Imposing contiguity constraints in political districting models

Hamidreza Validi[1] Austin Buchanan[1] Eugene Lykhovyd[2]

[1]Industrial Engineering & Management
Oklahoma State University

[2]Industrial and Systems Engineering
Texas A&M University

AMS Sectional Meeting, University of Virginia, March 2020

Why optimization?

Use (exact) optimization to determine what is *possible*.

1. How many counties can be kept intact? (NC Whole County Provision)
2. How many communities can. . . ? (CA Citizens Redistricting Commission)
3. How compact can it get? (VRA districts should be compact by *Gingles*)
4. Can multiple criteria be satisfied simultaneously?

Before answering these questions, need to be able to handle contiguity.

Is contiguity hard to handle?

Ricca et al. (2008):

> *[Contiguity] constraints make [districting] much more difficult than other partitioning problems in combinatorial optimization, such as coloring or frequency assignment.*

Ricca et al. (2013):

> *[Contiguity] is particularly difficult to deal with and, sometimes, it is even discarded from [political districting] models and considered only a posteriori.*

Goderbauer and Winandy (2017):

> *Ensuring contiguity efficiently seems to be an issue in exact methods [for political districting].*

Swamy et al. (2019):

> *For exact methods, contiguity enforcement has been a major challenge.*

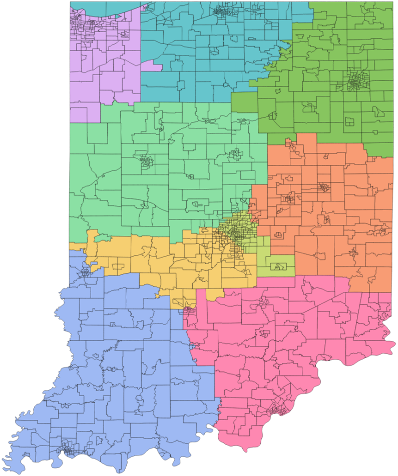
Validi et al. (2020):

Not really.

We solve a classical integer program for:

- all county-level instances in USA
- half of tract-level instances

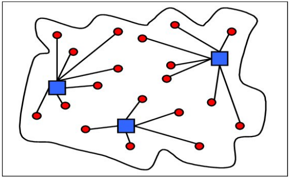
Largest solved is Indiana ($n = 1,511$).



A stylized problem to analyze

SW Hess, JB Weaver, HJ Siegfeldt, JN Whelan, and PA Zitlau. Nonpartisan political redistricting by computer. *Operations Research*, 13(6):998-1006, 1965.

OR volunteers developed a compactness measure and a "warehouse-location" heuristic to draw nonpartisan, Constitutional political districts. The heuristic maps compact and contiguous districts of equal population. The minimization criterion and compactness measure is population moment of inertia–the summed squared distances from each person to his district's center. The districting method is particularly useful when legislative impasse or indifference forces courts to intervene. Federal Courts have received a computer plan for possible use in Delaware and have asked for computer districts in Connecticut.



The Hess model

Model input:

k	number of districts desired
L	least allowable population in a district
U	most allowable population in a district
p_i	population in county i
d_{ij}	distance between county i and county j
$G = (V, E)$	undirected graph representing county adjacencies
$n := \|V\|$	

Decision variables:

$$x_{ij} = \begin{cases} 1 & \text{if vertex } i \text{ is assigned to (the district centered at) vertex } j \\ 0 & \text{otherwise} \end{cases}$$

Note: $x_{jj} = 1 \iff$ county j is chosen as a district center

The Hess model

Minimize moment-of-inertia subject to k districts and population balance.

$$\min \sum_{i \in V} \sum_{j \in V} \underbrace{p_i d_{ij}^2}_{w_{ij}} x_{ij} \tag{1a}$$

$$\sum_{j \in V} x_{ij} = 1 \qquad \forall i \in V \tag{1b}$$

$$\sum_{j \in V} x_{jj} = k \tag{1c}$$

$$L x_{jj} \leq \sum_{i \in V} p_i x_{ij} \leq U x_{jj} \qquad \forall j \in V \tag{1d}$$

