Balancing Graph Voronoi Diagrams

Shinichi Honiden*[†], Michael E. Houle[†], and Christian Sommer*[†]

*The University of Tokyo, Tokyo, Japan   [†]National Institute of Informatics, Tokyo, Japan

Abstract—**Many facility location problems are concerned with minimizing operation and transportation costs by partitioning territory into regions of similar size, each of which is served by a facility. For many optimization problems, the overall cost can be reduced by means of a partitioning into balanced subsets, especially in those cases where the cost associated with a subset is superlinear in its size. In this paper, we consider the problem of generating a Voronoi partition of a discrete graph so as to achieve balance conditions on the region sizes. Through experimentation, we first establish that the region sizes of randomly-generated graph Voronoi diagrams vary greatly in practice. We then show how to achieve a balanced partition of a graph via Voronoi site resampling. For bounded-degree graphs, where each of the n nodes has degree at most d, and for an initial randomly-chosen set of s Voronoi nodes, we prove that, by extending the set of Voronoi nodes using an algorithm by Thorup and Zwick, each Voronoi region has size at most $4dn/s + 1$ nodes, and that the expected size of the extended set of Voronoi nodes is at most $2s \log n$.**

Keywords-**graph Voronoi diagram; balancing; facility location; territorial design**

I. INTRODUCTION

Imagine a sales manager who aims at partitioning a country into different regions, the clients within each of which are to be served by a different sales team. He may want to both minimize the total travel time for all teams and to balance the workload among the different regions. The underlying problem of partitioning territory in order to efficiently allocate resources appears in many scenarios, such as sales force deployment [1], [2], ATM placement [3], political districting [4], [5], school districting [6], and salt spreading coordination [7]. The corresponding optimization problems are referred to as *facility location* [8], [9], [10], [11], [12] or *territorial design* [13]. In optimization, the objective of partitioning is to minimize cost. Whenever a cost function depends superlinearly on the size of the regions, balancing the region sizes can contribute to the minimization of the total cost.

The divide-and-conquer paradigm in algorithm design constitutes another application of balanced partitioning. In the context of graphs, divide-and-conquer strategies decompose the graph into components, solve the problem for each connected component, and combine the local solutions into a global solution. If the algorithm for each component requires superlinear time, a highly imbalanced partition would lead to increased computation cost. For some algorithms, the overall cost could be dominated by the time required to process the largest region. In computer science, partitioning

is also relevant for problems involving graph structures other than planar maps; applications include image processing [14], paging in operating systems [15], and routing in networks [16].

Finding a perfect partition of a graph is apparently hard: the problem of deciding whether a graph allows for an exact partition has been proven to be **NP**-complete [17, Thm. 1]. If the vertices are assigned weights, the problem remains **NP**-hard, even when restricted to classes of graphs as simple as series-parallel graphs [18], which often admit polynomial-time solutions for otherwise hard problems. For the relaxed problem where a graph is to be partitioned into balanced regions with sizes ranging between two threshold values, polynomial-time algorithms are known for (weighted) paths, trees [19], and graphs with bounded tree-width [18], [20]. For the case of general graphs, one may further relax the region size constraints so as to compute a partition for which the size of every region is bounded only from above, but the number of regions is not too large. The graph Voronoi diagram [21], [22] would constitute a solution to this relaxed partitioning problem if the region sizes could be made to respect the balancing condition.

The analysis of properties of Voronoi diagrams has received much attention for spatial settings. Gilbert [23] and Meijering [24], motivated by the problem of crystal formation, pioneered the theory of random partitions of space. The Poisson Voronoi tessellation in $\mathbb{R}^d$ is one whose sites have been generated according to a stationary Poisson point process [25], which has no dependence on process history and whose intensity does not change over time. Miles [26], Ambartzumian [27], and Møller [28] provided some of the first fundamental results for random Voronoi tessellations. Other results concerning the area or volume of Voronoi cells include the following. Mecke and Muche [29], Muche [30], [31], and Heinrich and Muche [32] analyzed several properties of random tessellations; among others, they derived representation formulae for the second moment of the number of vertices of the 'typical' cell. Hilhorst [33], [34], [35] analyzed the cell area for the planar case. Calka [36] analyzed the two-dimensional case, giving precise formulae for the area and perimeter of cells. Rathie [37] gave the exact distribution of the volume of cells for $d \leq 3$. Tanemura [38] evaluated various properties of cells under a simulation of the Poisson point process for two and three dimensions. Brakke derived a general scheme for integrals of probability density functions. He analyzed statistics for the plane [39],

for the three dimensional space [40], and for higher dimensions [41]. He also simulated random Voronoi tessellations of the plane [42], confirming his analytical findings. Research results are also known for random tessellations other than the Voronoi diagram. Cowan [43], [44], using ergodic theory, analyzed mosaic processes and Poisson point processes. Lautensack and Zuyev [45] evaluated Laguerre tessellations, which are weighted generalizations of Voronoi tessellations. To the best of our knowledge, the properties of random Voronoi diagrams for discrete graphs have not been investigated previously.

A. Our contribution

We first establish, experimentally, the extent to which the region sizes of random graph Voronoi diagrams, and the node degrees of their Delaunay duals, vary in practice for several different classes of graphs. The graphs under consideration stem from road networks of Europe and the USA, European public transportation networks, social networks, protein interactions, citations among scientific publications, and computer networks. The results reveal that, indeed, region sizes do vary greatly, and that the balancing problem for Voronoi diagrams is well-motivated.

