2006 PerCom FAWN 2006

Unbalanced Points and Vertices Problem

Zvi Lotker
Centrum voor Wiskunde en Informatica, Amsterdam
Netherlands
lotker@cwi.nl

Alfredo Navarra
Computer Science Department
University of L'Aquila
Italy
navarra@di.univaq.it

Outline

- Points and Vertices, balanced version
  - Definition, an example
  - Previous Results
- Unbalanced Version
  - Definition
  - Motivation and Possible scenarios
- Results of the paper
  - d-dimensional Grid, Trees, d-Cubes, Paths and General graphs
- Concluding remarks

Original Points and Vertices Problem

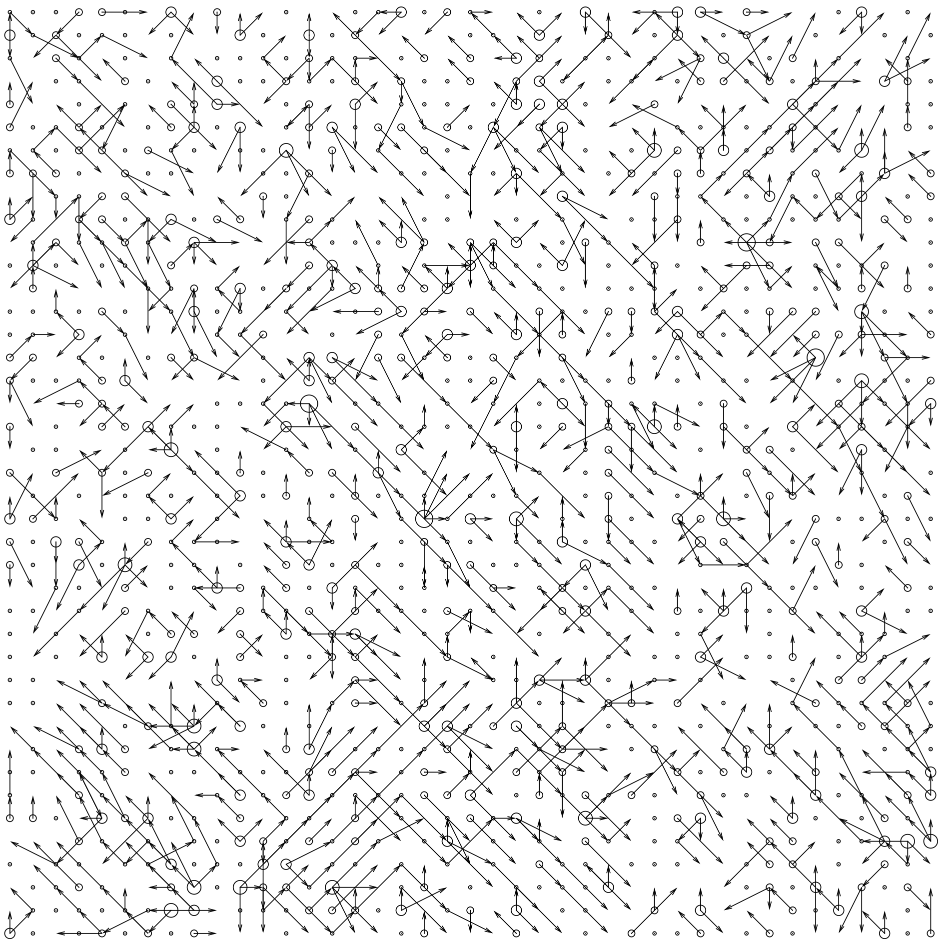
- Given a graph $G=(V,E)$,
- Throw $|V|$ points onto the vertices of G by means of the random uniform distribution
- The problem consists in remapping the points on G in such a way that each vertex contains exactly 1 point, while minimizing the maximum distance $\rho(G)$ that any point has to move on the edges of G.
- The aim is to bound the random variable $\rho(G)$

Remark 1

- There are many similarities with the well-known **Balls into Bins** problem (see [2,3,13])
- The main difference resides, in fact, in the added structural properties of the Bins (Vertices) since they are connected by the graph' edges so that the Balls (Points) cannot be moved as desired but following the graph paths
- Moreover, in the Balls and Bins problem, usually the aim is to study the distribution of the most loaded bin. In the Point and Vertices the aim is to study the accumulation on several vertices not just one
- Another interesting way of viewing the problem is in the opposite way of **discrepancy** since it captures in a natural way the "distance" between *randomness* of the thrown points and the *order* of the graph' vertices

Example on a Grid

- Each Grid vertex is represented by a circle whose radius is proportional to the thrown points inside it.
- The arrows represent the movements performed over the grid
- The final setting is a 1:1 matching between points and vertices