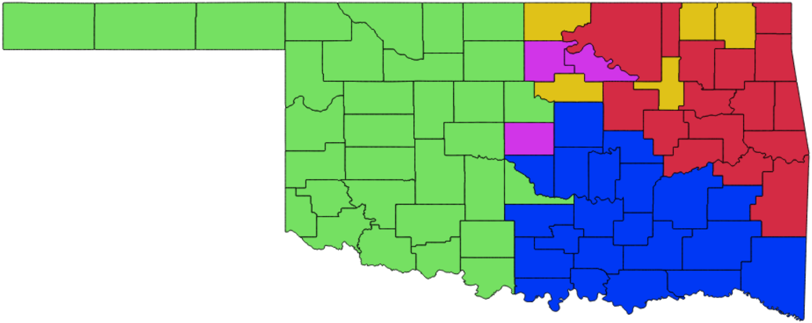
$$x_{ij} \leq x_{jj} \qquad \forall i, j \in V \tag{1e}$$

$$x_{ij} \in \{0, 1\} \qquad \forall i, j \in V. \tag{1f}$$

Feasible region for linear programming relaxation:

$$\mathcal{P}_{\mathrm{HESS}} := \left\{ x \in \mathbb{R}_+^{n \times n} \;\middle|\; x \text{ satisfies constraints (1b), (1c), (1d), (1e)} \right\}.$$

With 1% population deviation:



$$\bar{p} = \sum_{i \in V} p_i / k$$

$$L = 0.995\bar{p}$$

$$U = 1.005\bar{p}$$

Models for imposing contiguity

MIP	LP relaxation	description
SHIR	$\mathcal{P}_{\text{SHIR}}$	flow-based model by Shirabe (2005, 2009)
MCF	$\mathcal{P}_{\text{MCF}}$	new flow-based model
CUT	$\mathcal{P}_{\text{CUT}}$	cut-based model by Oehrlein and Haunert (2017)
LCUT	$\mathcal{P}_{\text{LCUT}}$	new cut-based model

Important to keep in mind:

- strength of LP bounds (affects # of branch-and-bound nodes)
- # of variables, constraints, and nonzeros (affects LP solve time)

Theorem (Strength)

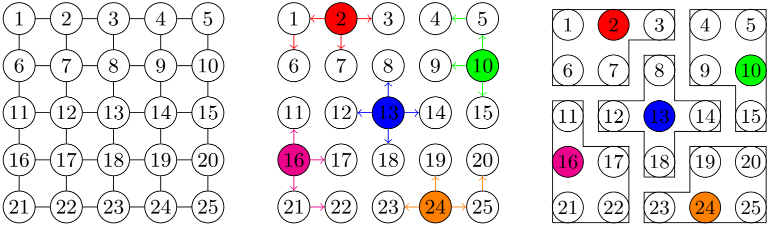
For every instance of districting,

$$\mathcal{P}_{\text{LCUT}} \subseteq \mathcal{P}_{\text{CUT}} = \text{proj}_x \, \mathcal{P}_{\text{MCF}} \subseteq \text{proj}_x \, \mathcal{P}_{\text{SHIR}},$$

and there exist instances for which the inclusions are strict.

SHIR

Suppose $k = L = U = 5$ and $p = \mathbf{1}$.



Model uses variables

$$f_{ij}^{v} = \text{the amount of flow, originating at district center } v\text{, that is sent across edge } (i,j).$$

SHIR

Credited to Shirabe (2005, 2009); detailed by Oehrlein and Haunert (2017)

$$
\begin{aligned}
& x \in \mathcal{P}_{\text{HESS}} && && \text{(2a)}\\
& f^j(\delta^-(i)) - f^j(\delta^+(i)) = x_{ij} && \forall i \in V \setminus \{j\},\ \forall j \in V && \text{(2b)}\\
& f^j(\delta^-(i)) \leq (n-1)x_{ij} && \forall i \in V \setminus \{j\},\ \forall j \in V && \text{(2c)}\\
& f^j(\delta^-(j)) = 0 && \forall j \in V && \text{(2d)}\\
& f^v_{ij} \geq 0 && \forall (i,j) \in A,\ \forall v \in V. && \text{(2e)}
\end{aligned}
$$

$$
\mathcal{P}_{\text{SHIR}} := \left\{ (x,f) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{2mn} \;\middle|\; (x,f) \text{ satisfies (2a)} - \text{(2e)} \right\}.
$$

