Computing Spanning Trees in a Social Choice Context

Andreas Darmann*, Christian Klamler* and Ulrich Pferschy**

*Institute of Public Economics, University of Graz
**Institute of Statistics and Operations Research, University of Graz

2nd International Workshop on Computational Social Choice, Liverpool 2008

Introduction

Combine Social Choice Theory with Discrete Optimization

- Given: individuals' preferences over edges of a graph
- Aim: Find a "socially best" spanning tree in the graph
- Applications:
  - oil pipeline construction
  - water network construction in a village
- Group ranking of edges may or may not allocate numerical values to the edges

Main result:

1. sets of best spanning trees for the discussed ranking rules coincide
2. a best spanning tree for each considered ranking rule can be determined efficiently

Formal Framework

- Undirected Graph $G = (V, E)$
- $T \subseteq E$ is a spanning tree $:\Longleftrightarrow$ subgraph (V, T) of G is acyclic and connected
- τ...set of spanning trees of G
- finite set of individuals $I = \{1, 2, ..., k\}$,
  linear orders P_i on E, $1 \leq i \leq k$
- $\pi = (P_1, P_2, ..., P_k)$ is a (voter) preference profile
- Complete order $\succsim$ on E:
  asymmetric part $\succ$ and symmetric part $\sim$
- Complete order $\trianglerighteq$ on τ:
  asymmetric part $\triangleright$ and symmetric part $\bowtie$

Definition

Let $\trianglerighteq$ be a complete order on τ.
$T \in \tau$ is a best tree with respect to $\trianglerighteq$ $:\Longleftrightarrow \nexists T' \in \tau : \; T' \triangleright T$

Examples

Let $w(e) \in \mathbb{R}$ for all $e \in E$.

- Minimum Spanning Tree Problem is to determine a best tree with respect to the relation $T_1 \trianglerighteq T_2 :\Longleftrightarrow \sum_{e \in T_1} w(e) \leq \sum_{e \in T_2} w(e)$.
- Maximum Spanning Tree Problem: find a best tree w.r.t. relation $T_1 \trianglerighteq T_2 :\Longleftrightarrow \sum_{e \in T_1} w(e) \geq \sum_{e \in T_2} w(e)$.

Basic orders on the edge-set

Definitions

- Borda's method (see [BF02]):
  Individual i's Borda count of edge e is given by $B_i(e) := |\{f \in E : eP_if\}|$. The total Borda count of edge e is defined by $B(e) := \sum_{i \in I} B_i(e)$. For $e, f \in E$ we define the Borda-order on E by $e \succsim_b f :\Longleftrightarrow B(e) \geq B(f)$.

- Simple Majority-order (see [BF02]):
  Let $e, f \in E$. Then we define the Simple Majority-order on E by $e \succsim_{sm} f :\Longleftrightarrow |\{i \in I : eP_if\}| \geq |\{i \in I : fP_ie\}|$.

Definitions

Let $e, f \in E$.
For all $i \in I$, partition edge-set E into a set $S_i \subset E$ of edges individual i approves of and a set $E \setminus S_i$ individual i disapproves of.

- Approval-order (see [BF83]):
  The Approval count of e is defined by $A(e) := |\{i \in I : e \in S_i\}|$.
  The Approval-order $\succsim_a$ is then defined by $e \succsim_a f :\Longleftrightarrow A(e) \geq A(f)$.

For all $i \in I$ the set $S_i^t := \{e \in E | eP_i f \; \forall f \in E \setminus \{e\}\}$ represents individual i's top choice

- Plurality-order (see [Rob91]):
  The Plurality count of e is $Pl(e) := |\{i \in I : e \in S_i^t\}|$. The Plurality-order $\succsim_{pl}$ is defined by $e \succsim_{pl} f :\Longleftrightarrow Pl(e) \geq Pl(f)$.

- Borda-, Approval- and Plurality-order are *weak orders* on E (complete and transitive).
- Simple Majority-order is in general not transitive $\Rightarrow$ preference cycles

Some complete orders on τ

- Idea: Derive weak orders on τ from weak orders on E

Definition

For $T \in \tau$ we define the Borda count of T by $B(T) := \sum_{e \in T} B(e)$. Then the Borda-order $\trianglerighteq_B$ on τ is defined by letting, for all $T_1, T_2 \in \tau$,

$$T_1 \trianglerighteq_B T_2 :\Longleftrightarrow B(T_1) \geq B(T_2) \ .$$

- Analogously: Approval-order $\trianglerighteq_A$ on τ, Plurality-order $\trianglerighteq_{Pl}$ on τ.

