2015 International Symposium on Nonlinear Theory and its Applications
NOLTA2015, Kowloon, Hong Kong, China, December 1-4, 2015

# Optimal Facility Location with Message Passing Algorithm

Bo Li†, K. Y. Michael Wong† and Chi Ho Yeung‡

†Department of Physics, The Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong SAR, P. R. China
‡Department of Science and Environmental Studies, The Hong Kong Institute of Education
Tai Po, Hong Kong SAR, P. R. China
Email: bliaf@ust.hk, phkywong@ust.hk, chyeung@ied.edu.hk

**Abstract**—In this paper we study a facility location problem in the form of transportation network. Based on the cavity method developed in statistical physics, we derive a message passing (MP) algorithm for solving this problem. Moreover, we develop an MP guided decimation strategy to facilitate convergence of the algorithm in loopy networks. Optimal locations of facilities in tilted square lattices and the corresponding scaling behaviours in different regimes are also discussed.

## 1. Introduction

Optimizing the locations of facilities is crucial in communication and supply networks. On one hand, a comprehensive coverage on a relevant region with facilities, e.g., outlets or sensors, is essential in supply networks or sensor networks [1, 2]. On the other hand, maintaining high level of coverage can be costly since extensive resources need to be transported to destinations to satisfy the demands. It is of great practical importance to strike a balance between expanding network coverage and increasing transportation cost. Traditional methods from operations research rely on global optimizers like linear or quadratic programming, which can be computationally demanding [3].

In our approach, we borrow the cavity method from the studies of disordered systems by statistical physics and derive a local and efficient message passing algorithm for solving instances of the problem [4]. The cavity approach applied to transportation network is exact on tree graphs but an approximation in loopy graphs [5, 6]. Interestingly, we found that a simple decimation strategy can make the algorithm converge and yield satisfactory results. Based on this approach, we identify the optimal facility locations and study the behaviours of the systems in different regimes.

## 2. The Facility Location Model

Similar to [7], we consider a connected network of $N$ nodes. A node can be in active or idle state, denoted as $n_i = 1, 0$ respectively. An active node corresponds to an outlet that serves the surrounding area. The commodities are supplied from a single warehouse and need to be transported to the outlet. We denote the warehouse as a terminal node $T$, and the commodity flow from node $j$ to node $i$ as $y_{ij}(= -y_{ji})$ which is a continuous variable. The cost function or energy function to be minimized is

$$E = U \sum_i (1 - n_i) + V \sum_{(i,j)} n_i n_j + \sum_{(i,j)} \frac{1}{2} R_{ij} y_{ij}^2. \tag{1}$$

In the first term, an idle node with the neighbourhood unserved is penalized by a cost $U$. In the second term, a short range repulsion of intensity $V$ is introduced for adjacent active nodes, which tends to spread the outlets and enhance the coverage. This corresponds to a tendency to avoid setting up outlets redundantly. The third term is the transportation cost with weight $R_{ij}$ analogous to resistance in resistor networks. Such convex nonlinear transportation cost tends to distribute flows on different routes to mitigate congestion in supply networks [8]. To satisfy the demand of each outlet, the variables are subject to the flow conservation constraints $\sum_{j\in\partial i} y_{ij} = n_i$ for $i \neq T$ where $\partial i$ is the set of nodes adjacent to node $i$.

It is interesting to study the behaviours of the systems which balance the objectives of saving transportation costs and increasing network coverage.

## 3. Message Passing Algorithm

In this section, we apply the cavity method to tackle the problem. For latter convenience, let $d_i = |\partial i|$ be the degree of node $i$ and define $f_{ij}(y_{ij}, n_i, n_j) = 1/2 R_{ij} y_{ij}^2 + V n_i n_j + U(1-n_i)/d_i + U(1-n_j)/d_j$, based on which one can rewrite the cost function in Eq. (1) as

$$E = \sum_{(i,j)} f_{ij}(y_{ij}, n_i, n_j), \tag{2}$$

Consider a node $j$ and one of its neighbours node $i$ as the ancestor. Denote the set of nodes adjacent to node $j$ excluding node $i$, or the descendants of node $j$, as $\partial j\backslash i$. We write the summation over descendants $\sum_{k\in\partial j\backslash i}$ as $\sum'_k$. Based on the tree approximation of the graph, breaking the edge $(i, j)$ will result in two sub-trees that are locally independent of each other, which is a good approximation in the absence of short loops. The cavity energy, i.e., the optimal energy