MCF

New variables:

$$f_{ij}^{ab} = \begin{cases} 1 & \text{if edge } (i,j) \in A \text{ is on the path to vertex } a \text{ from its district's center } b \\ 0 & \text{otherwise.} \end{cases}$$

New multicommodity flow model:

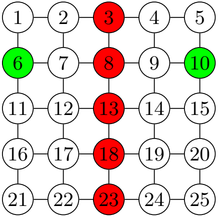
$$
\begin{aligned}
& x \in \mathcal{P}_{\text{HESS}} && && \text{(3a)} \\
& f^{ab}(\delta^+(b)) - f^{ab}(\delta^-(b)) = x_{ab} && \forall a \in V \setminus \{b\},\ \forall b \in V && \text{(3b)} \\
& f^{ab}(\delta^+(i)) - f^{ab}(\delta^-(i)) = 0 && \forall i \in V \setminus \{a,b\},\ \forall a \in V \setminus \{b\},\ \forall b \in V && \text{(3c)} \\
& f^{ab}(\delta^-(b)) = 0 && \forall a \in V \setminus \{b\},\ \forall b \in V && \text{(3d)} \\
& f^{ab}(\delta^-(j)) \leq x_{jb} && \forall j \in V \setminus \{b\},\ \forall a \in V \setminus \{b\},\ \forall b \in V && \text{(3e)} \\
& f_{ij}^{ab} \geq 0 && \forall (i,j) \in A,\ \forall a \in V \setminus \{b\},\ \forall b \in V. && \text{(3f)}
\end{aligned}
$$

$$\mathcal{P}_{\text{MCF}} := \left\{ (x,f) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{2mn(n-1)} \;\middle|\; (x,f) \text{ satisfies (3a) -- (3f)} \right\}.$$

CUT

Definition

A subset $C \subseteq V \setminus \{a, b\}$ of vertices is called an a, b-separator for $G = (V, E)$ if there is no a, b-path in $G - C$.



Model by Oehrlein and Haunert (2017):

$$x \in \mathcal{P}_{\mathrm{HESS}} \tag{4a}$$

$$x_{ab} \leq \sum_{c \in C} x_{cb} \qquad \forall (a, b, C). \tag{4b}$$

$$\mathcal{P}_{\mathrm{CUT}} := \left\{ x \in \mathbb{R}^{n \times n} \;\middle|\; x \text{ satisfies (4a)} - \text{(4b)} \right\}.$$

Proposition (CUT separation)

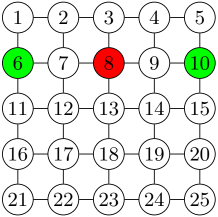
The separation problem for CUT is solvable in time

- $O(n^2 \log^3 n)$ *when* x^* *is fractional;*
- $O(n^2)$ *when* x^* *is integer.*

LCUT

Definition

A subset $C \subseteq V \setminus \{a, b\}$ of vertices is called a length-U a, b-separator in $G = (V, E)$, with respect to vertex weights p, if $\mathrm{dist}_{G-C,p}(a, b) > U$.



$$x \in \mathcal{P}_{\mathrm{HESS}} \tag{5a}$$

$$x_{ab} \leq \sum_{c \in C} x_{cb} \qquad \forall (a, b, C). \tag{5b}$$

$$\mathcal{P}_{\mathrm{LCUT}} := \left\{ x \in \mathbb{R}^{n \times n} \;\middle|\; x \text{ satisfies (5a)} - \text{(5b)} \right\}.$$

Theorem (LCUT separation)