In the second part of this paper, we address the problem of finding balanced Voronoi partitions of graphs. Our approach is to resample using an algorithm due to Thorup and Zwick that was originally proposed for bounding routing table sizes [16]. We prove that, for a graph $G = (V, E)$ with $|V| = n$ and maximum vertex degree d, Thorup and Zwick's original algorithm can, in addition to solving the problem they intended to solve, also extend a random sample $S \subseteq V$ of size $|S| = s$ such that the expected size of the extended sample is at most $2s \log n$ and every Voronoi region contains at most $\frac{4dn}{s} + 1$ nodes.

II. PRELIMINARIES

A *graph* $G = (V, E)$ consists of a set of *vertices* V and *edges* $E \subseteq \binom{V}{2}$. A *weighted* graph $G = (V, E, \omega)$ consists of a graph (V, E) together with a *weight function* $\omega : E \to \mathbb{R}^+$. We assume positive edge weights. The *neighbors* of v are the members of $N(v) = \{u : (v, u) \in E\}$. The *degree* of a node v, $\deg(v) = |N(v)|$, is the number of its neighbors. In an edge-weighted graph $G = (V, E, \omega)$, a *path* from $s = u_0 \in V$ to $t = u_h \in V$ is a sequence of nodes $(u_0, u_1, \ldots, u_h)$ for which $(u_i, u_{i+1}) \in E$ for all $i \in \{0, 1, \ldots h-1\}$. The *length* of a path P is the sum of its edge weights $\ell(P) := \sum_{i=0}^{h-1} \omega(u_i, u_{i+1})$. A subpath P' of a path P is a subsequence of its nodes $P' = (u_i, u_{i+1}, \ldots u_j)$, $0 \leq i < j \leq h$. A *simple path* is a path without repeated vertices. Let $\mathcal{P}_G(u, v)$ denote the set of paths from u to v in a graph G. The *distance* $d(u, v)$ between two nodes u, v is the length of a shortest path from u to v; that is, $d(u, v) = \min_{P \in \mathcal{P}_G(u,v)} \ell(P)$.

The classical Voronoi diagram is a distance-based decomposition of a metric space relative to a discrete set, the Voronoi sites. Given a set of points (the Voronoi sites), the Voronoi decomposition leads to regions (the Voronoi regions) consisting of all points that are closest to a specific site. Mehlhorn [21] and Erwig [22] proposed an analogous decomposition, the *Graph Voronoi Diagram*, for undirected and directed graphs respectively.

Definition 1 (Graph Voronoi Diagram [21], [22]). *In a graph $G = (V, E, \omega)$, the* Voronoi diagram *for a set of nodes $S = \{v_1, \ldots, v_s\} \subseteq V$ is a disjoint partition* $\mathsf{Vor}_{(G,S)} := \{V_1, \ldots, V_s\}$ *of V such that for each node $u \in V_i$, $d(u, v_i) \leq d(u, v_j)$ for all $j \in \{1, \ldots, s\}$.*

The V_i are called *Voronoi regions*. The graph Voronoi diagram is not necessarily unique, as a node u may have the same distance to more than one Voronoi node.

Analogously to the Delaunay triangulation dual for classical Voronoi diagrams of point sets, one can define the Voronoi dual for graphs.

Definition 2. *Let $G = (V, E, \omega)$ be a weighted graph and* $\mathsf{Vor}_{G,S}$ *its Voronoi diagram. The* Voronoi dual *is the graph $G^* = (S, E^*, \omega^*)$ with edge set $E^* := \{(v_i, v_j) : v_i, v_j \in S$ and $\exists u \in V_i \land \exists w \in V_j : (u, w) \in E\}$, and edge weights*
$$\omega^*(v_i, v_j) := \min_{\substack{u \in V_i, w \in V_j \\ (u,w) \in E}} \{d(v_i, u) + \omega(u, w) + d(w, v_j)\}.$$

Figure 1 illustrates two graph Voronoi diagrams for the same (planar) graph but with different edge weights. Although the classical Voronoi dual of a non-degenerate set of points in the plane is always a triangulation, the graph Voronoi dual is not necessarily a triangulation, even for planar graphs. For example, a graph Voronoi dual may have nodes whose removal would disconnect the graph.

Erwig [22, Thm. 2] shows that the graph Voronoi diagram can be constructed via a single Dijkstra search in time $O(m + n \cdot \log n)$. Honiden et al. [46] provide an example of how the construction of a Voronoi diagram [22, sec. 3.1] may be adapted to compute a graph Voronoi dual — that is, to also compute E_{G^*} and ω_{G^*}.

III. VARIATION IN VORONOI REGION SIZES

In order to show the extent of the variance of region sizes in randomly-generated graph Voronoi diagrams, we applied the following generation algorithm to the data sets described below. Given a graph $G = (V, E)$ with $|V| = n$, the algorithm generates a set of Voronoi sites $S \subseteq V$ by selecting each node of V independently at random with probability $p = \lceil \sqrt{n} \rceil / n$, and computes the graph Voronoi dual of S as per Definition 2. We report the mean values and standard deviations of the node degrees in the Voronoi dual, as well as the mean values and standard deviations of the number of Voronoi region nodes, produced over 100 repetitions of the algorithm.

A. Data sets

For the experimentation, we make use of several graphs derived from five transportation network data sets. Two of the graphs represent road networks: one from Western Europe, made available for scientific use by the company PTV AG, and the other from the USA, a data set figuring in the 9th DIMACS implementation challenge [47]. For some road network graphs, there are two different edge weightings, one representing geographical distances (denoted by **d** in the names of the graphs listed in the result tables) and the other representing driving time (denoted by **t**). The other three graphs represent European public transportation networks used previously for experiments conducted in [48], [49]: one of the three represent long-distance railway connections, and the other two represent the bus networks of two companies, the Rhein-Main-Verkehrsverbund RMV, and the Verkehrsverbund Berlin Brandenburg VBB. For the experimentation, all graphs derived from transportation networks were converted into their undirected forms.