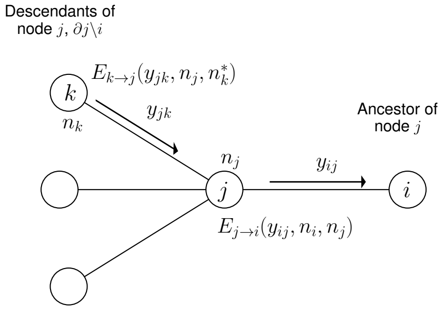



Figure 1: An illustrative diagram for Eq. (3). The cavity energy $E_{j\to i}(y_{ij}, n_i, n_j)$ is obtained by collecting and optimizing the cavity energies of the descendants of node $j$.

of the sub-tree terminated at node $j$ is,

$$\begin{aligned} E_{j\to i}(y_{ij}, n_i, n_j) &= \min_{\{y_{jk}|\sum_k' y_{jk}=n_j+y_{ij}\}} \Big\{ \sum_k{}' E_{k\to j}(y_{jk}, n_j, n_k^*) \Big\} \\ &\quad + f_{ij}(y_{ij}, n_i, n_j), \end{aligned} \tag{3}$$

where the optimal value $n_k^*$ of $n_k$ was found in the update of node $k$. An illustrative diagram is shown in Fig. 1. The cavity energy $E_{j\to i}(y_{ij}, n_i, n_j)$ depends on the cavity energies of the descendants of node $i$, which indicates that a recursion process to determine the energy of all the sub-trees is needed.

As in [5], we approximate the cavity energy up to second order in $y_{ij}$ and express it in the following quadratic form [6]

$$E_{j\to i}(y_{ij}, n_i, n_j) = \frac{1}{2} a_{ij}(y_{ij} - \tilde{y}_{ij})^2 + d_{ij}. \tag{4}$$

We refer to $\tilde{y}_{ij}$ as the preferred cavity flow of edge $(i, j)$, which is favoured by the cavity energy $E_{j\to i}$. We term $a_{ij}$ as the cavity resistance of the sub-tree terminated at node $i$ whose meaning will be clear in the following expressions.

By solving the cavity energy $E_{j\to i}$ in Eq. (3) and expressing it in the form of Eq. (4), we can identify the messages as follows,

$$a_{ij} = R_{ij} + \frac{1}{\sum_k' a_{jk}^{-1}}, \tag{5}$$

$$\tilde{y}_{ij}(n_j) = \frac{-n_j + \sum_k' \tilde{y}_{jk}}{1 + R_{ij}\sum_k' a_{jk}^{-1}}, \tag{6}$$

$$\begin{aligned} d_{ij}(n_i, n_j) &= \frac{(-n_j + \sum_k' \tilde{y}_{jk})^2 R_{ij}}{2(1 + R_{ij}\sum_k' a_{jk}^{-1})} + \sum_k{}' d_{jk}(n_j, n_k^*) \\ &\quad + V n_i n_j + \frac{U}{d_i}(1 - n_i) + \frac{U}{d_j}(1 - n_j). \end{aligned} \tag{7}$$

It is interesting to draw an analogy between the cavity resistance $a_{ij}$ of the sub-tree and the connections of electric circuit. One first connects the cavity resistors of the descendants of node $j$ in parallel, yielding the parallel resistance $(\sum_k' a_{jk}^{-1})^{-1}$, and further connects it to edge $(i, j)$ with resistance $R_{ij}$ in series, which gives the cavity resistance of the sub-tree as expressed in Eq. (5). Once the network structure is given, all the cavity resistances are determined without knowing the state of nodes.

From Eq. (6), the preferred cavity flows of edge $(i, j)$ is obtained by first collecting the upstream preferred cavity flows $\tilde{y}_{jk}$, subtracting the resource absorbed in node $i$ of amount $n_i$, which is then re-weighted by a factor smaller than one, i.e., $(1 + R_{ij}\sum_k' a_{jk}^{-1})^{-1}$. The re-weighting factor is due to addition of a new edge of resistance $R_{ij}$ such that the new sub-tree favours a smaller cavity flow on edge $(i, j)$. The iterative updates of $y_{ij}$ and $d_{ij}$ depend on the choices of states of the nodes.

