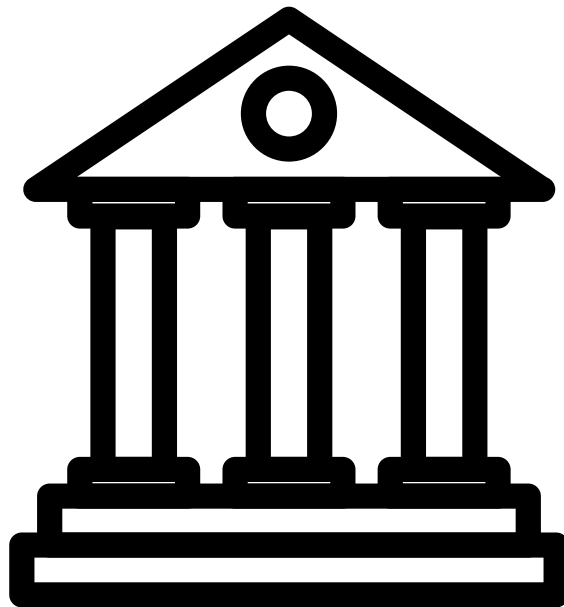


Background

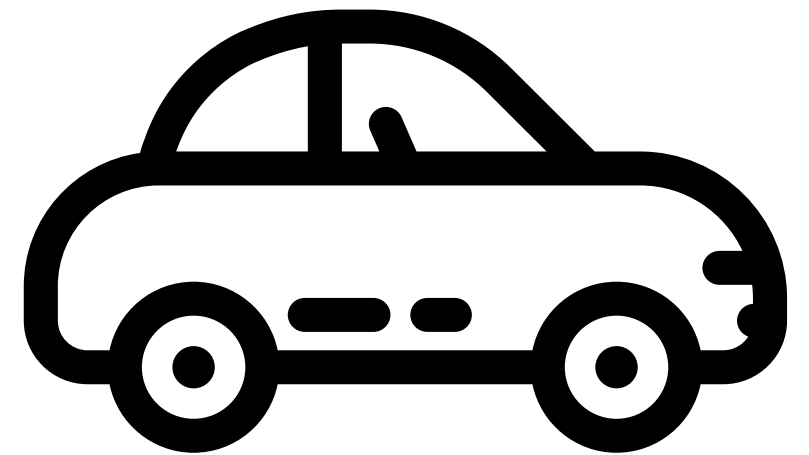
- Multi-criteria decision making



Employee recruitment



University selection



Car comparison

Background

- Suppose you are a college basketball coach, how do you recruit the best players?



PLAYER	TEAM	AGE	GP	W	L	MIN	OFFRTG	DEFRTG	NETRTG	AST%	AST/TO	AST RATIO	OREB%
AJ Hammons	DAL	24	22	4	18	7.4	102.2	102.8	-0.6	3.8	0.40	6.2	4.9
Aaron Brooks	IND	32	65	36	29	13.7	101.5	104.6	-3.0	21.6	1.89	24.6	2.2
Aaron Gordon	ORL	21	80	29	51	28.7	105.4	108.2	-2.8	9.7	1.69	12.5	5.4
Aaron Harrison	CHA	22	5	2	3	3.3	83.3	101.9	-18.6	37.5	0.00	38.1	0.0
Adreian Payne	MIN	26	18	5	13	7.5	102.6	101.8	0.8	8.9	0.88	9.0	6.9
Al Horford	BOS	31	68	46	22	32.3	110.7	105.8	5.0	23.9	2.93	25.7	4.9
Al Jefferson	IND	32	66	33	33	14.1	102.3	108.1	-5.8	11.4	1.73	9.5	8.6
Al-Farouq Aminu	POR	26	61	33	28	29.1	107.7	105.9	1.8	8.2	1.05	13.8	4.9
Alan Anderson	LAC	34	30	20	10	10.3	103.1	114.0	-10.8	5.2	1.57	10.5	1.1
Alan Williams	PHX	24	47	11	36	15.1	105.6	105.8	-0.3	4.9	0.62	6.1	13.8
Alec Burks	UTA	25	42	26	16	15.5	105.0	104.9	0.1	7.4	0.86	8.6	2.9
Alex Abrines	OKC	23	68	37	31	15.5	106.0	108.3	-2.3	5.5	1.21	9.2	1.9
Alex Len	PHX	24	77	21	56	20.3	99.4	110.5	-11.1	4.3	0.43	6.3	10.4

Introduction – Skyline

- **Skyline algorithm**: automatically select the **skyline** of the dataset
- In database, skyline algorithm is an important and extensively studied problem