The separation problem for LCUT is

- *NP-hard when x^* is fractional, and*
- *solvable in time $O(n^2)$ when x^* is integer.*

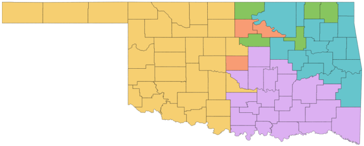
Computational setup:

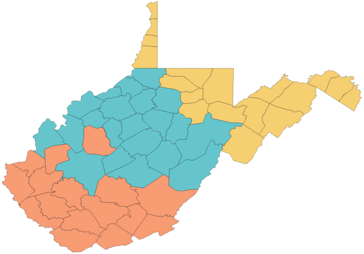
- Intel Xeon E3-1270 v6 “Kaby Lake” 3.80 GHz CPU with 8 cores and 32 GB RAM
- Gurobi 8.1.1 (8 threads, 10 GB per thread, concurrent, zero MIP gap)
- 3600 second time limit (TL)
- Use 2010 Census data; processed using QGIS; map projections from EPSG dataset

Table: A breakdown of the 50 county-level instances

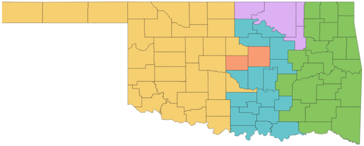
Description	#	States
Overtly infeasible	27	AZ, CA, CT, FL, GA, HI, IL, IN, KY, MA, MC, MI, MN, MO, NC, NJ, NV, NY, OH, PA, RI, TN, TX, UT, VA, WA, WI.
Trivial ($k = 1$)	7	AK, DE, MT, ND, SD, VT, WY.
Remaining	16	AL, AR, CO, IA, ID, KS, LA, ME, MS, NE, NH, NM, OK, OR, SC, WV.

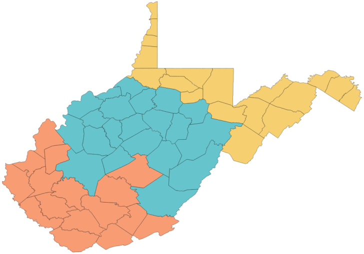
Optimal solution w/o contiguity





Optimal solution w/ contiguity





County-level MIP times

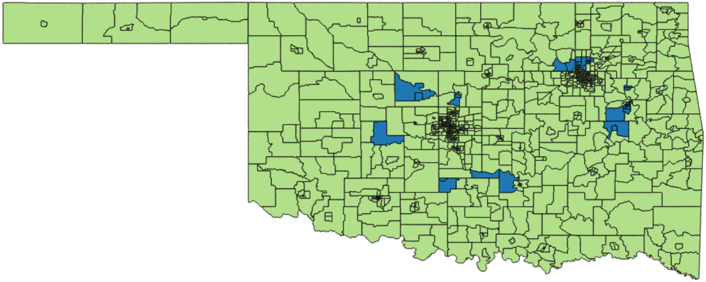
State	n	k	Hess	MCF	SHIR	CUT	LCUT
NH	10	2	0.03	0.23	0.12	0.12	0.08
ME	16	2	0.03	2.42	0.78	0.72	0.66
NM	33	3	0.07	0.63	0.11	0.10	0.10
OR	36	5	0.40	55.80	0.59	3.08	0.32
ID	44	2	0.06	2.79	0.14	0.09	0.09
SC	46	7	8.67	TL	TL	TL	TL
WV	55	3	0.82	62.16	1.68	0.86	0.84
LA	64	6	2.14	66.87	4.08	2.22	2.23
CO	64	7	2.46	215.25	2.74	TL	1.94
AL	67	7	3.83	457.52	64.29	20.94	19.41
AR	75	4	0.53	29.04	1.62	0.96	0.96
OK	77	5	1.01	101.64	5.06	1.64	1.97
MS	82	4	0.41	8.31	0.54	0.49	0.49
NE	93	3	0.28	13.33	0.57	0.45	0.47
IA	99	4	0.85	29.23	1.47	0.87	0.87
KS	105	4	1.34	48.98	1.78	1.10	1.11

Tract-level instances are bigger; harder to solve out-of-the-box.