We also conduct experiments for a number of instances of power-law graphs, also known as scale-free networks. Power-law graphs generally obey the power-law distribution, which assumes that every node has degree x with probability proportional to $x^{-\tau}$ ($Pr[\deg(v) > x] = x^{-\tau+1}$). The power-law graph model seems to apply to many real-world structures [50], [51] such as social networks, the web graph, protein interaction networks, semantic networks, and others. For our experimentation, power-law graphs were generated from the following databases: a co-authorship graph derived from the DBLP computer science bibliography [52], in which two authors are connected by an edge if they share at least one joint publication (dblp20080824); a citation graph used as a dataset in the KDD Cup 2003 competition [53], taken from the high-energy physics research literature (hep-th-cite); a graph derived from the CAIDA [54] database of the router-level topology of a portion of the Internet (itdk0304); four graphs based on two random topologies generated using BRITE [55], two for each of the Barabási [56] and Waxman [57] models (ASBarabasi, ASWaxman, RTBarabasi, RTWaxman); a graph of experimentally-determined protein-protein interactions taken from the Database of Interacting Proteins [58] (dip20090126); a graph of physical and genetic interactions from the freely-accessible database BioGRID [59]. For each data set, the power-law coefficient τ was estimated using the maximum-likelihood method [60].

B. Results

The results of the experimentation are listed in Tables I, II, and III. The columns from left to right list the name of the graphs considered, their numbers of nodes and edges, and the means and standard deviations of their node degrees. The next columns show the means and standard deviations of two quantities taken over 100 random graph Voronoi diagram generations: the Voronoi dual node degrees, and the Voronoi region sizes. For the latter, the standard deviation is normalized by dividing by the mean number of nodes of the Voronoi regions, and the mean is consequently shown as 1.

The results for the US and European road networks are presented in Table I. The standard deviation of Voronoi region sizes is fairly consistent across all graphs studied, at approximately 0.7 of the mean value. Note that for those road networks with edges weighted by distance, the standard deviation is consistently smaller than for those graphs with edges weighted by driving times. This is perhaps due to the fact that the distance-weighted graph is embeddable in the Euclidean plane, whereas large time-weighted graphs are (in general) not embeddable.

For the power-law graphs (Table II) and the railway networks (Table III), the variation among Voronoi region sizes is substantially larger, although the variation does not appear to be correlated with the power-law coefficient τ.

IV. Balancing the Region Sizes by Resampling

In this section, we analyze the resampling algorithm due to Thorup and Zwick [16, Algo. 2]. The original application of their algorithm is in the limitation of the routing table sizes required for a compact routing scheme, also introduced in [16].

A. Thorup and Zwick resampling

Recall that for a sample $S \subseteq V$, the Voronoi diagram is a partition of the node set into Voronoi regions V_i such that for each node $u \in V_i$, $d(u, v_i) \leq d(u, v_j)$ for all $j \in \{1, \ldots, k\}$. Ties in distances may be broken arbitrarily (but consistently). Using the notation of Thorup and Zwick, let the method $\mathsf{sample}(W, s)$ return a random subset of W by selecting each node independently at random with probability $\min\{1, s/|W|\}$. If $s \geq |W|$, the entire set W is returned. The rebalancing algorithm is stated below. The description requires the determination of the *cluster* $C_S(w)$ of a node $w \in V$, defined by Thorup and Zwick to be the set of nodes that would lie in the Voronoi region of w if it were to be inserted into the set of Voronoi nodes S. Their interest in bounding the size of $C_S(w)$ is motivated by the need to store references to these nodes in the routing table of w in their compact routing strategy.

Algorithm 1 (center(G, s) [16, Algo. 2])**.**

- $S := \emptyset$
  $W := V$
- *while* $W \neq \emptyset$
  - $S := S \cup \mathsf{sample}(W, s)$
  - *for all* $w \in V$: $C_S(w) := \{v \in V : d(w, v) < d(S, v)\}$
  - $W := \{w \in V : |C_S(w)| > \frac{4n}{s}\}$
- `return` S