Introduction – Skyline

- **Skyline algorithm:** automatically select the **skyline** of the dataset



Introduction – Skyline Definition

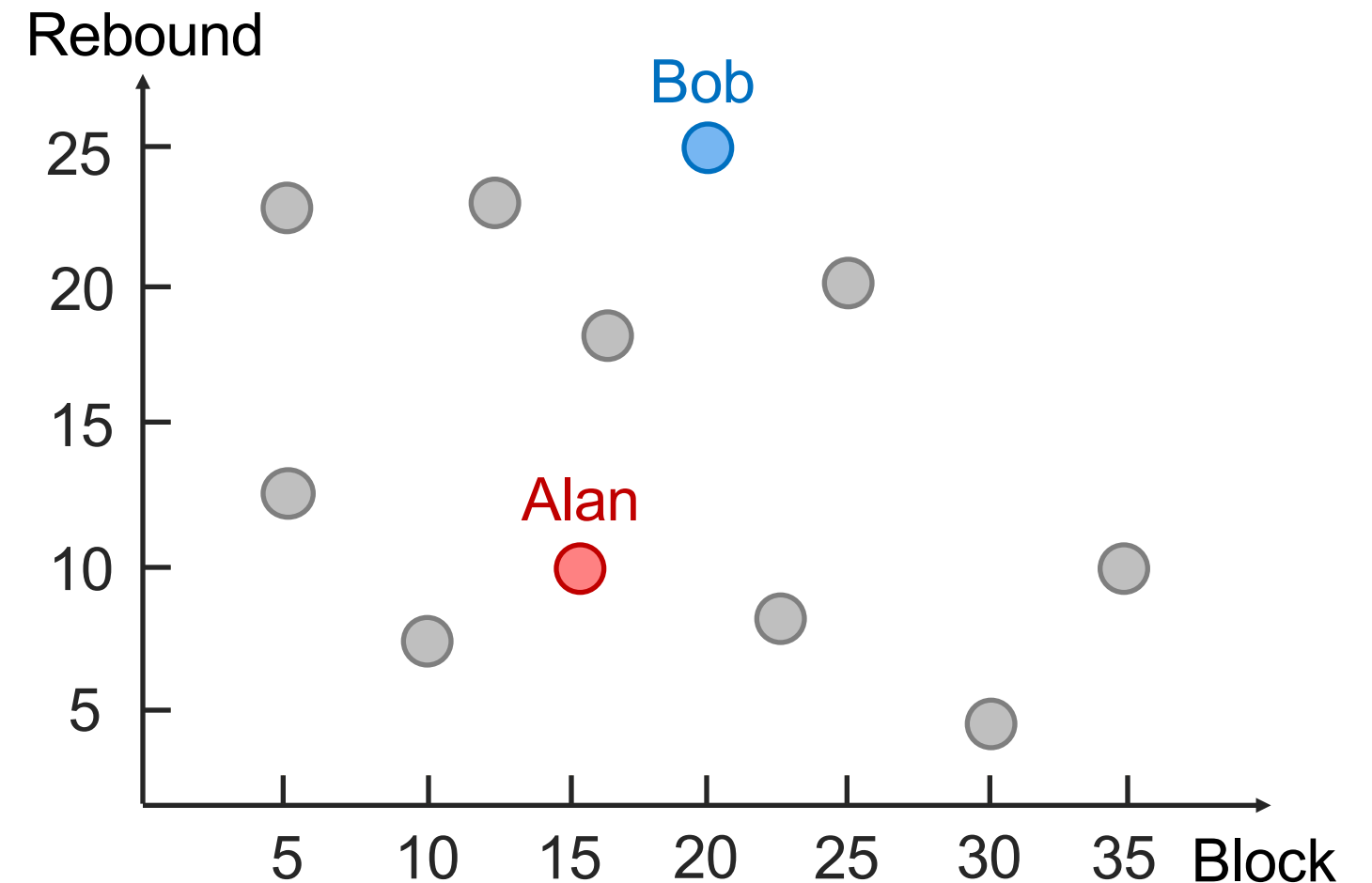
- **Skyline:** a set of **superior** points that are not **dominated** by other points in the dataset
- **Dominance:**
 - If p **dominates** q , then:
 - p is not worse than q in all attributes
 - p is at least better than q in one attribute

Introduction – Skyline Example

- **Skyline**: a set of **superior** points that are not **dominated** by other points in the dataset

Players	Block	Rebound
Alan	15	10
Bob	20	25

Bob dominates Alan (**block** & **rebound**)

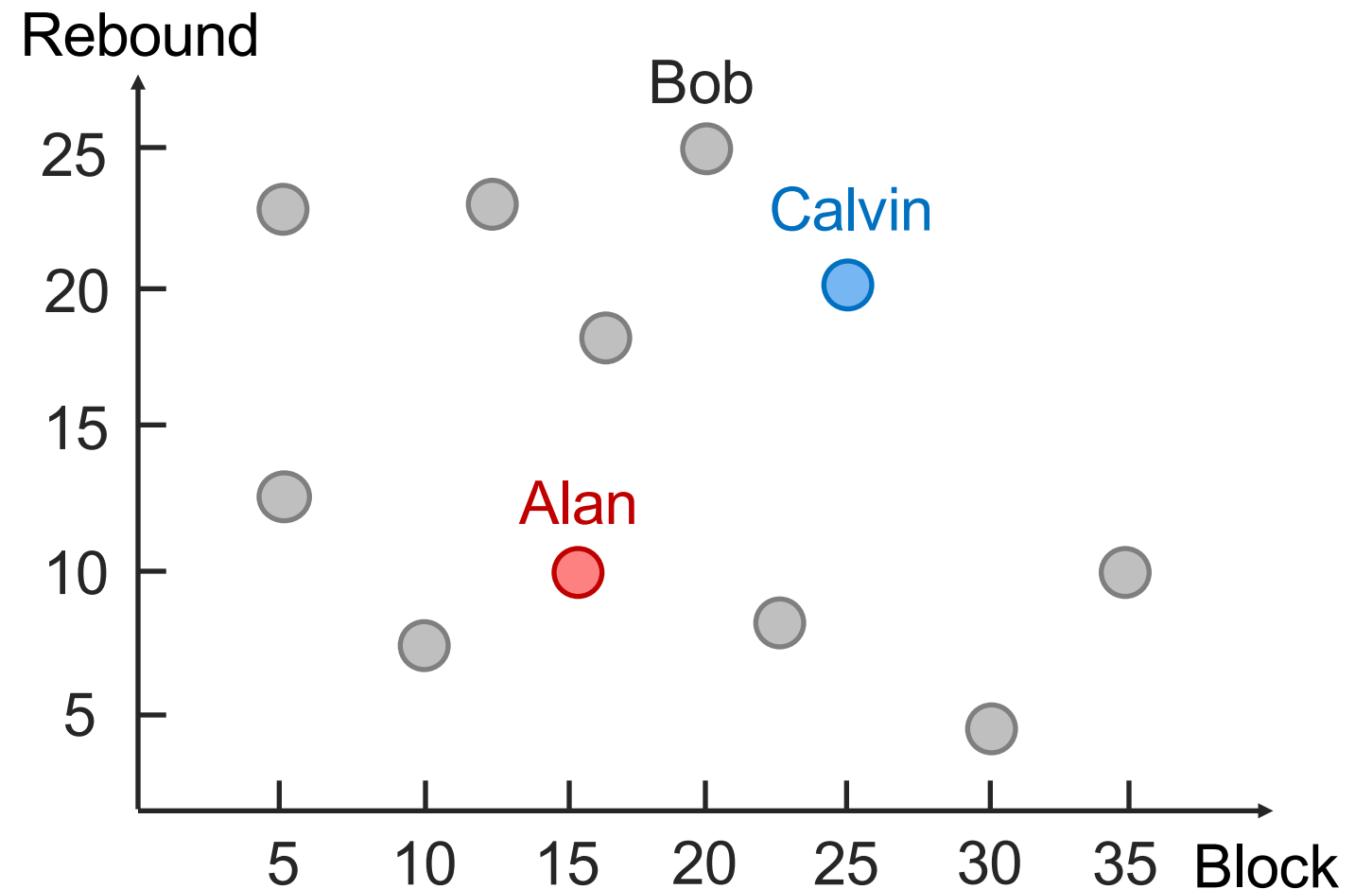


Introduction – Skyline Example

- **Skyline**: a set of **superior** points that are not **dominated** by other points in the dataset