The updates of the state of node $n_j$ and the corresponding messages are determined by minimizing the full energy, defined by joining the cavity energies of the two disconnected sub-trees and subtracting the double-counted energy $f_{ij}$,

$$\begin{aligned} E_{ij}^{\text{full}}(y_{ij}, n_i, n_j) &= E_{j\to i}(y_{ij}, n_i, n_j) + E_{i\to j}(-y_{ij}, n_i, n_j) \\ &\quad - f_{ij}(y_{ij}, n_i, n_j). \end{aligned} \tag{8}$$

We then find out the optimal flow $y_{ij}^*$ and the optimal state of node $n_j$ at the current step by

$$y_{ij}^*(n_i, n_j) = \arg\min_{y_{ij}} E_{ij}^{\text{full}}(y_{i,j}, n_i, n_j) = \frac{a_{ij}\tilde{y}_{ij} - a_{ji}\tilde{y}_{ji}}{a_{ij} + a_{ji} - R_{ij}}, \tag{9}$$

$$n_j^* = \arg\min_{n_j} E_{ij}^{\text{full}}(y_{i,j}^*, n_i, n_j). \tag{10}$$

Special considerations are needed for the messages from terminal nodes and dangling nodes, which we omit here. The resulting message passing algorithm is summarized in Algorithm 1.

Initialize the messages and states of nodes;
**for** $t = 1$ *to* $t_{max}$ **do**

- Randomly pick a node $j$ and one of its neighbours node $i$ as the ancestor;
- Collect the messages $a_{jk}$, $\tilde{y}_{jk}$ and $d_{jk}$ for $k \in \partial j\backslash i$, and compute $a_{ij}$, $\tilde{y}_{ij}$, $d_{ij}$ by Eq. (5)-(7);
- Compute $y_{ij}^*$ and $n_j^*$ by Eq. (9)-(10); update the messages of $E_{j\to i}$ as $a_{ij}$, $\tilde{y}_{ij}(n_j^*)$, $d_{ij}(n_i, n_j^*)$; update the state of node $j$ as $n_j^*$;
- If the messages converge, exit and output results of all $y_{ij}^*$ and $n_i^*$.

**end**

**Algorithm 1**: The Message Passing Algorithm

## 4. Results on Tilted Square Lattices

In this section, we use the above message passing algorithm to identify the optimal facility locations on a tilted square lattice constructed layer by layer starting from the terminal node. We assume $R_{ij} = 1$ for all edges.

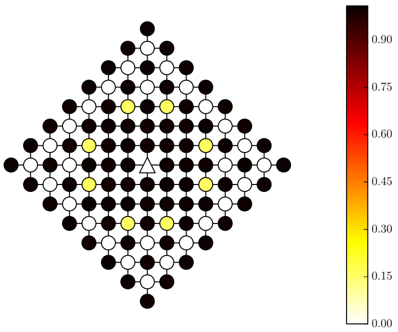



Figure 2: The active frequencies of nodes in the 7-layer tilted square lattice during the message passing process with $U = 40, V = 2$. The central triangular node is the terminal node.

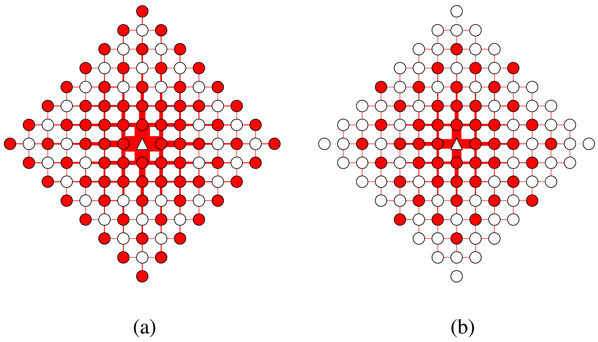



Figure 3: Optimal state with (a) $U = 40, V = 2$, (b) $U = 20, V = 0.8$. Results are obtained by MP guided decimation. Red nodes are active, and white nodes are idle. The width of an edge indicates the flow intensity.

### 4.1. MP Guided Decimation

Due to the presence of numerous short loops, the algorithm fails to converge in some regimes. We found that the non-convergence instances are usually associated with degenerate ground states or multiple local minima, from which conflicting messages are sent along different routes and confuse certain states of nodes.