Previous Results

- In [10,14,15] by P.Shor et al. 1986 - 1991,
  - the relation between two basic structures like **Uniform Random** points and ***d*-dimensional Grid** vertices was studied. The expected minimax grid matching evaluated distance is
    - $\rho(G) = \Theta(\sqrt{\log n})$ for $d = 2$,
    - $\rho(G) = \Theta((\log n)^{3/4})$ for $d > 2$
- In [9] by R. Klasing et al. 2005,
  - *#P*-hardness for the general case
  - A Fully Polynomial Randomized Approximation Scheme when the graph admits a polynomial-size family of witness cuts
  - $\rho(G) = O(\sqrt{n})$ w.h.p. for any connected graph *G*
  - A greedy algorithm ✌ that remaps the points on any tree *T* with remapping distance $\rho_{✌}(T) \leq 2\rho(T)$

Original Points and Vertices Problem

Unbalanced Points and Vertices Problem

- Given a graph $G=(V,E)$, and a number $0<\varepsilon<1$
- Throw $(1-\varepsilon)|V|$ points onto the vertices of G by means of the random uniform distribution
- The problem consists in remapping the points on G in such a way that each vertex contains at most 1 point, while minimizing the maximal distance $\rho(G)$ that any point has to move on the edges of G.
- The aim is to bound the random variable $\rho(G)$

Motivations and Possible Scenarios

- Coverage issues
  - Robotics, Equidistant formation
  - Wireless Networks
    - Given an area of interest that can be partitioned in a set S of sub-areas such that $|S| = n$, and a set D of devices such that $|D| = (1-\varepsilon)n$, the goal is to minimize the spent transmission range and/or energy needed for movements of the devices in order to monitoring as much sub-areas as possible
- Vertex Capacity
    - Given n sensors distributed over a graph of n vertices it is possible to consider, as in reality, base stations (graph vertices) that can handle many sensors, i.e., capacity for each grid vertex hence obtaining an unbalanced version of the Points and Vertices
- Hierarchy
    - Given a variety of sensors with different capabilities like GPS, AOA, Thermometer and so on, the idea is to spread each type of sensor as much as possible hence solving several time the Unbalanced Points and Vertices problem

d-dimensional Grids

Theorem: $\rho(G)=\Theta(\sqrt[d]{\log n})$ *w.h.p.*

Sketch of the proof [upper bound]:

• Partition the grid into subgrids (boxes) containing, on average, $c_\varepsilon \log n$ points, i.e., $c_\varepsilon \frac{\log n}{1-\varepsilon}$ vertices, for some constant c_ε

• Let X_i be the number of random points belonging to the *i*-th box (the points can be rearranged in such a box whenever $X_i \leq c_\varepsilon \frac{\log n}{1-\varepsilon}$)

•From Chernoff [7] we have: $\Pr[X_i \geq \left(1+\frac{\varepsilon}{1-\varepsilon}\right)c_\varepsilon \log n] \leq e^{-\left(\left(\frac{\varepsilon}{1-\varepsilon}\right)^2 c_\varepsilon \log n\right)/3}$

•Hence for $c_\varepsilon \geq 3\left(\frac{1-\varepsilon}{\varepsilon}\right)^2$, *Pr[$X_i$ cannot be rearranged locally] ≤ 1/n*

•The claim holds by observing that there are *Θ(n / log n)* boxes and the diameter of each one is $d\left(\frac{c_\varepsilon}{1-\varepsilon}\right)^{\frac{1}{d}} \sqrt[d]{\log n}$

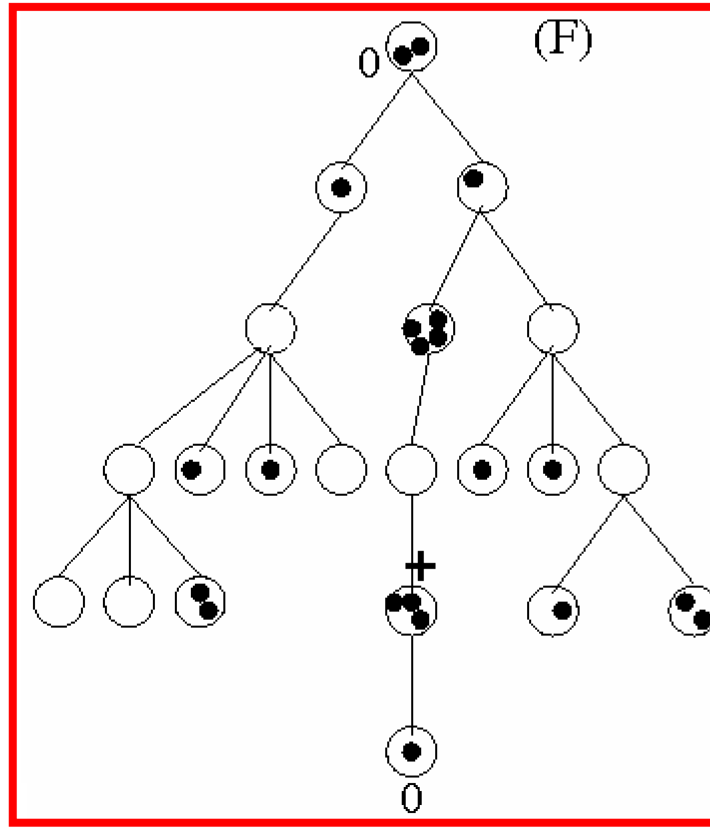
•An easy but not tight **lower bound** can be derived from [5] in the case of *(1+ε)n* bins and *n* balls