Players	Block	Rebound
Alan	15	10
Bob	20	25
Calvin	25	20

Calvin dominates Alan (**block** & **rebound**)

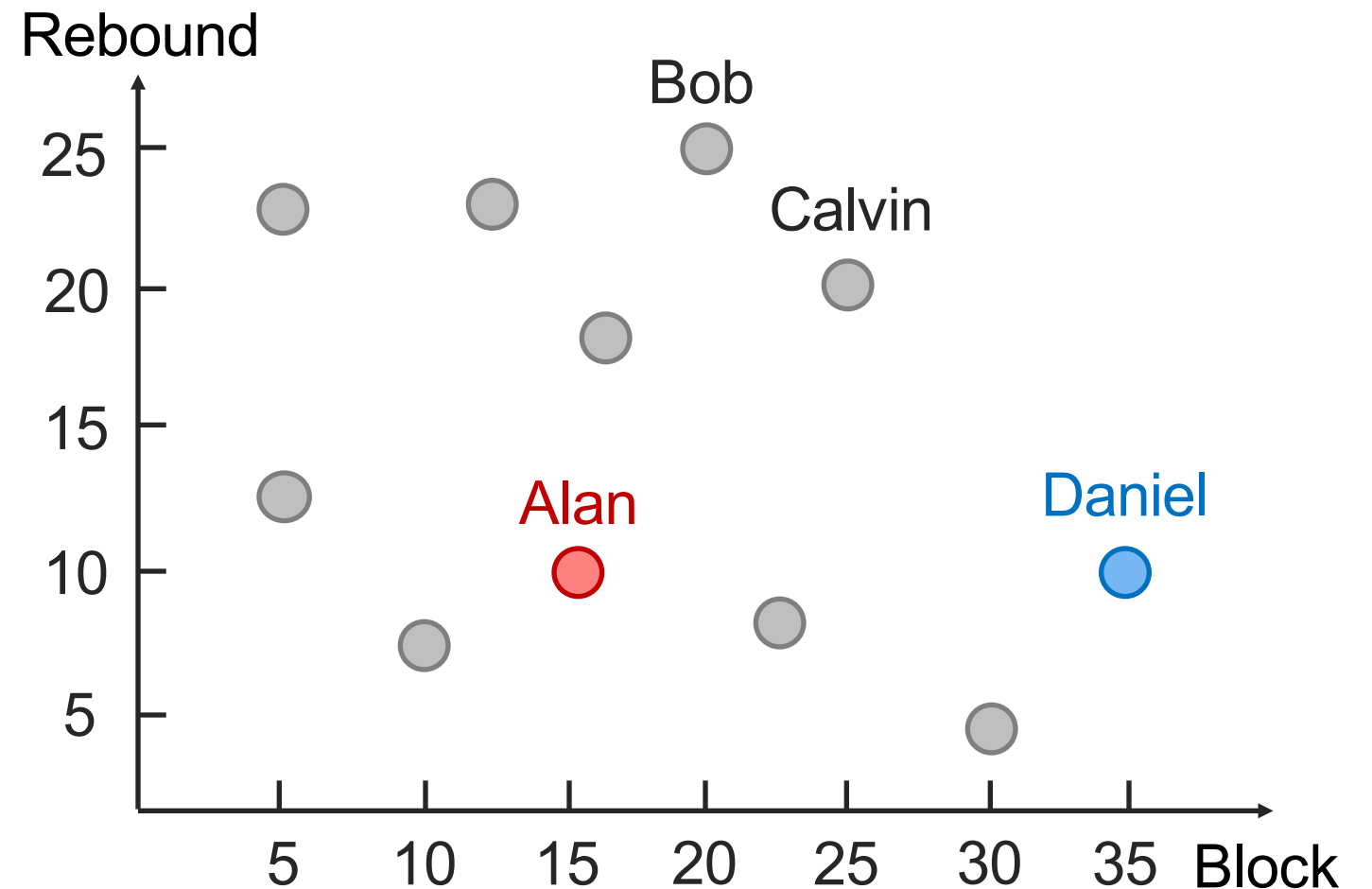


Introduction – Skyline Example

- **Skyline**: a set of **superior** points that are not **dominated** by other points in the dataset

Players	Block	Rebound
Alan	15	10
Bob	20	25
Calvin	25	20
Daniel	30	10

Daniel dominates Alan (**block**)