Graph	$|V|$	$|E|$	Degree	Vor. Dual Degree	Vor. Region Size
cze	23094	53412	2.631±0.953	4.432±1.712	1±0.759
lux	30047	70189	2.439±0.896	4.756±1.729	1±0.739
irl	32868	71655	2.349±0.966	4.688±1.804	1±0.770
prt	159945	372129	2.563±0.889	5.040±1.794	1±0.715
fin	458231	1019975	2.209±1.009	4.600±1.642	1±0.700
bel	458403	1092727	2.538±0.906	5.487±1.658	1±0.649
dnk	469110	1074929	2.324±0.944	4.791±1.721	1±0.727
che	585514	1343980	2.355±0.917	4.689±1.738	1±0.746
esp	695399	1546253	2.685±0.868	4.965±1.786	1±0.830
aut	722512	1667569	2.350±0.932	4.780±1.709	1±0.712
nld	849537	2033788	2.520±0.899	5.077±1.725	1±0.740
nor	898196	1946881	2.169±1.008	3.856±1.498	1±0.835
swe	1546705	3579738	2.253±0.992	4.886±1.664	1±0.708
ita	2077709	4824125	2.493±0.955	5.122±1.806	1±0.761
gbr	2142853	4896816	2.406±0.935	5.269±1.738	1±0.719
fra	2793047	6558521	2.526±0.886	5.508±1.738	1±0.722
deu	4047577	9572939	2.425±0.929	5.493±1.580	1±0.640
d.eur	18010173	42560279	2.467±0.942	5.347±1.681	1±0.676
t.eur	18010173	42560279	2.467±0.942	5.440±2.261	1±0.856
d.NY	264346	733846	2.762±0.979	4.886±1.577	1±0.655
t.NY	264346	733846	2.762±0.979	4.965±1.881	1±0.708
d.BAY	321270	800172	2.474±0.987	4.568±1.562	1±0.748
t.BAY	321270	800172	2.474±0.987	4.628±1.837	1±0.796
d.COL	435666	1057066	2.393±0.915	4.573±1.643	1±0.768
t.COL	435666	1057066	2.393±0.915	4.614±1.894	1±0.861
d.FLA	1070376	2712798	2.511±0.955	4.690±1.561	1±0.723
t.FLA	1070376	2712798	2.511±0.955	4.778±1.878	1±0.808
d.NW	1207945	2840208	2.335±0.938	4.742±1.635	1±0.735
t.NW	1207945	2840208	2.335±0.938	4.789±1.882	1±0.825
d.NE	1524453	3897636	2.537±0.951	5.155±1.601	1±0.657
t.NE	1524453	3897636	2.537±0.951	5.217±1.916	1±0.732
d.CAL	1890815	4657742	2.449±0.941	4.763±1.563	1±0.728
t.CAL	1890815	4657742	2.449±0.941	4.817±1.847	1±0.814
d.LKS	2758119	6885658	2.464±0.927	5.286±1.588	1±0.650
t.LKS	2758119	6885658	2.464±0.927	5.329±1.863	1±0.742
d.E	3598623	8778114	2.420±0.942	5.132±1.608	1±0.662
t.E	3598623	8778114	2.420±0.942	5.199±1.908	1±0.753
d.W	6262104	15248146	2.414±0.921	4.936±1.572	1±0.712
t.W	6262104	15248146	2.414±0.921	4.985±1.889	1±0.818
d.CTR	14081816	34292496	2.405±0.929	5.427±1.555	1±0.646
t.CTR	14081816	34292496	2.405±0.929	5.463±1.892	1±0.740
1.USA	23947347	58333344	2.410±0.929	5.306±1.760	1±0.640
d.USA	23947347	58333344	2.410±0.929	5.334±1.569	1±0.652
t.USA	23947347	58333344	2.410±0.929	5.381±1.889	1±0.747

Table I

VARIATION OF REGION SIZE AND DISTRIBUTION FOR 30 ROAD NETWORKS OF EUROPE AND THE USA.

Graph	$\hat{\tau}$ [60]	$|V|$	$|E|$	degree	Vor. Dual Degree	Vor. Region Size
BIOGRID	2.551	4039	43854	10.855±13.435	15.285± 10.253	1±1.275
ASBarabasi	2.893	10000	39994	3.999± 6.529	44.120± 23.774	1±1.831
ASWaxman	2.806	10000	40000	4.000± 2.486	60.273± 20.672	1±1.072
RTBarabasi	2.892	10000	39994	3.999± 6.159	45.215± 23.566	1±1.559
RTWaxman	2.806	10000	40000	4.000± 2.479	60.460± 20.275	1±1.047
dip20090126	3.079	19928	82406	4.135± 5.842	34.335± 17.862	1±1.456
hep-th-cite	2.385	27400	705084	25.732±45.558	69.836± 37.055	1±1.841
itdk0304	2.820	190914	1215220	6.365±14.181	35.737± 48.355	1±3.456
dblp20080824	2.686	511163	3742140	7.321±11.785	143.868±124.423	1±2.412

Table II

VARIATION OF REGION SIZE AND DISTRIBUTION FOR NINE POWER-LAW GRAPHS, TOGETHER WITH THE ESTIMATED POWER-LAW COEFFICIENT $\hat{\tau}$.

Theorem 1 (Thorup and Zwick [16, Thm. 3.1])**.** *The expected size of the set S returned by Algorithm* center(G, s) *is at most $2s \log n$. For every $w \in V$, the cluster size is bounded by $|C_S(w)| \leq \frac{4n}{s}$.*

Note that if w also happens to be a member of S, the condition $\forall v : d(w, v) \geq d(S, v)$ implies that its cluster $C_S(w)$ is empty.

The bound on cluster size stated in the theorem does not directly apply to the region sizes themselves: for a node $v_i \in S$ it may happen that its Voronoi region V_i has size $|V_i| > \frac{4n}{s}$ even though $|C_S(w)| \leq \frac{4n}{s}$ may hold for all $w \in V$. However, we prove in the following that the union of all sets $C_S(v')$ over all neighbors v' of v_i does in fact contain all nodes of V_i.

Theorem 2. *For a graph $G = (V, E)$ with $|V| = n$, positive edge weights, and maximum degree d, the expected size of the set S returned by Algorithm* center(G, s) *is at most $2s \log n$. For every $v_i \in S$ with Voronoi region V_i, we then have $|V_i| \leq \frac{4dn}{s} + 1$. Also, for every $w \in V$ it holds that $|C_S(w)| \leq \frac{4n}{s}$.*

Proof: Except for the bounds on the region sizes $|V_i|$, the remainder of the claims follow directly from Theorem 1. It suffices to prove that no region contains more than $\frac{4dn}{s} + 1$ nodes. Recall that for each node $u \in V_i \setminus \{v_i\}$, we have $d(u, v_i) \leq d(u, v_j)$ for all $j \in \{1, \ldots, s\}$. Every shortest path from v_i to $u \in V_i$ must include at least one of the neighbors of v_i (call it v'). As the edge weights are positive, the length of the subpath from v' to u is strictly less than the length of any shortest path from v' to u. Therefore, $u \in C_S(v')$, and thus every $u \in V_i \setminus \{v_i\}$ is contained in $C_S(v')$ for some choice of neighbor v' of v_i.

$$V_i \subseteq v_i \cup \bigcup_{v \in N(v_i)} C_S(v)$$

As $|C_S(w)| \leq \frac{4n}{s}$ for all $w \in V$, we obtain

$$|V_i| \leq 1 + |N(v_i)| \cdot \frac{4n}{s} \leq 1 + \frac{4dn}{s}$$

∎

In general, for graphs with unbounded degree, Algorithm 1 cannot generate a balanced partition where all region sizes are bounded by $O(\frac{n}{s})$. As a counter-example, consider an n-node star graph with distinct edge weights. If the center node is selected as a Voronoi site (which happens with probability $\frac{s}{n}$), its Voronoi region includes all nodes except S. Thus, its region size can only be reduced by adding nodes to S. Reducing its region size down to $O(\frac{n}{s})$ requires the addition of $n - s - O(\frac{n}{s})$ Voronoi nodes. This extended set of Voronoi nodes may be much larger than the desired bound of $2s \log n$. If the center node is not selected as a Voronoi site, the Voronoi node closest to the center will have the largest Voronoi region, which would include all nodes of $V \setminus S$.