			Hess			SHIR		
State	n	k	Vars	LP time	IP time	Vars	LP time	IP time
NH	295	2	87,025	76.73	77.13	551,060	160.30	180.17
NM	499	3	249,001	930.61	1167.75	1,576,175	1560.83	2434.04
AR	686	4	470,596	2042.92	3428.91	3,007,159	TL	TL
OR	834	5	695,556	TL	TL	4,553,181	TL	TL
OK	1046	5	1,094,116	TL	TL	6,842,685	TL	TL

Bottleneck is solving the root LP. How to speed up?

Helpful observation: most x_{ij} variables have no chance of being used.



Make rigorous with Lagrangian reduced cost fixing:
96% of the x_{jj} variables fixed to zero
99% of the x_{ij} variables fixed to zero

Proposition

For any multipliers $\alpha, \lambda, \upsilon \in \mathbb{R}^n$, the following Lagrangian lower bounds the contiguity-constrained Hess model in which x_{st} is fixed to one.

$$\min \sum_{i \in V} \sum_{j \in V} w_{ij} x_{ij} + \sum_{i \in V} \alpha_i \left(1 - \sum_{j \in V} x_{ij}\right)$$

$$+ \sum_{j \in V} |\lambda_j| \left(L x_{jj} - \sum_{i \in V} p_i x_{ij}\right) + \sum_{j \in V} |\upsilon_j| \left(\sum_{i \in V} p_i x_{ij} - U x_{jj}\right)$$

$$\sum_{j \in V} x_{jj} = k$$

$$x_{ij} \leq x_{jj} \qquad \forall i, j \in V$$

$$x_{ij} \in \{0, 1\} \qquad \forall i, j \in V$$

there is a path from s to t across nodes assigned to t

$x_{st} = 1.$

Proposition

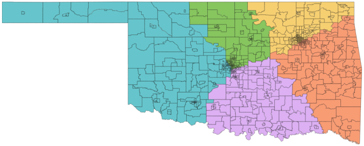
For given $(\alpha, \lambda, \upsilon)$, the bounds for all $s, t \in V$ can be found in time $O(n^2)$.

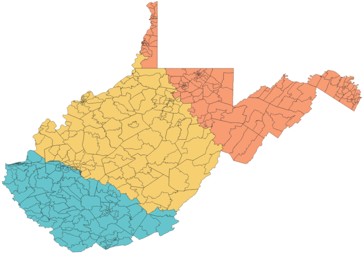
Lagrangian reduced cost fixing

Table: Time for heuristic and Lagrangian, and % of variables fixed.

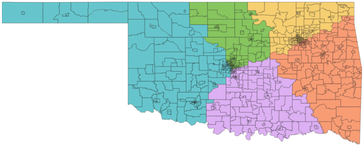
			Lagrangian		Heuristic w/o contiguity			Heuristic w/ contiguity		
State	n	k	LB	time	UB	time	fixed	UB	time	fixed
NH	295	2	2,687.72	0.09	2,689.07	0.18	100	2,688.28	0.30	100
NM	499	3	31,575.28	0.53	31,598.10	1.72	95	31,608.19	1.91	95
AR	686	4	15,490.50	1.02	15,565.41	3.85	96	15,569.97	4.09	97
OR	834	5	26,742.01	1.87	26,750.23	9.55	98	26,750.15	10.36	98
OK	1,046	5	19,106.02	3.34	19,132.60	3.60	99	19,131.99	4.35	99