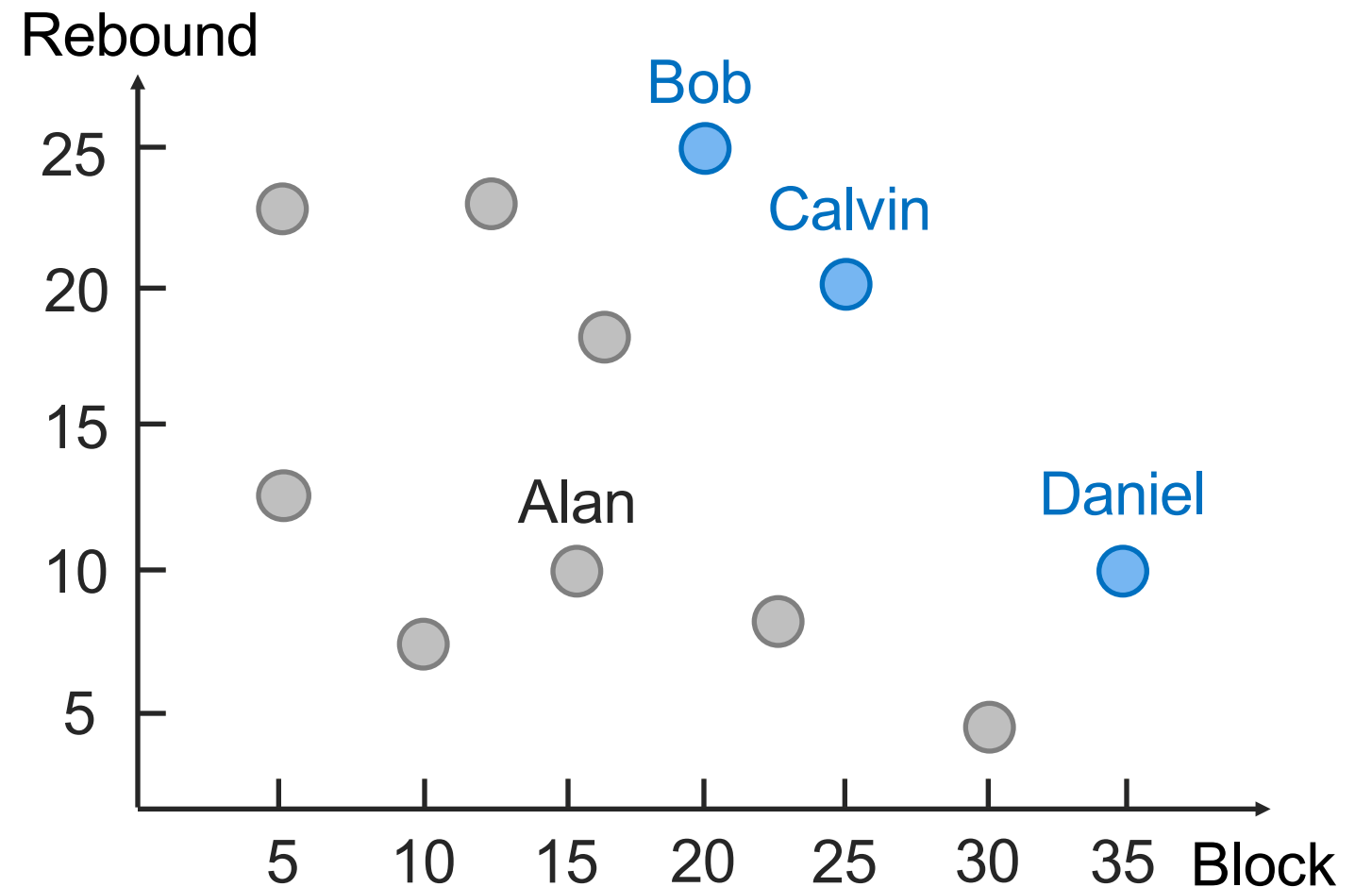


Introduction – Skyline Example

- **Skyline**: a set of **superior** points that are not **dominated** by other points in the dataset

Players	Block	Rebound
Alan	15	10
Bob	20	25
Calvin	25	20
Daniel	30	10

Points: Daniel > Calvin > Bob
Rebound: Bob > Calvin > Daniel

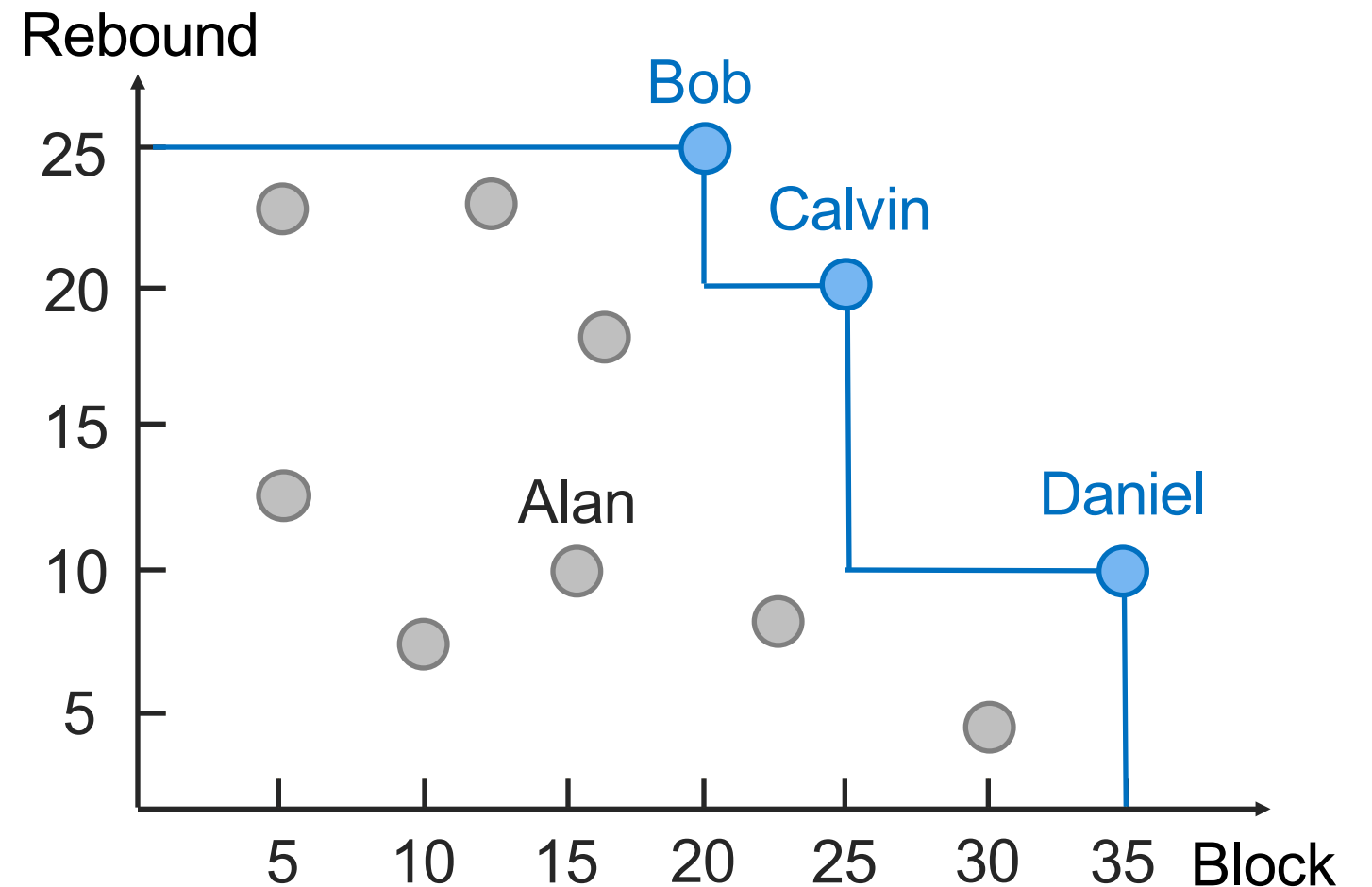


Introduction – Skyline Example

- **Skyline**: a set of **superior** points that are not **dominated** by other points in the dataset