Thorup and Zwick bound the run time of their algorithm as follows:

Theorem 3 (Thorup and Zwick [16, proof of Thm. 3.1])**.** *The expected run time of Algorithm* center(G, s) *is at most $O(\frac{mn}{s} \log n)$.*

V. Experiments

For the experiments, the resampling algorithm is given an edge-weighted input graph $G = (V, E, \omega)$ with $|V| = n$ and target sample size $s = \lceil \sqrt{n} \rceil$. As test data, we use the same graphs as in Section III. For each graph, we report the mean values and standard deviations of the Voronoi dual node degrees, and the means and standard deviations of the Voronoi region size. Here, the values for the region size are normalized by dividing by the expected size of Voronoi regions *without* resampling; that is, by $\frac{n}{s} = \frac{n}{\lceil \sqrt{n} \rceil} = \frac{1}{p}$. For example, a reported value of 0.5 ± 0.25 would indicate a mean region size of $\frac{n}{2s}$ (instead of $\frac{n}{s}$), with a standard deviation of $\frac{n}{4s}$.

The experimental results, shown in Tables IV, V, and VI, indicate an effective reduction in the variance of the Voronoi region sizes. Throughout our experimentation, we observed that only a few rounds of resampling were required for all region sizes to fall below the bound $\frac{4n}{s}$. For the road networks (Table IV), the number of Voronoi sites roughly doubled due to the resampling, and for the public transportation networks (Table VI), the number of Voronoi sites roughly tripled. For the power-law graphs, with unbounded degrees, the resampling expanded the number of Voronoi nodes by a factor of up to 10 to 20 times (note that $2 \log n \geq 20$ for all the graphs evaluated). As expected, the variance for these power-law graphs could be bounded only by introducing a substantial number of additional Voronoi nodes.

VI. Conclusion

Regions of random graph Voronoi diagrams vary greatly in size. For graphs with bounded degree, the resampling algorithm by Thorup and Zwick can be used to obtain a graph Voronoi diagram with bounded region sizes. A number of practical improvements to the algorithm seem possible. The original algorithm actually enforces stronger conditions than what is necessary for a balanced partition. The resampling algorithm repeatedly computes the size $|C_S(w)|$ for every node w. This is not necessary, since, for a balanced partition, the size of $C_S(w)$ does not have to be bounded for all w. Instead of simply reiterating the resampling process, it may be possible to restrict the resampling and the computation of the sets $C_S(w)$ to only those nodes w contained in large Voronoi regions. In addition, instead of resampling among nodes w with large sets $C_S(w)$, a heuristic algorithm could resample among nodes in large Voronoi regions without the explicit (and costly) computation of their clusters $C_S(w)$.

Graph	$|V|$	$|E|$	Degree	Vor. Dual Degree	Vor. Region Size
longRail	1215750	1823609	3.000±0.006	16.984± 9.711	1±1.251
RMV	2277812	3416597	3.000±0.010	15.379± 9.767	1±1.209
VBB	2599953	3899807	3.000±0.010	18.770±11.978	1±1.035

Table III

VARIATION OF REGION SIZE AND DISTRIBUTION FOR 3 EUROPEAN PUBLIC TRANSPORTATION GRAPHS.

