Large-Scale Density-Friendly Decomposition via Convex Programming [1]

Mauro Sozio

Telecom ParisTech

January 16, 2020

LP formulation for Densest Subgraph

$$
\mathrm{LP}(G) \qquad \max \sum_{e \in E} w_e x_e
$$

$$
\begin{aligned}
\text{s.t.} \qquad & x_e \le y_u, \quad \forall u \in e \\
& \sum_{u \in V} y_u = 1, \\
& x_e, y_u \ge 0, \quad \forall u \in V, e \in E
\end{aligned}
$$

Dual of the LP

$$
\begin{array}{lll}
\mathrm{DP}(G) & & \\
\text{min} & \rho & \\
\text{s.t.} & \rho \geq \sum_{e:u \in e} \alpha_u^e, & \forall u \in V \\
& \sum_{u \in e} \alpha_u^e \geq 1, & \forall e \in E \\
& \alpha_u^e \geq 0, & \forall u \in e \in E
\end{array}
$$

Convex Program $\mathrm{CP}(G)$

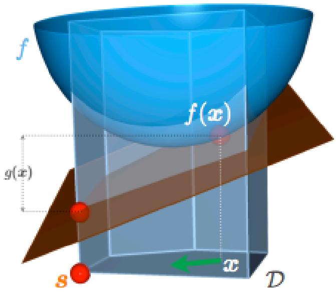
Let $Q_G(\alpha) := \sum_{u \in V} r(u)^2$, where (r, α) is an invariant pair, i.e., $r(u) = \sum_{e \in E : u \in e} \alpha_u^e$ and $\alpha_u^e + \alpha_v^e = 1$ for every $e = uv \in E$.

The convex program $\mathrm{CP}(G)$ is defined as follows:

$\mathrm{CP}(G) := \min\{Q_G(\alpha) : \alpha \text{ is feasible for } \mathrm{DP}(G)\}.$

The Frank-Wolfe Algorithm [2]

1: **Input:** function f convex and continuously differentiable, a compact convex set $\mathcal{D}$, integer T
2: Let $x^{(0)} \in \mathcal{D}$
3: **for** $t = 1, \ldots, T$ **do**
4: $\quad \gamma_t \leftarrow \frac{2}{t+2}$
5: $\quad$ Compute $s := \arg\min_{s \in \mathcal{D}} \langle s, \nabla f(x^{(k)}) \rangle$
6: $\quad$ Update $x^{(k+1)} = (1 - \gamma) x^{(k)} + \gamma s$



Frank-Wolfe-Based Algorithm for Densest Subgraph

for each $e = uv$ in E in parallel **do**
    $\alpha_u^{e(0)}, \alpha_v^{e(0)} \leftarrow \frac{1}{2}$
for each $u \in V$ in parallel **do**
    $r^{(0)}(u) \leftarrow \sum_{e \in E: u \in e} \alpha_u^{e(0)}$
for each iteration $t = 1, \ldots, T$ **do**
    $\gamma_t \leftarrow \frac{2}{t+2}$
    for each e in E in parallel **do**
        $x \leftarrow \arg\min_{v \in e} r^{(t-1)}(v)$
        for each $u \in e$ **do**
            $\widehat{\alpha}_u^e \leftarrow 1$, if $u = x$ and 0 otherwise.
    $\alpha^{(t)} \leftarrow (1 - \gamma_t) \cdot \alpha^{(t-1)} + \gamma_t \cdot \widehat{\alpha}$
    for each $u \in V$ in parallel **do**
        $r^{(t)}(u) \leftarrow \sum_{e \in E: u \in e} \alpha_u^{e(t)}$
return $(\alpha^{(t)}, r^{(t)})$

Convergence

Theorem 1

(Convergence of the Frank-Wolfe-based Algorithm.) *Let* $G = (V, E)$ *be an undirected graph with maximum degree* Δ*. Let* (r^*, α^*) *be an invariant pair for* G *where* α^* *is an optimal solution for* $\mathrm{CP}(G)$*. In Algorithm 2, for any* $\epsilon > 0$ *for any* $t > \frac{4\Delta|E|}{\epsilon^2}$*, we have* $||r^{(t)} - r^*||_2 \leq \epsilon$*.*