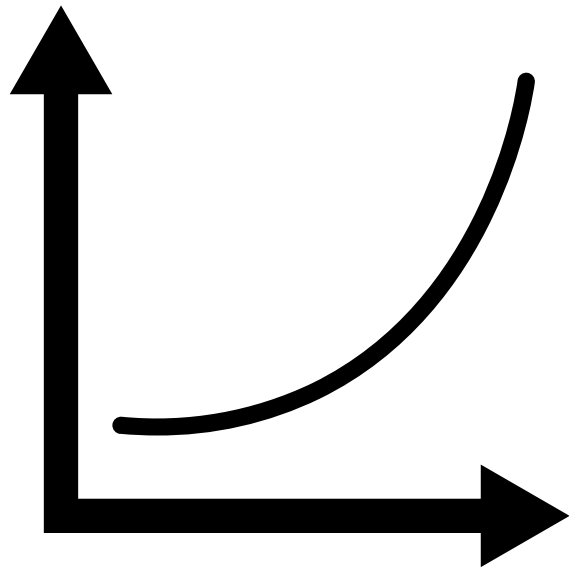
Players	Block	Rebound
Alan	15	10
Bob	20	25
Calvin	25	20
Daniel	30	10

Skyline: Bob, Calvin, Daniel



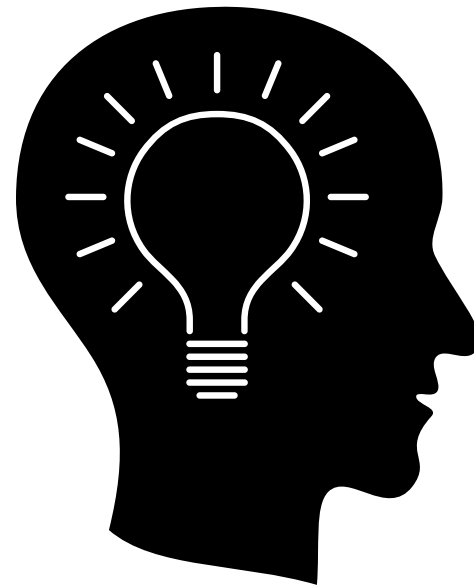
Introduction – Challenges

Scalability



The **size of skyline** increases with the number of attributes

Interpretation



The **reasons** that make a point in skyline is unclear

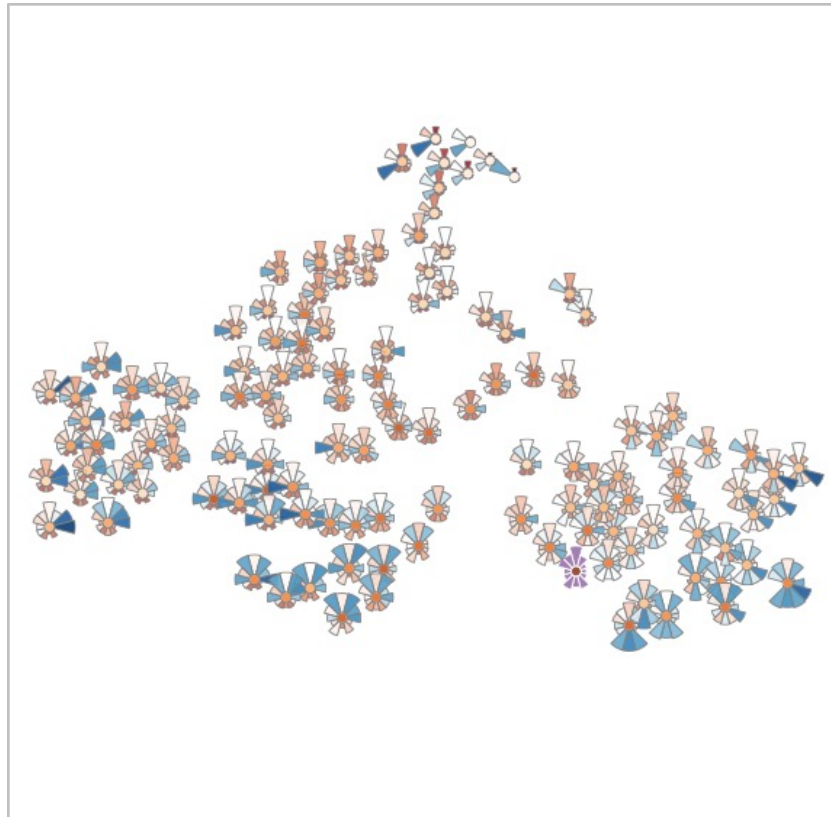
Comparison



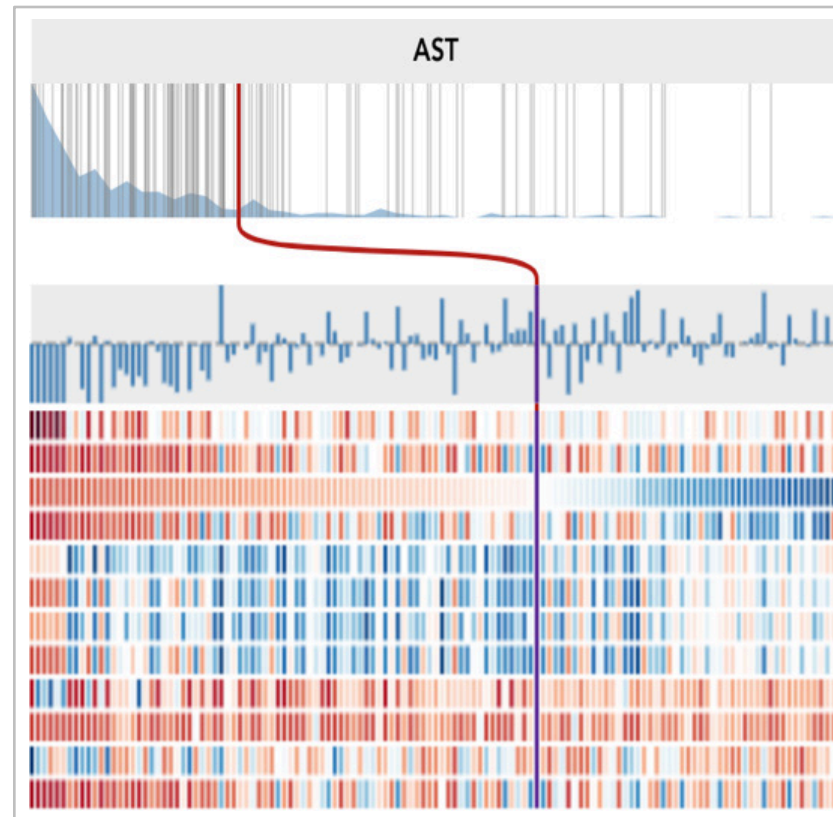
The **strength and weakness** of each skyline point is implicit