Optimal solution w/o contiguity





Optimal solution w/ contiguity



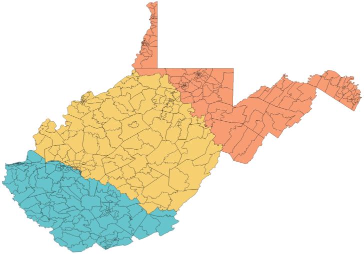


Table: Some MIP results under 3,600 second time limit

State	n	k	Hess obj	Hess time	SHIR obj	SHIR time	CUT obj	CUT time
NH	295	2	2,688.28	0.03	2,688.28	0.13	2,688.28	0.01
NM	499	3	31,598.10	24.79	31,603.15	216.76	31,603.15	44.78
AR	686	4	15,563.52	17.72	15,569.97	119.47	15,569.97	14.99
OR	834	5	26,749.59	81.37	26,749.74	733.67	26,749.74	83.74
OK	1,046	5	19,107.72	2.09	19,107.72	17.94	19,107.72	1.65

Table: MIP summary under 3,600 second time limit

Status	Hess	SHIR	CUT
# MIPs solved	16	15	16
# LPs solved (but not MIP)	11	6	13
# LPs not solved	16	22	14

Conclusion

- SHIR works well on county-level instances
- CUT/LCUT are generally faster
- Contiguity need not be the bottleneck.

preprint

GitHub repo

Future work

- Solve all tract-level instances
- Preserving communities/counties; other compactness measures; . . .
- What would you do with exact methods?

References

1 Rodolfo Carvajal, Miguel Constantino, Marcos Goycoolea, Juan Pablo Vielma, and Andrés Weintraub. Imposing connectivity constraints in forest planning models. *Operations Research*, 61(4):824–836, 2013.

2 Sebastian Goderbauer and Jeff Winandy. Political districting problem: Literature review and discussion with regard to federal elections in Germany, 2018.

3 SW Hess, JB Weaver, HJ Siegfeldt, JN Whelan, and PA Zitlau. Nonpartisan political redistricting by computer. *Operations Research*, 13(6):998–1006, 1965.

4 Johannes Oehrlein and Jan-Henrik Haunert. A cutting-plane method for contiguity-constrained spatial aggregation. *Journal of Spatial Information Science*, 2017(15):89–120, 2017.

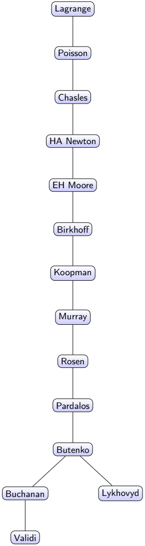
5 Federica Ricca and Bruno Simeone. Local search algorithms for political districting. *European Journal of Operational Research*, 189(3):1409–1426, 2008.

6 Federica Ricca, Andrea Scozzari, and Bruno Simeone. Political districting: from classical models to recent approaches. *Annals of Operations Research*, 204(1):271–299, 2013.

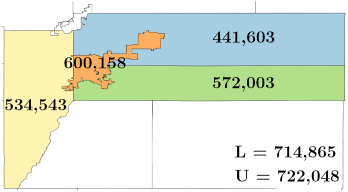
7 Takeshi Shirabe. A model of contiguity for spatial unit allocation. *Geographical Analysis*, 37(1):2–16, 2005.

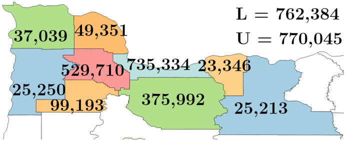
8 Takeshi Shirabe. Districting modeling with exact contiguity constraints. *Environment and Planning B: Planning and Design*, 36(6):1053–1066, 2009.

9 Rahul Swamy, Douglas M. King, and Sheldon H. Jacobson. Multi-objective optimization for political districting: a scalable multilevel approach, 2019.



CO and OR are infeasible at the county level. Why?





Denver County (pop. 600,158) and Maltnomah County (pop. 735,334).

Table: Some MIP results under no time limit

State	n	k	Hess obj	Hess time	CUT obj	CUT time
WV	484	3	11,801.26	27,056.92	11,834.84	28,984.58
MN	1,338	8	24,220.16	2,204.75	24,221.70	92,312.17
MD	1,406	8	5,084.32	36,553.32	5,084.44	75,539.58
WI	1,409	8	16,311.75	173,216.16	16,312.72	293,674.57
IN	1,511	9	11,081.72	10,570.75	11,081.72	13,435.36