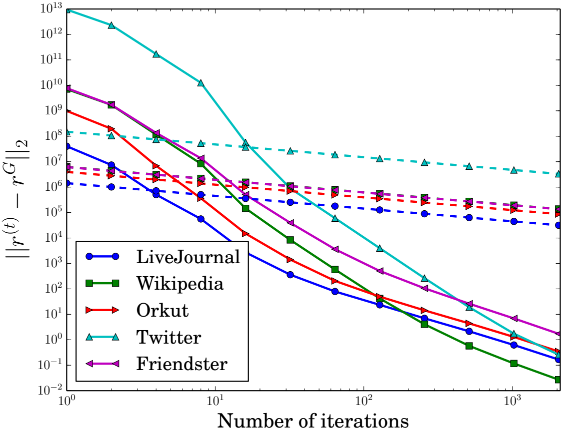
Experiments: Settings

networks	n	m
LiveJournal	4,036,538	34,681,189
Wikipedia	2,080,370	42,336,692
Orkut	3,072,627	117,185,083
Twitter	52,579,683	1,614,106,500
Friendster	124,836,180	1,806,067,135
gsh-2015	988,490,691	25,690,705,119

Table: Our set of large graphs.

Used a linux machine with 2 processors Intel Xeon CPU E5-2660 @ 2.60 GHz with 10 cores split in 2 threads each, as well as 64G of RAM DDR4 2133 MHz. We employ 10 threads.

Convergence to the r^* vector



where r^t is the r vector at step t and $r^G = r^*$.

Densest Subgraph

How to extract the densest subgraph from a sufficiently good r^t? Use the fact that for any graph H_1, H_2, $|\rho(H_1) - \rho(H_2)| \geq \frac{1}{|V|^2}$.

Densest Subgraph

How to extract the densest subgraph from a sufficiently good r^t? Use the fact that for any graph H_1, H_2, $|\rho(H_1) - \rho(H_2)| \geq \frac{1}{|V|^2}$.

From Thm. 1, after some step the nodes in a densest subgraph will have max r^t score. Sort the r^t's non-increasingly: $r(v_1)^t \geq r(v_2)^t, \ldots, \geq r(v_n)^t$. Recall that $r(v_1)^t$ gives an upper bound on the max density. Let G_k be the graph induced by $v_1, \ldots, v_k$. As soon as we find a graph G_k such that $|\rho(G_k) - r(v_1)^t| < \frac{1}{n^2}$ we know that G_k is densest.

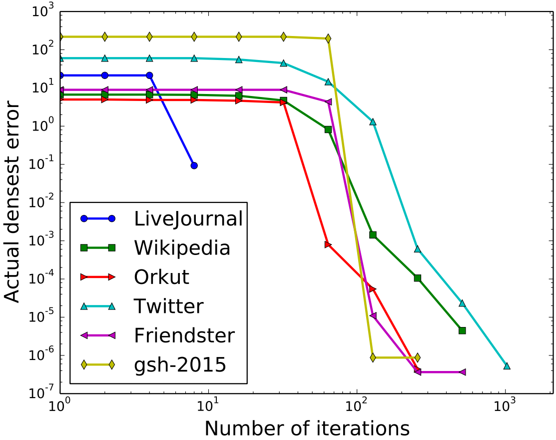
Densest Subgraph

How to extract the densest subgraph from a sufficiently good r^t? Use the fact that for any graph H_1, H_2, $|\rho(H_1) - \rho(H_2)| \geq \frac{1}{|V|^2}$.

From Thm. 1, after some step the nodes in a densest subgraph will have max r^t score. Sort the r^t's non-increasingly: $r(v_1)^t \geq r(v_2)^t, \ldots, \geq r(v_n)^t$. Recall that $r(v_1)^t$ gives an upper bound on the max density. Let G_k be the graph induced by $v_1, \ldots, v_k$. As soon as we find a graph G_k such that $|\rho(G_k) - r(v_1)^t| < \frac{1}{n^2}$ we know that G_k is densest.

In practice: as soon as we find a "sufficiently" small stable subset H we compute a densest subgraph in H via the LP-based algorithm or maximum flow. This works well...

Computation of the densest subgraph



where r^t is the r vector at step t and $r^G = r^*$.

K-Core and Density-Friendly Decomposition

- k-core decomposition: compute for each node v the largest integer c_v such that v is in an induced subgraph with minimum degree c_v.
- k-core decomp. might reveal the structural organization of a graph.
- It has been applied to the analysis of the internet topology [5], social network analysis [8], bionformatics [7], analysis of the human brain [4], as well as influence analysis [3]. There is also a startup based in NYC using such an algorithm `http://www.kcore-analytics.com/`.
- In the k-core decomposition, outer cores might be denser than inner cores, which is not ideal.
- Here, we will show that the Frank-Wolfe based algorithm actually computes a so-called diminishingly-dense decomposition where inner cores are always denser than outer cores.