SkyLens – Visual Components

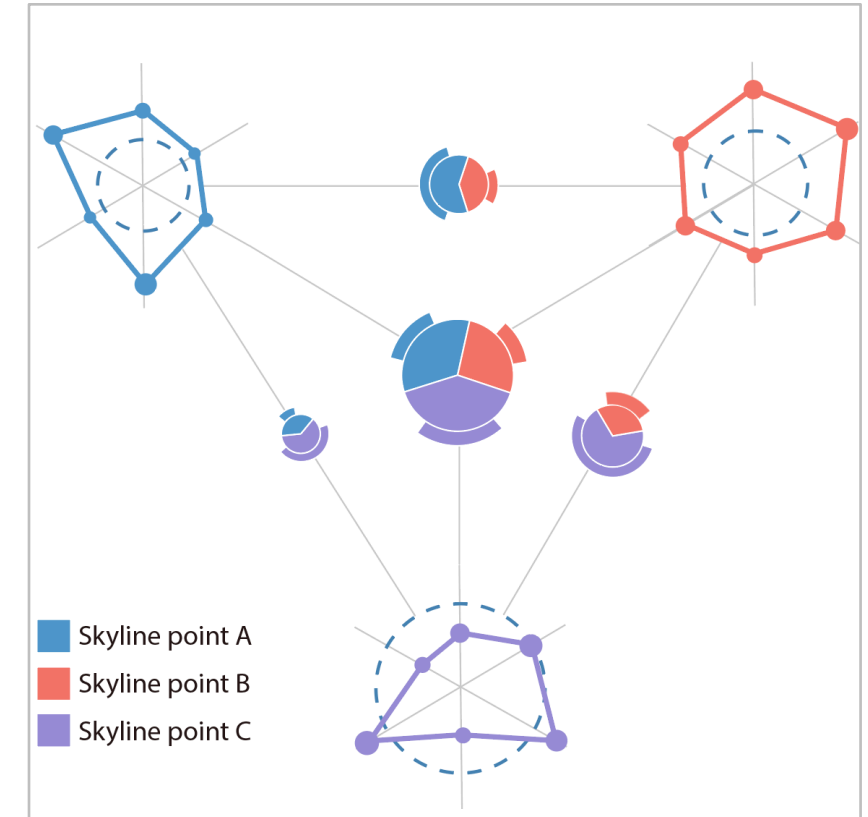
Projection View



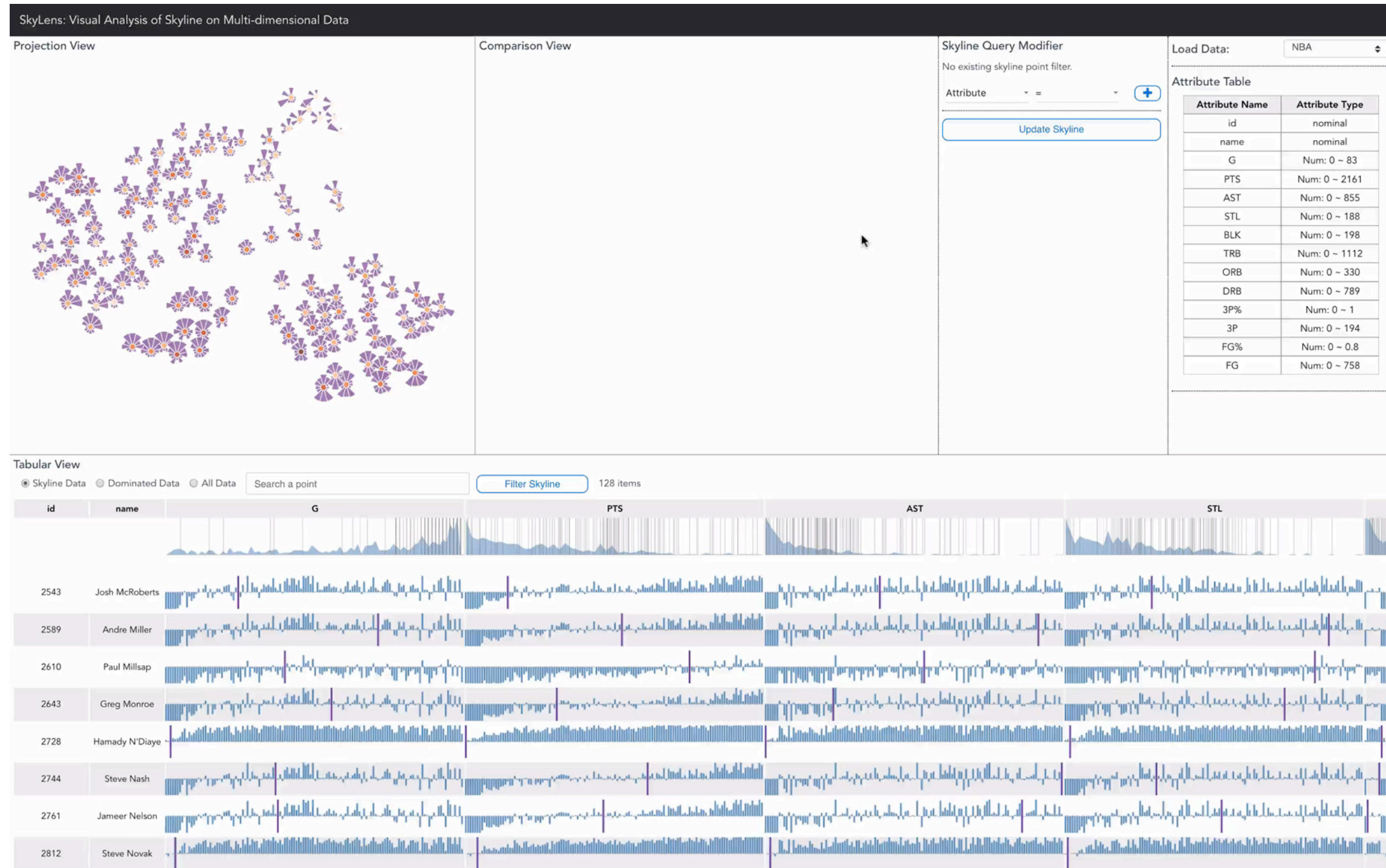
Tabular View



Comparison View

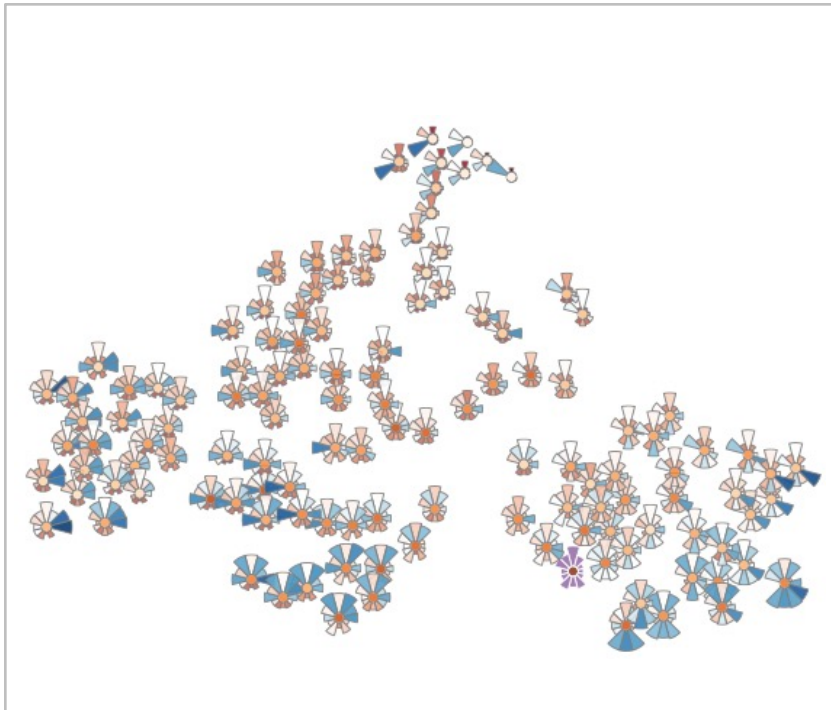


