y = y_{jk}$ respectively. The forward passing of the message from node j to i is then followed by a backward message from node j to k for updating the currents y_{jk} according to $y_{jk} \leftarrow y_{jk} - \frac{\phi'_{jk} + A_{jk} + \mu_{ij}}{\phi''_{jk} + B_{jk}}$. For the quadratic load balancing task considered here, an independent exact optimisation is available for comparison. The Kühn-Tucker conditions for the optimal solution yields

$$\mu_i = \min \left[\frac{1}{c} \left(\sum_j \mathcal{A}_{ij} \mu_j + \Lambda_i \right), 0 \right]. \tag{3}$$

Numerical results: We exploit both the theoretical framework developed using methods of statistical physics and the message passing techniques mentioned above to study properties of the resource allocation problem with a quadratic cost function. The iterative solution of the free energy is obtained numerically using the recursion relation obtained from the Bethe approximation. We generate 1000 nodes at each iteration with capacities randomly drawn from the distribution $\rho(\Lambda) = \mathcal{N}(\langle \Lambda \rangle, 1)$, and each is being fed by $c-1$ nodes randomly drawn from the previous iteration.

Figure 1(a) illustrates the current distribution for various average capacities. The distribution $P(y)$ consists of a delta function component at $y = 0$ and a continuous component whose breadth decreases with average capacity. The fraction of links with zero currents increases with the average capacity. Hence at a low average capacity, links with nonzero currents form a percolating cluster, whereas at a high average capacity, it breaks into isolated clusters. As shown in Fig. 1(b), both the analytic results and the message passing algorithm Eq.(1) yield excellent agreement with the iteration of Eq.(3). Besides the case of $c = 3$, Fig. 1(b) also shows the simulation results of the average energy for $c = 4, 5$, using both Eqs. (1) and (3). We see that the average energy decreases when the connectivity increases. This is because the increase in links connecting a node provides more freedom to allocate resources. When the average capacity is 0.2 or above, an exponential fit $\langle E \rangle \sim \exp(-k \langle \Lambda \rangle)$ is applicable, where k lies in the range 2.5 to 2.7. Remarkably, multiplying by a factor of $(c-2)$, we find that the 3 curves collapse in this regime of average capacity, showing that the average energy scales as $(c-2)^{-1}$ in this regime (inset).

Further properties of the optimised networks have been studied by simulations, and will be presented elsewhere. Here we merely summarise the main results: (a) When the average capacity drops below