Graph	$|V|$	$|E|$	Degree	Vor. Dual Degree	Vor. Region Size
cze	23094	53412	2.631±0.953	4.242±1.702	0.528±0.434
lux	30047	70189	2.439±0.896	4.581±1.696	0.568±0.440
irl	32868	71655	2.349±0.966	4.553±1.764	0.593±0.444
prt	159945	372129	2.563±0.889	4.970±1.779	0.578±0.420
fin	458231	1019975	2.209±1.009	4.387±1.646	0.524±0.427
bel	458403	1092727	2.538±0.906	5.416±1.719	0.578±0.407
dnk	469110	1074929	2.324±0.944	4.806±1.754	0.570±0.431
che	585514	1343980	2.355±0.917	4.624±1.763	0.573±0.431
esp	695399	1546253	2.685±0.868	4.921±1.767	0.545±0.426
aut	722512	1667569	2.350±0.932	4.695±1.737	0.565±0.433
nld	849537	2033788	2.520±0.899	5.031±1.772	0.553±0.429
nor	898196	1946881	2.169±1.008	3.703±1.468	0.538±0.446
swe	1546705	3579738	2.253±0.992	4.748±1.693	0.566±0.433
ita	2077709	4824125	2.493±0.955	5.050±1.819	0.572±0.442
gbr	2142853	4896816	2.406±0.935	5.219±1.743	0.592±0.428
fra	2793047	6558521	2.526±0.886	5.498±1.759	0.572±0.421
deu	4047577	9572939	2.425±0.929	5.426±1.623	0.626±0.413
d.eur	18010173	42560279	2.467±0.942	5.331±1.699	0.608±0.420
t.eur	18010173	42560279	2.467±0.942	5.432±2.390	0.516±0.442
d.NY	264346	733846	2.762±0.979	4.859±1.604	0.617±0.421
t.NY	264346	733846	2.762±0.979	4.924±1.866	0.583±0.416
d.BAY	321270	800172	2.474±0.987	4.538±1.597	0.559±0.431
t.BAY	321270	800172	2.474±0.987	4.585±1.873	0.533±0.423
d.COL	435666	1057066	2.393±0.915	4.472±1.650	0.559±0.431
t.COL	435666	1057066	2.393±0.915	4.510±1.854	0.527±0.429
d.FLA	1070376	2712798	2.511±0.955	4.626±1.598	0.558±0.423
t.FLA	1070376	2712798	2.511±0.955	4.712±1.902	0.523±0.421
d.NW	1207945	2840208	2.335±0.938	4.643±1.659	0.550±0.425
t.NW	1207945	2840208	2.335±0.938	4.679±1.854	0.514±0.418
d.NE	1524453	3897636	2.537±0.951	5.099±1.644	0.573±0.419
t.NE	1524453	3897636	2.537±0.951	5.164±1.949	0.541±0.419
d.CAL	1890815	4657742	2.449±0.941	4.705±1.610	0.540±0.428
t.CAL	1890815	4657742	2.449±0.941	4.746±1.853	0.507±0.421
d.LKS	2758119	6885658	2.464±0.927	5.259±1.637	0.591±0.426
t.LKS	2758119	6885658	2.464±0.927	5.298±1.920	0.547±0.423
d.E	3598623	8778114	2.420±0.942	5.083±1.656	0.578±0.424
t.E	3598623	8778114	2.420±0.942	5.141±1.953	0.531±0.421
d.W	6262104	15248146	2.414±0.921	4.857±1.623	0.557±0.428
t.W	6262104	15248146	2.414±0.921	4.885±1.889	0.510±0.427
d.CTR	14081816	34292496	2.405±0.929	5.228±1.730	0.487±0.434
t.CTR	14081816	34292496	2.405±0.929	5.225±2.036	0.446±0.428
1.USA	23947347	58333344	2.410±0.929	5.245±1.832	0.583±0.421
d.USA	23947347	58333344	2.410±0.929	5.263±1.629	0.580±0.425
t.USA	23947347	58333344	2.410±0.929	5.297±1.957	0.533±0.427

Table IV

AFTER RESAMPLING: VARIATION OF REGION SIZE AND DISTRIBUTION FOR 30 ROAD NETWORKS OF EUROPE AND THE USA.

REFERENCES

[1] A. Drexl and A. Drexl, "Fast approximation methods for sales force deployment," *Management Science*, vol. 45, pp. 1307–1323, 1999.

[2] S. W. Hess and S. A. Samuels, "Experiences with a sales districting model: Criteria and implementation," *Management Science*, vol. 18, pp. 998–1006, 1971.

[3] G. Cornuejols, M. L. Fisher, and G. L. Nemhauser, "Location of bank accounts to optimize float: an analytic study of exact and approximation algorithms," *Management Science*, vol. 23, no. 8, pp. 789–810, 1977.

[4] B. Bozkaya, E. Erkut, and G. Laporte, "A tabu search heuristic and adaptive memory procedure for political districting," *European Journal of Operational Research*, vol. 144, no. 1, pp. 12–26, 2003.

[5] M. Hojati, "Optimal political districting," *Computers & Operations Research*, vol. 23, no. 12, pp. 1147–1161, 1996.

[6] F. Caro, T. Shirabe, M. Guignard, and A. Weintraub, "School redistricting: Embedding GIS tools with integer programming," *Journal of the Operational Research Society*, vol. 55, pp. 836–849, 2004.

[7] L. Muyldermans, D. Cattrysse, D. V. Oudheusden, and T. Lotan, "Districting for salt spreading operations," *European Journal of Operational Research*, vol. 139, no. 3, pp. 521–532, 2002.

[8] K. Aardal, F. A. Chudak, and D. B. Shmoys, "A 3-approximation algorithm for the k-level uncapacitated facility location problem," *Information Processing Letters*, vol. 72, pp. 161–167, 1999.

[9] V. Arya, N. Garg, R. Khandekar, A. Meyerson, K. Munagala, and V. Pandit, "Local search heuristics for k-median and facility location problems," *SIAM Journal on Computing*, vol. 33, no. 3, pp. 544–562, 2004.

[10] N. Garg, R. Khandekar, and V. Pandit, "Improved approximation for universal facility location," in *Symposium on Discrete Algorithms*, 2005, pp. 959–960.

[11] M. I. Shamos, "Geometric complexity," in *Symposium on Theory of Computing*, 1975, pp. 224–233.

[12] Z. Svitkina, "Lower-bounded facility location," in *Symposium on Discrete Algorithms*, 2008, pp. 1154–1163.

[13] J. Kalcsics, S. Nickel, and M. Schröder, "Towards a unified territorial design approach - applications, algorithms and gis integration," *TOP: An Official Journal of the Spanish Society of Statistics and Operations Research*, vol. 13, no. 1, pp. 1–56, 2005.

[14] M. Lucertini, Y. Perl, and B. Simeone, "Most uniform path partitioning and its use in image processing," *Discrete Applied Mathematics*, vol. 42, no. 2, pp. 227–256, 1993.

[15] A. S. Tanenbaum, *Operating systems: design and implementation; 3rd ed.*

[16] M. Thorup and U. Zwick, "Compact routing schemes," in *Symposium on Parallelism in Algorithms and Architectures*, 2001, pp. 1–10.

[17] C. Dürr and N. K. Thang, "Nash equilibria in Voronoi games on graphs," in *European Symposium on Algorithms*, 2007, pp. 17–28.