For simplicity, we consider a 7-layer lattice with $U = 40, V = 2$. We show in Fig. 2 the relative frequency of occurrence of active state of each node $i$ during the message passing process. In Fig. 2, most of the nodes have a relative frequency close to either 1 or 0, i.e., they are certain about their states, except that there are eight nodes with intermediate relative frequencies on the 4th layer. In fact, only one of the eight nodes will be active in the ground state (see Fig. 3(a)), but the algorithm could not fix a single active node since the eight nodes are symmetric. Similar behaviours also appear in other non-convergent regimes.

Based on this observation, we propose to fix the state of oscillating nodes one by one heuristically according to their relative frequencies, reminiscent of belief/survey propagation guided decimation [9]. Our MP guided decimation

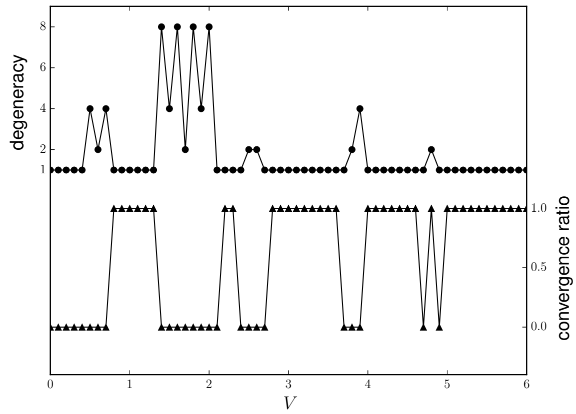



Figure 4: Upper panel : the ground state degeneracy of a 7-layer lattice as a function of $V$ for $U = 40$. Lower panel: the convergence ratio of the MP algorithm without decimation from 100 trials.

strategy is summarized as follows,

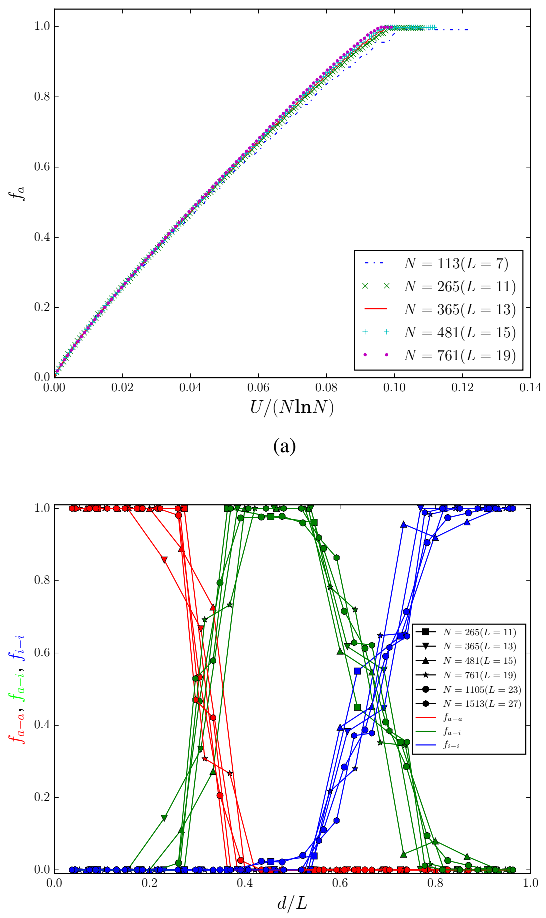
1. Initialize the messages and states of nodes;
2. Run message passing iterations for $t$ times;
3. If the MP iterations in Step 2 converge, exit and output $y_{ij}^*, n_i^*$; otherwise, go to Step 4;
4. Find the set of oscillating nodes, pick the most polarized one and fix its state;
5. Repeat Step 2 to 4 until the system converges.

Remarkably, after such decimation process, the MP algorithm always converges despite the numerous short loops. With the help of this strategy, we are able to identify the optimal states and reveal the ground state symmetry. We then verify in Fig. 4 that the non-convergence of the original MP algorithm (Algorithm 1) is associated with degenerate ground states, indicating a direct connection between the algorithmic behaviours and the energy landscape. Local minima other than degenerate ground states could also hamper algorithm convergence similarly. We remark that our algorithm also works well in random graphs.