General concept of a best tree w.r.t. orders, that are based on summing up numerical values of the edges:

Definition

Let τ be the set of spanning trees of G and let $\succsim$ be a weak order on E. A tree $M \in \tau$ is called *max-spanning tree* if and only if for every edge $f = \{i,j\}$, $f \notin M$,

$$f \precsim e$$

holds for all $e \in M$ that are part of the unique simple path between i and j in M.

Remarks:

1. Above definition generalizes the path optimality condition for the maximum spanning tree problem stated in [AMO93]
2. A max-spanning tree can be determined efficiently by a greedy algorithm (e.g. Kruskal's algorithm)
3. Note that for above definition $\succsim$ does not need to be based on numerical values
4. Simple-Majority order does not fit in this concept

Alternative idea to rank two trees:

1. Those edges that are simultaneously contained in both trees should not play a role. Thus, we simply remove those edges that both trees have in common.
2. Rank trees $T_1, T_2 \in \tau$ according to the sum of wins of edges of $\tilde{T}_1$ against those of $\tilde{T}_2$, where $\tilde{T}_1 := T_1 \setminus T_2$ and $\tilde{T}_2 := T_2 \setminus T_1$

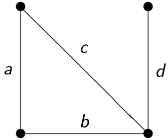
Definition

Let $\succsim$ be a complete order on E. For $e, f \in E$ let

$$s(e,f) := \begin{cases} 1 & \text{if } e \succ f \\ 0 & \text{if } e \sim f \\ -1 & \text{if } e \prec f \end{cases}$$

be the score of e versus f.
For $T_1,\ T_2 \in \tau$ we define

$$T_1 \trianglerighteq_S T_2 \ :\Longleftrightarrow \sum_{a \in \tilde{T}_1} \sum_{b \in \tilde{T}_2} s(a,b) \geq 0\ .$$



1	2	3
a	b	c
b	c	a
c	a	b
d	d	d

Example

3 spanning trees: $T_1 := \{a, b, d\}$, $T_2 := \{b, c, d\}$ and $T_3 := \{a, c, d\}$.

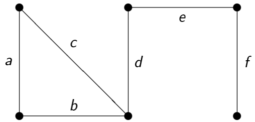
Preference cycle $a \succ_{sm} b \succ_{sm} c \succ_{sm} a$.

We get $T_1 \setminus T_2 = \{a\}$ and $T_2 \setminus T_1 = \{c\}$,

$T_1 \setminus T_3 = \{b\}$ and $T_3 \setminus T_1 = \{c\}$,

$T_2 \setminus T_3 = \{b\}$ and $T_3 \setminus T_2 = \{a\}$.

Thus we have $T_3 \rhd_S T_2 \rhd_S T_1 \rhd_S T_3$.
Hence a best tree with respect to $\unrhd_S$ does not exist in this example.



1	2	3
c	f	b
a	e	f
b	d	a
d	c	d
e	a	e
f	b	c

Example

For the above graph the Simple Majority-order on $E = \{a, b, c, d, e, f\}$ is of the following form:

$a \succ_{sm} b$	$b \succ_{sm} d$	$c \succ_{sm} a$	$d \succ_{sm} e$	$f \succ_{sm} c$
$a \succ_{sm} d$	$b \succ_{sm} e$	$c \succ_{sm} b$	$e \succ_{sm} c$	$f \succ_{sm} d$
$a \succ_{sm} e$	$b \succ_{sm} f$	$d \succ_{sm} c$	$f \succ_{sm} a$	$f \succ_{sm} e$

edge	# of inf. edges
f	4
a, b	3
c, d	2
e	1

$\Rightarrow$ Kruskal's algorithm according to SM-wins outputs $T_g = \{f, a, b, d, e\}$.