Quotient Graph

Definition 2 (Quotient Graph)

Given an undirected graph $G = (V, E)$, and a subset $B \subseteq V$, the quotient graph of G with respect to B is a weighted graph $G \setminus B = (\widehat{V}, \widehat{E}, \widehat{w})$, which is defined as follows.

- $\widehat{V} := V \setminus B$.
- $\widehat{E} := \{e \cap \widehat{V} : e \in E, e \cap \widehat{V} \neq \emptyset\}$, i.e., every edge $e \in E$ not contained in B contributes towards $\widehat{E}$.
- For $e' \in \widehat{E}$, $\widehat{w}(e') := |\{e \in E : e' = e \cap \widehat{V}\}|$.

Diminishingly-dense Decomposition

Definition 3 (Diminishingly-dense Decomposition)

Given an undirected graph $G = (V, E, w)$, we define the diminishingly-dense decomposition $\mathcal{B}$ of G as the sequence $\emptyset = B_0 \subsetneq B_1 \subsetneq B_2 \subsetneq \cdots \subsetneq B_k = V$ as follows:

Diminishingly-dense Decomposition

Definition 3 (Diminishingly-dense Decomposition)

Given an undirected graph $G = (V, E, w)$, we define the diminishingly-dense decomposition $\mathcal{B}$ of G as the sequence $\emptyset = B_0 \subsetneq B_1 \subsetneq B_2 \subsetneq \cdots \subsetneq B_k = V$ as follows:

Initially, we set $B_0 := \emptyset$ and $G_0 := G$. For $i \geq 1$, if $B_{i-1} = V$, the decomposition is fully defined.

Diminishingly-dense Decomposition

Definition 3 (Diminishingly-dense Decomposition)

Given an undirected graph $G = (V, E, w)$, we define the diminishingly-dense decomposition $\mathcal{B}$ of G as the sequence $\emptyset = B_0 \subsetneq B_1 \subsetneq B_2 \subsetneq \cdots \subsetneq B_k = V$ as follows:

Initially, we set $B_0 := \emptyset$ and $G_0 := G$. For $i \geq 1$, if $B_{i-1} = V$, the decomposition is fully defined.

Otherwise, let $G_i := G_{i-1} \setminus B_{i-1} = (V_i, E_i, w_i)$ be the quotient graph of G_{i-1} with respect to B_{i-1}. Let S_i be the maximal densest subset in G_i (with respect to w_i). We define $B_i := B_{i-1} \cup S_i$. For each $i = 1, \ldots, k$, we denote $r_i = \rho_i(S_i)$. Moreover, we define the *maximal density vector* $r^G \in \mathbb{R}^V$ such that if $u \in S_i$, then $r^G(u) := r_i$.

Properties of the Decomposition

Lemma 4 (the decomposition is unique)

Given a graph G, there is a unique diminishingly-dense decomposition.

Properties of the Decomposition

Lemma 4 (the decomposition is unique)

Given a graph G, there is a unique diminishingly-dense decomposition.

Lemma 5 (Diminishing r_i's)

In the diminishingly-dense decomposition in Definition 3, if $B_i \subsetneq V$, then

$$r_i > r_{i+1} \quad i = 1, \ldots, k-1.$$

Lemma 6 (Diminishing Densities)

In the diminishingly-dense decomposition in Definition 3, if $B_i \subsetneq V$, then

$$\rho(B_i) > \rho(B_{i+1}) \quad i = 1, \ldots, k-1.$$

References I

[1] Maximilien Danisch, Hubert Chan, Mauro Sozio
Large Scale Density-Friendly Decomposition via Convex Programming
Proceedings of the 26th international conference on World wide web. ACM, 2017.

[2] Jaggi, Martin.
Revisiting Frank-Wolfe: Projection-Free Sparse Convex Optimization.
ICML, 2013.

[3] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse.
Identification of influential spreaders in complex networks.
Nature physics, 6(11):888?893, 2010.

[4] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns.
Mapping the structural core of human cerebral cortex.
PLoS, Biology, 6(7):888?893, 2008

References II

[5] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir.
A model of internet topology using k-shell decomposition.
Proceedings of the National Academy of Sciences, 104(27):11150?11154, 2007.

[6] Tatti, Nikolaj, and Aristides Gionis.
Density-friendly graph decomposition.
Proceedings of the 24th International Conference on World Wide Web. ACM, 2015.

[7] G. Bader and C. Hogue.
An automated method for finding molecular complexes in large protein interaction networks.
BMC Bioinformatics, 4(1), 2003.

[8] Seidman, Stephen B.
Network structure and minimum degree.
Social networks 5.3 (1983): 269-287.