[18] T. Ito, X. Zhou, and T. Nishizeki, "Partitioning a graph of bounded tree-width to connected subgraphs of almost uniform size," *Journal of Discrete Algorithms*, vol. 4, no. 1, pp. 142–154, 2006.

[19] T. Ito, T. Uno, X. Zhou, and T. Nishizeki, "Partitioning a weighted tree to subtrees of almost uniform size," in *International Symposium on Algorithms and Computation*, 2008, pp. 196–207.

[20] T. Ito, K. Goto, X. Zhou, and T. Nishizeki, "Partitioning a multi-weighted graph to connected subgraphs of almost uniform size," *IEICE Transactions*, vol. 90-D, no. 2, pp. 449–456, 2007.

[21] K. Mehlhorn, "A faster approximation algorithm for the Steiner problem in graphs," *Information Processing Letters*, vol. 27, no. 3, pp. 125–128, 1988.

[22] M. Erwig, "The graph Voronoi diagram with applications," *Networks*, vol. 36, no. 3, pp. 156–163, 2000.

[23] E. Gilbert, "Random subdivisions of space into crystals," *Annals of Mathematical Statistics*, vol. 33, pp. 958–972, 1962.

[24] J. Meijering, "Interface area, edge length, and number of vertices in crystal aggregates with random nucleation," *Philips Research Reports*, vol. 8, pp. 270–290, 1953.

[25] M. P. Quine and D. F. Watson, "Radial generation of n-dimensional Poisson procesess," *Journal of Applied Probability*, vol. 21, pp. 548–557, 1984.

[26] R. E. Miles, "The random division of space," *Advances in Applied Probability*, vol. 4, pp. 243–266, 1972.

[27] R. V. Ambartzumian, "Convex polygons and random tessellations," *Stochastic Geometry*, 1974.

[28] J. Møller, "Random tessellations in $\mathbb{R}^d$," *Advances in Applied Probability*, vol. 21, pp. 37–73, 1989.

[29] J. Mecke and L. Muche, "The Poisson Voronoi tessellation I. A basic identity," *Mathematische Nachrichten*, vol. 176, no. 1, pp. 199–208, 1995.

[30] L. Muche, "The Poisson Voronoi tessellation II. Edge length distribution functions," *Mathematische Nachrichten*, vol. 178, pp. 271–283, 1996.

[31] ——, "The Poisson Voronoi tessellation III. Miles' formula," *Mathematische Nachrichten*, vol. 191, no. 1, pp. 247–267, 1998.

[32] L. Heinrich and L. Muche, "Second-order properties of the point process of nodes in a stationary Voronoi tessellation," *Mathematische Nachrichten*, vol. 281, no. 3, pp. 350–375, 2008.

[33] H. J. Hilhorst, "Statistical properties of planar Voronoi tessellations," *The European Physical Journal B - Condensed Matter and Complex Systems*, vol. 64, no. 3-4, pp. 437–441, 2008.

[34] ——, "Planar Voronoi cells: the violation of Aboav's law explained," *Journal of Physics A: Mathematical and General*, vol. 39, pp. 7227–7243, 2006.

[35] ——, "New Monte Carlo method for planar Poisson-Voronoi cells," *Journal of Physics A: Mathematical and Theoretical*, vol. 40, pp. 2615–2638, 2007.

[36] P. Calka, "Precise formulae for the distributions of the principal geometric characteristics of the typical cells of a two-dimensional Poisson-Voronoi tessellation and a Poisson line process," *Advances in Applied Probability*, vol. 35, no. 3, pp. 551–562, 2003.

[37] P. N. Rathie, "On the volume distribution of the typical Poisson-Delaunay cell," *Journal of Applied Probability*, vol. 29, pp. 740–744, 1992.

[38] M. Tanemura, "Statistical distributions of Poisson Voronoi cells in two and three dimensions," *Forma*, vol. 18, no. 4, pp. 221–247, 2003.

[39] K. A. Brakke, "Statistics of random plane Voronoi tessellations," 2005.

[40] ——, "Statistics of three dimensional random Voronoi tessellations," 2005.

[41] ——, "Random Voronoi tessellations in arbitrary dimension," 2005.

[42] ——, "200,000,000 random Voronoi polygons," 2005.

[43] R. Cowan, "Properties of ergodic random mosaic processes," *Mathematische Nachrichten*, vol. 97, no. 1, pp. 89–102, 1978.

[44] ——, "A more comprehensive complementary theorem for the analysis of Poisson point processes," *Advances in Applied Probability*, vol. 38, no. 3, pp. 581–601, 2006.

[45] C. Lautensack and S. Zuyev, "Random Laguerre tessellations," *Advances in Applied Probability*, vol. 40, no. 3, 2008.

[46] S. Honiden, M. E. Houle, C. Sommer, and M. Wolff, "Approximate shortest path queries in graphs using Voronoi duals," *International Symposium on Voronoi Diagrams in science and engineering*, 2009.

[47] C. Demetrescu, A. V. Goldberg, and D. S. Johnson, "Implementation challenge for shortest paths," in *Encyclopedia of Algorithms*, 2008.

[48] R. Bauer and D. Delling, "SHARC: Fast and robust unidirectional routing," in *Workshop on Algorithm Engineering and Experiments*, 2008, pp. 13–26.

[49] R. Bauer, D. Delling, P. Sanders, D. Schieferdecker, D. Schultes, and D. Wagner, "Combining hierarchical and goal-directed speed-up techniques for Dijkstra's algorithm," in *Workshop on Experimental Algorithms*, 2008, pp. 303–318.

[50] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," arXiv:0706.1062, 2007.

[51] M. Mitzenmacher, "A brief history of generative models for power law and lognormal distributions," *Internet Mathematics*, vol. 1, no. 2, 2003.