The two other spanning trees are $T_1 = \{b, c, d, e, f\}$ and $T_2 = \{a, c, d, e, f\}$.
$T_1 \setminus T_g = \{c\}$ and $T_g \setminus T_1 = \{a\} \Rightarrow T_1 \rhd_S T_g$.

$T_2 \setminus T_g = \{c\}$ and $T_g \setminus T_2 = \{b\} \Rightarrow T_2 \rhd_S T_g$.
I.e. according to $\trianglerighteq_S$ every other spanning tree of the graph is strictly preferred to T_g.

Proposition

Let $\succsim = \succsim_{sm}$. Then the following statements hold:

1. There exist a graph $G = (V, E)$ and a voter preference profile on E such that a best tree with respect to $\trianglerighteq_S$ does not exist.
2. There exist a graph $G = (V, E)$ and a voter preference profile on E such that a best tree with respect to $\trianglerighteq_S$ exists but the generalized version of Kruskal's algorithm fails to determine such a best tree. In fact, the tree determined by the algorithm may even be the worst tree with respect to $\trianglerighteq_S$.

Comparing trees

Three more complete orders on τ

Idea:

- Compare trees on basis of a *given weak order* $\succsim$ on E.
- When comparing two trees, those edges that are simultaneously contained in both trees should not play a role.

Notation:
Given $T_1, T_2 \in \tau$, we use the notation $\tilde{T}_1 := T_1 \setminus T_2$, $\tilde{T}_2 := T_2 \setminus T_1$ and $r := |\tilde{T}_1|$

Three more complete orders on τ

Concept 1: Derived from the maxmin-order on sets presented in [BBP04]

Definition

Let $T_1, T_2 \in \tau$. Then we define the maxmin-order $\trianglerighteq_{mxn}$ on τ by

$$T_1 \trianglerighteq_{mxn} T_2 \;:\Longleftrightarrow\; \begin{array}{l} [\tilde{T}_1 = \emptyset \text{ or} \\ \max \tilde{T}_1 \succ \max \tilde{T}_2 \text{ or} \\ (\max \tilde{T}_1 \sim \max \tilde{T}_2 \text{ and } \min \tilde{T}_1 \succsim \min \tilde{T}_2)] \end{array}$$

Concept 2: Derived from the leximax order on sets presented in [BBP04]

Definition

Let $T_1, T_2 \in \tau$.
Let $\tilde{T}_1 := \{e_1, e_2, ..., e_r\}$, $\tilde{T}_2 := \{f_1, f_2, ..., f_r\}$ such that $e_i \succsim e_{i+1}$ and $f_i \succsim f_{i+1}$ holds for $1 \leq i \leq r-1$.
Then the leximax order $\trianglerighteq_{lex}$ on τ is defined by

$$T_1 \trianglerighteq_{lex} T_2 \;:\Longleftrightarrow\; \begin{array}{l} [\tilde{T}_1 = \emptyset \text{ or} \\ e_i \sim f_i \text{ for all } 1 \leq i \leq r \text{ or} \\ (\exists j \in \{1, ..., r\} \text{ such that} \\ e_i \sim f_i \text{ for all } i < j \text{ and } e_j \succ f_j)] \end{array}$$

Concept 3: Rank the edges of the disjoint union of $T_1, T_2 \in \tau$ according to $\succsim$. For the resulting ranking use a positional scoring concept to compare the trees.

Definition

Let $T_1, T_2 \in \tau$.
Let $\tilde{T}_1 \cup \tilde{T}_2 := \{d_1, d_2, ..., d_{2r}\}$ such that $d_i \succsim d_{i+1}$ holds for $1 \leq i \leq 2r - 1$.
Let $b : E \to \mathbb{R}$ be strictly increasing according to $\succsim$, that is, for $1 \leq i \leq 2r - 1$,

$$\begin{array}{ll} b(d_i) = b(d_{i+1}) & \text{if } d_i \sim d_{i+1} \\ b(d_i) > b(d_{i+1}) & \text{if } d_i \succ d_{i+1} \end{array}$$

Let $b(\tilde{T}_1) := \sum_{e \in \tilde{T}_1} b(e)$ and $b(\tilde{T}_2) := \sum_{f \in \tilde{T}_2} b(f)$.
Then we define

$$T_1 \trianglerighteq_{ps} T_2 \ :\Longleftrightarrow b(\tilde{T}_1) \geq b(\tilde{T}_2) \ .$$

Remark. This approach adapts the concept of the positional scoring procedures presented in [BF02]

Remarks.