### 4.2. Optimal States

By applying the MP algorithm with decimation, we further examine the optimal facility location in square lattices. In the asymptotic limit of large $U$ but small $V$, the system favours an all active state, similar to the ferromagnetic state in spin system; while in the limit of large $U$ and $V$ with $V/U \gg 1$, the system exhibits alternative active-idle pattern, which corresponds to the anti-ferromagnetic state [7]. In the intermediate regime, the system exhibits interesting ground state patterns.

In the two instances shown in Fig. 3, an active magnetized domain and an outer anti-ferromagnetic band coexist in the ground states. Physically speaking, the systems set up facilities extensively near the warehouse but only partly far away from the warehouse in order to save supply cost while still maintain comprehensive coverage.




(a)

(b)

Figure 5: (a) The fraction of active node $f_a$ as a function of $\tilde{U} = U/(N \ln N)$ for $\tilde{V} = 0$. (b) The fractions $f_{a-a}, f_{a-i}, f_{i-i}$ as a function of normalized Manhattan distance $d/L$ from the terminal node with $\tilde{U} = 0.2322, \tilde{V} = 0.0004644$.

A continuous approximation shows that the transportation cost scales as $N^2 \ln N$ in leading order in two dimension, which suggests a universal behavior as a function of the rescaled parameters $\tilde{U} = U/(N \ln N), \tilde{V} = V/(N \ln N)$ [7]. The presence of $\ln N$ in the rescaling factor is ubiquitous in two dimensional systems. In Fig. 5a, we show the fraction of active node $f_a$ versus $\tilde{U}$ for $\tilde{V} = 0$. After the rescaling, $f_a$ for different $N$ tends to collapse into a universal function. In Fig. 5b, we plot $f_{a-a}, f_{a-i}, f_{i-i}$, the fraction of edges connecting two active nodes, one active and one idle node, two idle nodes respectively, as a function of their normalized Manhattan distance from the terminal node with fixed $\tilde{U} = 0.2322, \tilde{V} = 0.0004644$. In this case, the system admits a mixed phase with the coexistence of an active core, an anti-ferromagnetic band and an outer idle area which looks like Fig. 3b. The data collapse in Fig. 5b indicates the optimal states of different system sizes behave similarly at the same $\tilde{U}$ and $\tilde{V}$.

## 5. Conclusion

We have derived a local message passing algorithm to identify the optimal locations of facilities in networks. A heuristic decimation strategy turns out to be effective to fix the convergence problem and allows us to apply the MP algorithm in loopy networks. In square lattices, it is observed that an active core and an anti-ferromagnetic domain can coexist to balance coverage and the transportation cost. We further verify that two-dimensional systems of different sizes behave similarly after applying the rescaling factors of $N \ln N$.


## Acknowledgments

This work was supported by the Research Grants Council of Hong Kong (Grants No. 604512 and No. 605813).


## References


[1] S. Rangwala, R. Gummadi, R. Govindan, and K. Psounis, "Interference-aware Fair Rate Control in Wireless Sensor Networks," *Comput. Commun. Rev.*, vol.36, pp.63-74, 2006.

[2] N. Megiddo, E. Zemel, and S. L. Hakimi, "The Maximum Coverage Location Problem," *SIAM J. Algebraic Discrete Methods*, vol.4, pp.253-261, 1983.

[3] D. Bertsekas, "Linear Network Optimization," *MIT Press*, 1991.

[4] M. Mézard, G. Parisi, M. Virasoro, "Spin Glass Theory and Beyond," *World Scientific*, 1987.

[5] K. Y. M. Wong and D. Saad, "Inference and optimization of real edges on sparse graphs: A statistical physics perspective," *Phys. Rev. E*, vol.76, pp.011115, 2007.

[6] C. H. Yeung and K. Y. M. Wong, "Optimal location of sources in transportation networks," *J. Stat. Mech. Theor. Exp.*, vol.2010, pp.P04017, 2010.

[7] C. H. Yeung, K. Y. M. Wong and B. Li, "Coverage versus supply cost in facility location: Physics of frustrated spin systems," *Phys. Rev. E*, vol.89, pp.062805, 2014.

[8] S. Bohn, M. O. Magnasco, "Structure, Scaling, and Phase Transition in the Optimal Transport Network," *Phys. Rev. Lett.*, vol.98, pp.088702, 2007.

[9] M. Mézard and R. Zecchina, "Random K-satisfiability problem: From an analytic solution to an efficient algorithm," *Phys. Rev. E*, vol.66, pp.056126, 2002.