SkyLens – Video

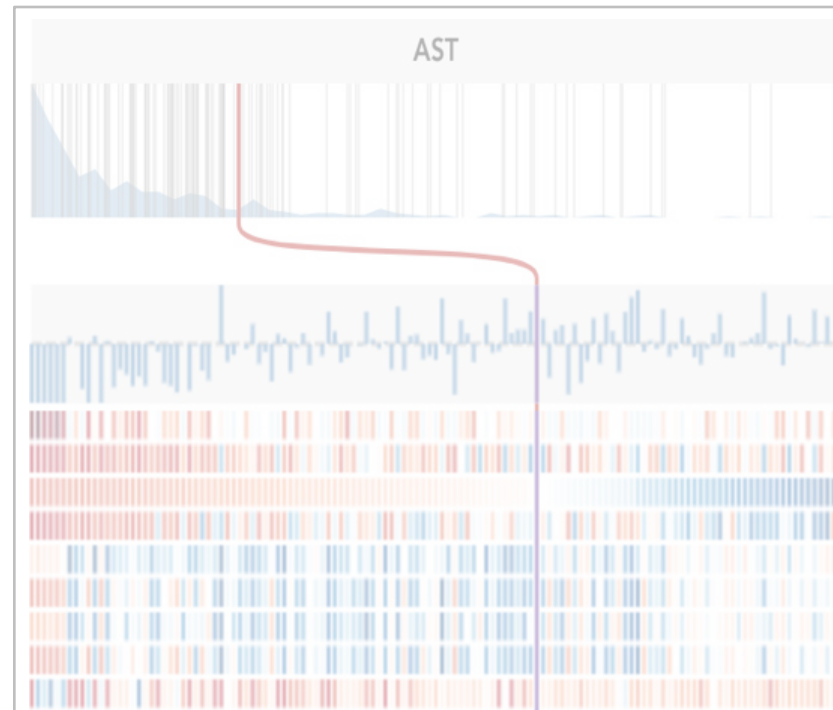


SkyLens – Projection View

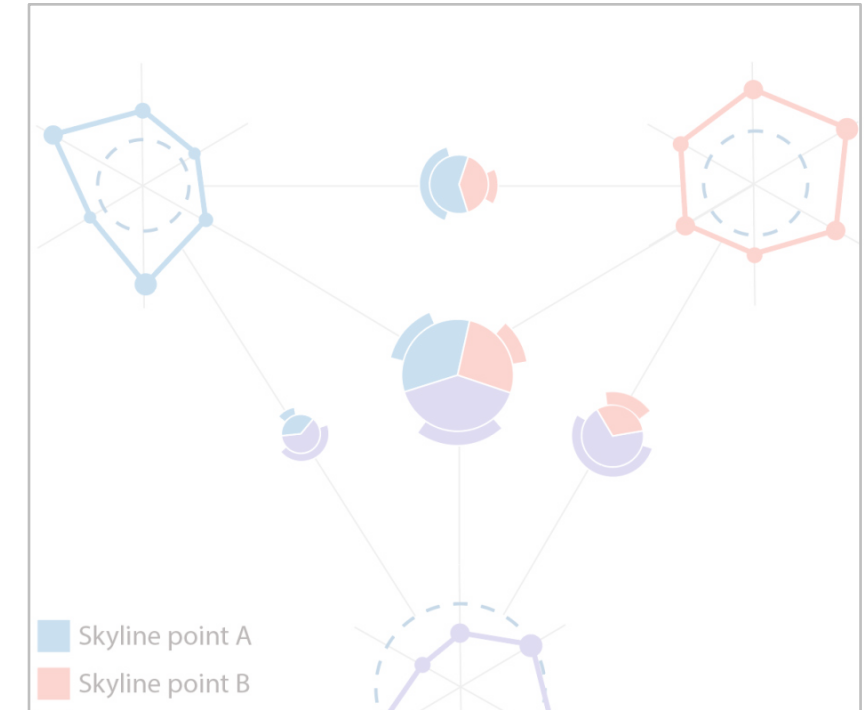
Projection View



Tabular View



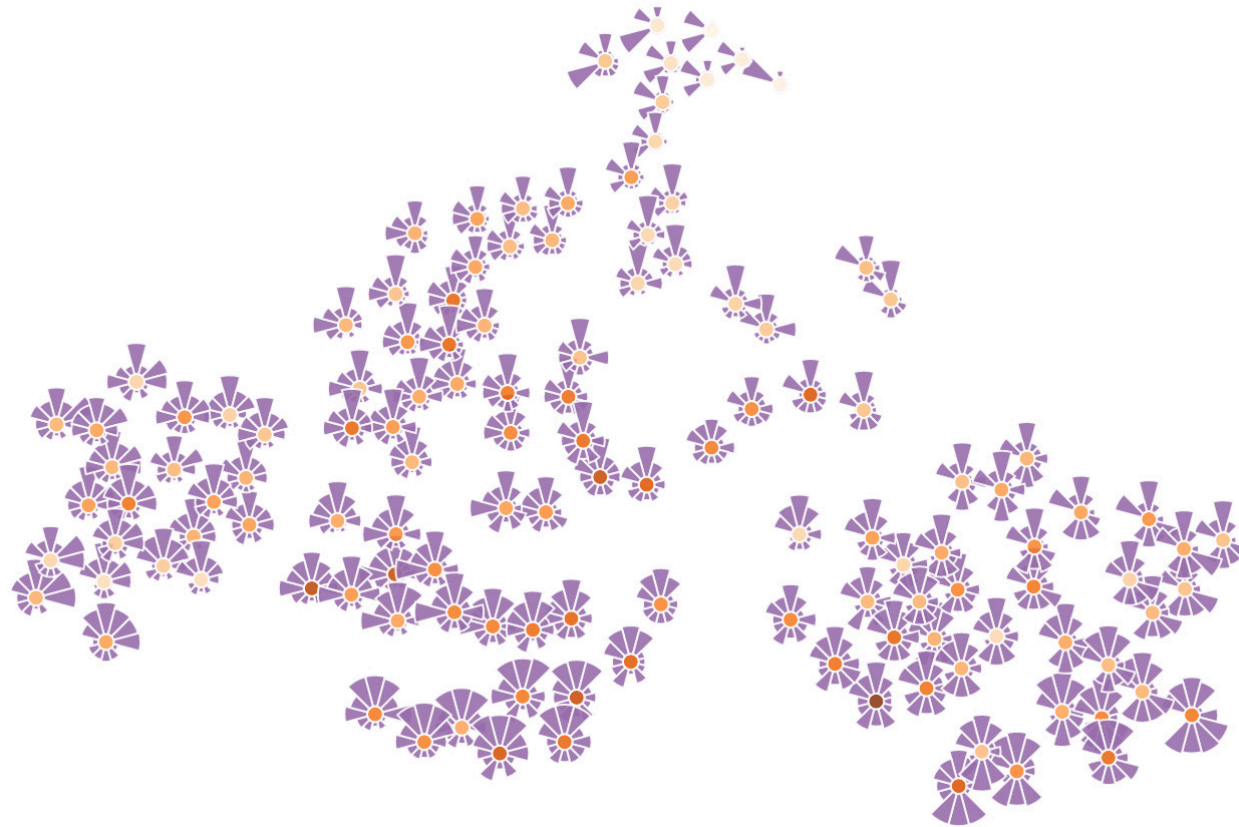