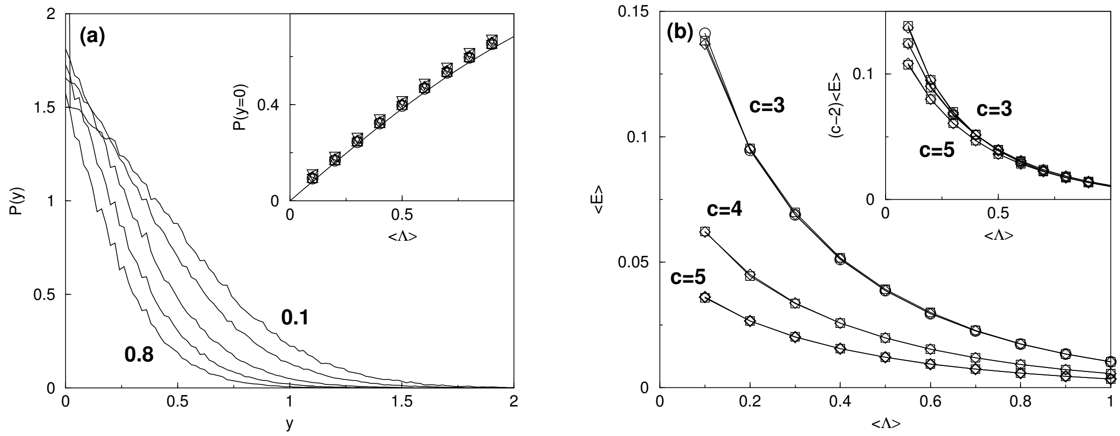

Figure 1: Results for system size $N = 1000$ and $\phi(y) = y^2/2$. (a) The distribution $P(y)$ obtained by iterating the recursive equations to steady states for $\langle \Lambda \rangle = $0.1, 0.2, 0.4, 0.6, 0.8 from right to left. Inset: $P(y=0)$ as a function of $\langle \Lambda \rangle$. Symbols: $c=3$ ($\bigcirc$) and ($\Box$), $c=4$ ($\Diamond$) and ($\triangle$), $c=5$ ($\triangleleft$) and (∇); each pair obtained from Eqs. (1) and (3) respectively. Line: $\mathrm{erf}(\langle \Lambda \rangle/\sqrt{2})$. (b) Mean cost $\langle E \rangle$ as a function of $\langle \Lambda \rangle$ for $c = 3, 4, 5$. Symbols: results obtained by iterating the recursive equations to steady states ($\bigcirc$), Eq.(1) ($\Box$), and Eq. (3) ($\Diamond$). Inset: $\langle E \rangle$ multiplied by $(c-2)$ as a function of $\langle \Lambda \rangle$ for the same conditions.

0.1, the energy rises above the exponential fit applicable to the average capacity above 0.2. (b) The fraction of links with zero currents increases with the average capacity, and is rather insensitive to the connectivity. Remarkably, except for very small average capacities, the function $\mathrm{erf}(\langle \Lambda \rangle/\sqrt{2})$ has a very good fit with the data. Indeed, in the limit of large $\langle \Lambda \rangle$, this function approaches the fraction of links with both vertices unsaturated, that is, $[\int_0^\infty d\Lambda \rho(\Lambda)]^2$. (c) The fraction of unsaturated nodes increases with the average capacity, and is rather insensitive to the connectivity. In the limit of large average capacities, it approaches the upper bound of $\int_0^\infty d\Lambda \rho(\Lambda)$, which is the probability that the capacity of a node is non-negative. (d) The convergence time of both the Bethe recursion equations and Eq. (3) follows a power-law dependence on the average capacity when the average capacity is 0.2 or above; the exponent is ranging from -1 for $c=3$ to -0.8 for $c=5$ for Eq. (3), and being about -0.5 for $c=3, 4, 5$ for Eq. (1). When the average capacity decreases further, the convergence time deviates above the power laws.

Summary: We have studied a prototype problem of resource allocation on sparsely connected networks using the replica method and the Bethe approximation. The resultant recursion relation leads to a message passing algorithm for optimising the average energy, which significantly reduces the computational complexity of the global optimisation task and is suitable for online distributive control. The suggested 2-parameter approximation produces results with excellent agreement with the original recursion relation. We have considered the simple but illustrative example of a quadratic cost function, where both Bethe recursion equations and message passing algorithm show remarkable agreement with the exact result. The suggested simple message passing algorithm can be generalised to more realistic cases of nonlinear cost functions and additional constraints on the capacities of nodes and links. This constitutes a rich area for further investigations with many potential applications.

Acknowledgements This work is partially supported by research grants HKUST6062/02P and DAG04/05.SC25 of the Research Grant Council of Hong Kong, and by EVERGROW, IP No. 1935, IST Priority, EU FP6.

References

[1] L. Peterson, B.S. Davie, *Computer Networks: A Systems Approach*, Academic Press, San Diego CA, (2000)
[2] Y.C. Ho, L. Servi and R. Suri, *Large Scale Systems* **1** 51 (1980)
[3] S. Shenker, D. Clark, D. Estrin and S. Herzog, *ACM Computer Communications Review* **26** 19 (1996)
[4] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing*, OUP, Oxford, UK (2001)
[5] K.Y.M. Wong, D. Saad and Z. Gao, submitted (2005)
[6] M. Opper and D. Saad *Advanced Mean Field Methods*, MIT press (2001)
[7] D.J.C. MacKay, *Information Theory, Inference and Learning Algorithms*, CUP UK(2003)