[52] M. Ley, "The DBLP computer science bibliography: Evolution, research issues, perspectives," in *Symposium on String Processing and Information Retrieval*, 2002, pp. 1–10.

[53] L. Getoor, T. E. Senator, P. Domingos, and C. Faloutsos, Eds., *Knowledge Discovery and Data Mining (SIGKDD)*, 2003.

[54] Cooperative Association for Internet Data Analysis, "Router-level topology measurements," Online at http://www.caida.org/tools/measurement/skitter/router_topology/, 2003.

[55] A. Medina, A. Lakhina, I. Matta, and J. W. Byers, "Brite: An approach to universal topology generation," in *Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, 2001, pp. 346–.

[56] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[57] B. M. Waxman, "Routing of multipoint connections," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 9, pp. 1617–1622, 1988.

[58] L. Salwinski, C. S. Miller, A. J. Smith, F. K. Pettit, J. U. Bowie, and D. Eisenberg, "DIP, the database of interacting proteins," *Nucleic Acids Research*, vol. 30, no. 1, pp. 303–305, 2002.

[59] C. Stark, B. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, and M. Tyers, "BioGRID: a general repository for interaction datasets," *Nucleic Acids Research*, vol. 34, no. 1, pp. 535–539, 2006.

[60] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary Physics*, vol. 46, pp. 323–351, 2005.
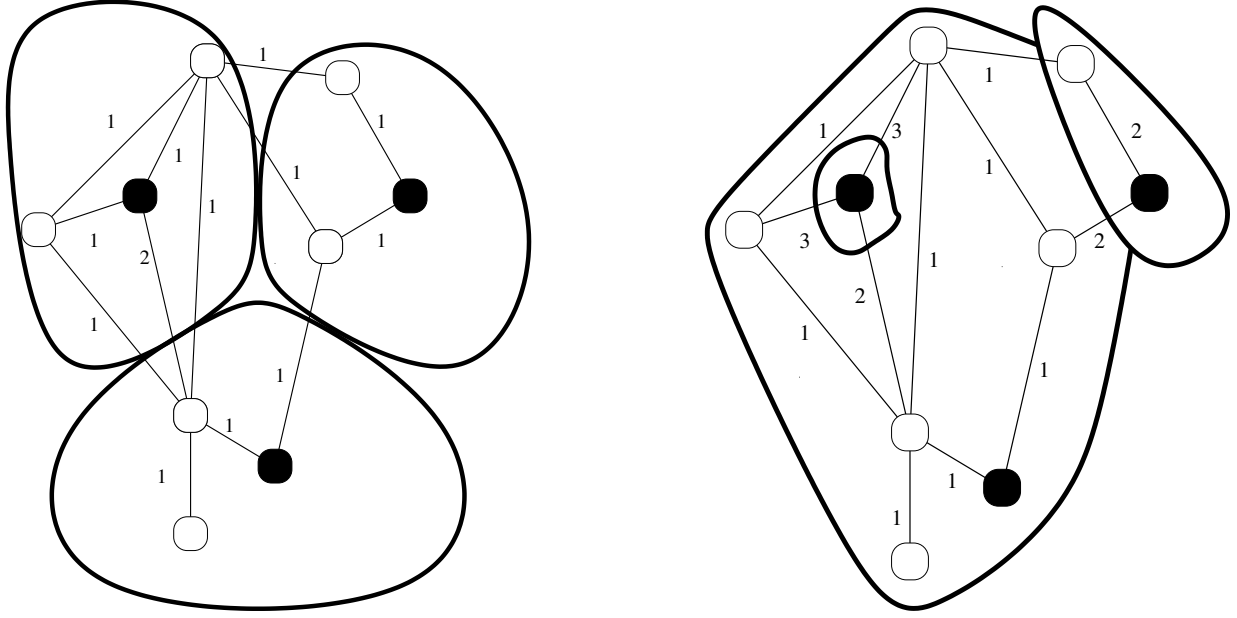
Figure 1. Two graph Voronoi diagrams for the same planar graph but with different edge weights. Voronoi nodes are black and the remaining nodes are white. Even though the graphs are structurally equivalent, the corresponding graph Voronoi diagrams are not.

Graph	$|V|$	$|E|$	Degree	Vor. Dual Degree	Vor. Region Size
BIOGRID	4039	43854	10.855±13.435	12.762±11.539	0.294±0.333
ASBarabasi	10000	39994	3.999± 6.529	27.571±34.150	0.156±0.206
ASWaxman	10000	40000	4.000± 2.486	46.792±33.707	0.286±0.240
RTBarabasi	10000	39994	3.999± 6.159	30.060±34.584	0.173±0.217
RTWaxman	10000	40000	4.000± 2.479	47.294±34.006	0.292±0.246
dip20090126	19928	82406	4.135± 5.842	36.522±36.825	0.174±0.248
hep-th-cite	27400	705084	25.732±45.558	80.889±94.295	0.114±0.168
itdk0304	190914	1215220	6.365±14.181	19.951±59.507	0.050±0.115
dblp20080824	511163	3742140	7.321±11.785	42.944±93.010	0.047±0.100

Table V
AFTER RESAMPLING: VARIATION OF REGION SIZE AND DISTRIBUTION FOR NINE POWER-LAW GRAPHS.

Graph	$|V|$	$|E|$	Degree	Vor. Dual Degree	Vor. Region Size
RMV	2277812	3416597	3.000±0.013	13.156± 8.671	0.307±0.387
VBB	2599953	3899807	3.000±0.014	13.568±10.633	0.289±0.382
LongDistance	1215750	1823609	3.000±0.007	14.284± 8.659	0.326±0.360

Table VI
AFTER RESAMPLING: VARIATION OF REGION SIZE AND DISTRIBUTION FOR 3 EUROPEAN PUBLIC TRANSPORTATION GRAPHS.