Trees

To each leaf we associate a label
-1 if there are no points inside.
0 if there is one point,
p-1 times +1 if there are p points.

For each labelled subtree but its root, let **s** be the difference among positive and negative labels in its sons.
If **s>0** we label the root with the smallest **s-1** positive numbers contained in the previous labels increased by 1 and a **+1** for each point contained in it.
If **s≤0**, let **s'** be the number of points contained in the root.

- If **s'>|s|** then we label the root with **s'+s-1 times +1** (hence with 0 if s'+s-1=0);
- if **s'<|s|** then with the biggest **|s|-s'-1** negative numbers contained in the previous labels decreased by 1 and a -1;
- if **s'=|s|** just with a -1



Among all the labels we look for the absolute biggest one (M=3 in the example). It holds **ρ(T)≥M**, and by means of matching arguments we achieve a **2**-approximation algorithm

Other Results

- d-dimensional Grid *G*:
  - $\rho(G)=\Theta(\sqrt[d]{\log n})$ *w.h.p.*
- Tree *T:*
  - *2-approxiamtion algorithm*
- d-dimensional Hypercubes *H*:
  - $\rho(H)=\Theta(1)$
- Paths *P:*
  - $\rho(P)=\Theta(\log n)$
- General Graph *G*:
  - $\rho(G)=O(\log n)$

Remark 2

- The grid remains the structure for which more deep techniques were required in order to obtain the bound
- Bounds for other topologies were more simple to derive, some of them are based on the grid method
- For trees, similar arguments to the 2-approximation algorithm proposed in [9] hold
- It is worth noting that for general graphs the value of *ρ(G)* is now *O(log n)* instead of $O(\sqrt{n})$ of the balanced version; This reveals a more local behavior of the random variable *ρ(G)*. This strengthens the fact that the unbalanced version is much more suitable for distributed environments like in the case of Sensor Networks

Conclusion

- We have proposed a variant of the *Points and Vertices Problem* by unbalancing the ratio between points and vertices
- We have derived several bounds for the remapping function among the points and vertices according to different topologies of the underlying graph
- The new remapping distance turned out to be much more local than the original one hence much more suitable for distributed environments
- As future work, an interesting direction could be to investigate the case in which the final setting of the points is not just one per vertex but it also preserves connectivity as much as possible

References

[2] R. Cole, A. M. Frieze, B. M. Maggs, M. Mitzenmacher, A. W. Richa, R. K. Sitaraman, and E. Upfal. **On balls and bins with deletions**. In *Proceedings of the 2nd International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 145–158. Springer-Verlag, 1998.

[3] E. Drinea, A. Frieze, and M. Mitzenmacher. **Balls and bins models with feedback**. In *Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 308–315. Society for Industrial and Applied Mathematics, 2002.

[9] R. Klasing, Z. Lotker, A. Navarra, and S. Perennes. **From Balls and Bins to Points and Vertices**. In *Proceedings of the the 16th Annual International Symposium on Algorithms and Computation (ISAAC)*, volume 3827 of *Lecture Notes in Computer Science*, pages 757–766. Springer-Verlag, 2005.

[10] F. T. Leighton and P. W. Shor. **Tight bounds for minimax grid matching with applications to the average case analysis of algorithms**. *Combinatorica*, 9(2):161–187, 1989.

[13] M. Raab and A. Steger. **"Balls into bins" - a simple and tight analysis**. In *Proceedings of the 2nd International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 159–170. Springer-Verlag, 1998.

[14] P. W. Shor. **The average-case analysis of some on-line algorithms for bin packing**. *Combinatorica*, 6(2):179–200, 1986.

[15] P. W. Shor and J. E. Yukich. **Minimax Grid Matching and Empirical Measures**. *The Annals of Probability*, 19(3):1338–1348, 1991.