- The orders $\trianglerighteq_S$, $\trianglerighteq_{lex}$, $\trianglerighteq_{mxn}$ and $\trianglerighteq_{ps}$ are complete orders on τ.
- In the above concepts the order $\succsim$ on E does not need to be of numerical nature, i.e. $\succsim$ does not have to allocate numbers to the edges.

Aim: Find a best tree w.r.t. the corresponding order

Results on max-spanning trees and best trees

Recall: $\succsim$ is assumed to be a given weak order on E,
i.e. $\succsim$ is complete and transitive

Theorem

A max-spanning tree can be computed in $\mathcal{O}(|E| + |V| \log |V|)$ time.

Proof.

Immediately follows from the fact that the maximum spanning tree problem can be solved in $\mathcal{O}(|E| + |V| \log |V|)$ time. □

Main Theorem

Theorem

Let $M \in \tau$ and let $\succsim$ be a weak order on E. Then the following statements are equivalent:

1. *M is a max-spanning tree*
2. $\nexists B \in \tau : B \rhd_{lex} M$
3. $\nexists B \in \tau : B \rhd_{S} M$
4. $\nexists B \in \tau : B \rhd_{mxn} M$
5. $\nexists B \in \tau : B \rhd_{ps} M$

Corollary

Every positional scoring method that yields the same ranking $\succsim$ on E yields the same set of best trees w.r.t. $\trianglerighteq_{ps}$, irrespective of the numerical values assigned to the edges.

Consequences

Let $\succsim$ be a weak order on E and let $\trianglerighteq \in \{\trianglerighteq_{lex}, \trianglerighteq_{S}, \trianglerighteq_{mxn}, \trianglerighteq_{ps}\}$

Consequences of the Theorem:

1. A best tree with respect to $\trianglerighteq$ always exists
2. A best tree with respect to $\trianglerighteq$ can be determined efficiently
3. For the orders $\trianglerighteq_{lex}, \trianglerighteq_{S}, \trianglerighteq_{mxn}, \trianglerighteq_{ps}$, the sets of best trees coincide

Concluding Remark: Results can be generalized to bases of matroids

References

[AMO93] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network flows: theory, algorithms, and applications*. Prentice Hall, 1993.

[BBP04] S. Barbera, W. Bossert, and P.K. Pattanaik. Ranking sets of objects. In *Handbook of Utility Theory*, volume 2, pages 893–977. 2004.

[BE93] M.W. Bern and D. Eppstein. Worst-case bounds for subadditive geometric graphs. In *Symposium on Computational Geometry*, pages 183–188, 1993.

[BF83] S.J. Brams and P.C. Fishburn. *Approval Voting*. Birkhäuser, Boston, 1983.

[BF02] S.J. Brams and P.C. Fishburn. Voting procedures. In K.J. Arrow, A.K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 1, pages 173–236. Elsevier, 2002.

[Bla58] D. Black. *The Theory of Committees and Elections*. Cambridge University Press, Cambridge, 1958.

[BTT89] J.J. Bartholdi, C.A. Tovey, and M.A. Trick. Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6:157–165, 1989.

[Cop51] A.H. Copeland. *A 'Reasonable' Social Welfare Function*. Notes from a Seminar on Applications of Mathematics to the Social Sciences, University of Michigan, 1951.

[Nur07] H. Nurmi. *Voting Procedures under Uncertainty*. Springer Verlag, Berlin, 2007.

[PS05] P. Perny and O. Spanjaard. A preference-based approach to spanning trees and shortest paths problems. *European Journal of Operational Research*, 127(3):584–601, May 2005.

[Rob91] F.S. Roberts. Characterizations of the plurality function. *Mathematical Social Sciences*, 21(2):101–127, April 1991.

[Saa95] D.G. Saari. *Basic Geometry of Voting*. Springer Verlag, Berlin, 1995.

[Sla61] P. Slater. Inconsistencies in a schedule of paired comparisons. *Biometrica*, 48:303–312, 1961.

[Wel76] D.J. Welsh. *Matroid theory*. Academic Press, London, New York, 1976.