Comparison View



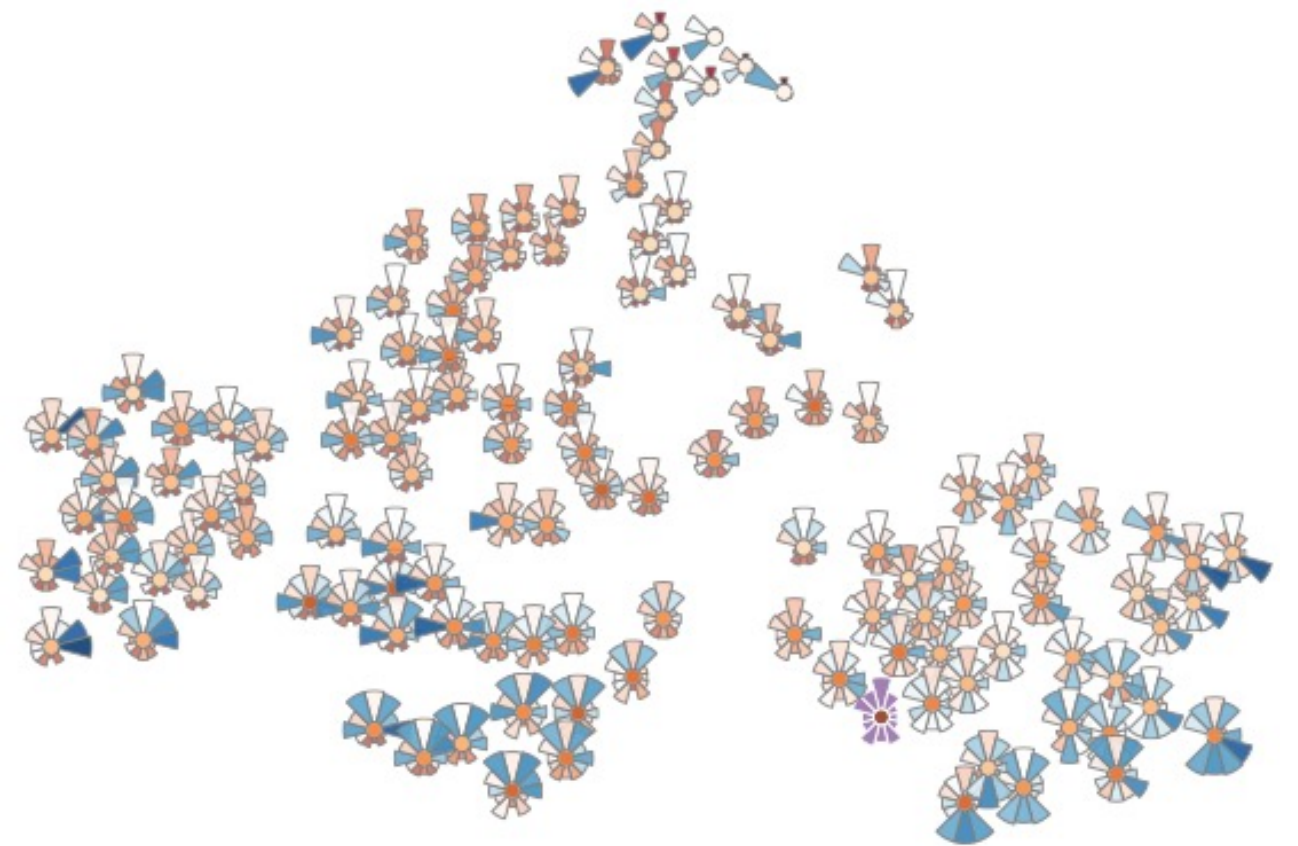
Projection View: provide an overview of skyline (clusters and outliers)

Projection View

- Methods: t-SNE projection and skyline glyphs



Normal mode

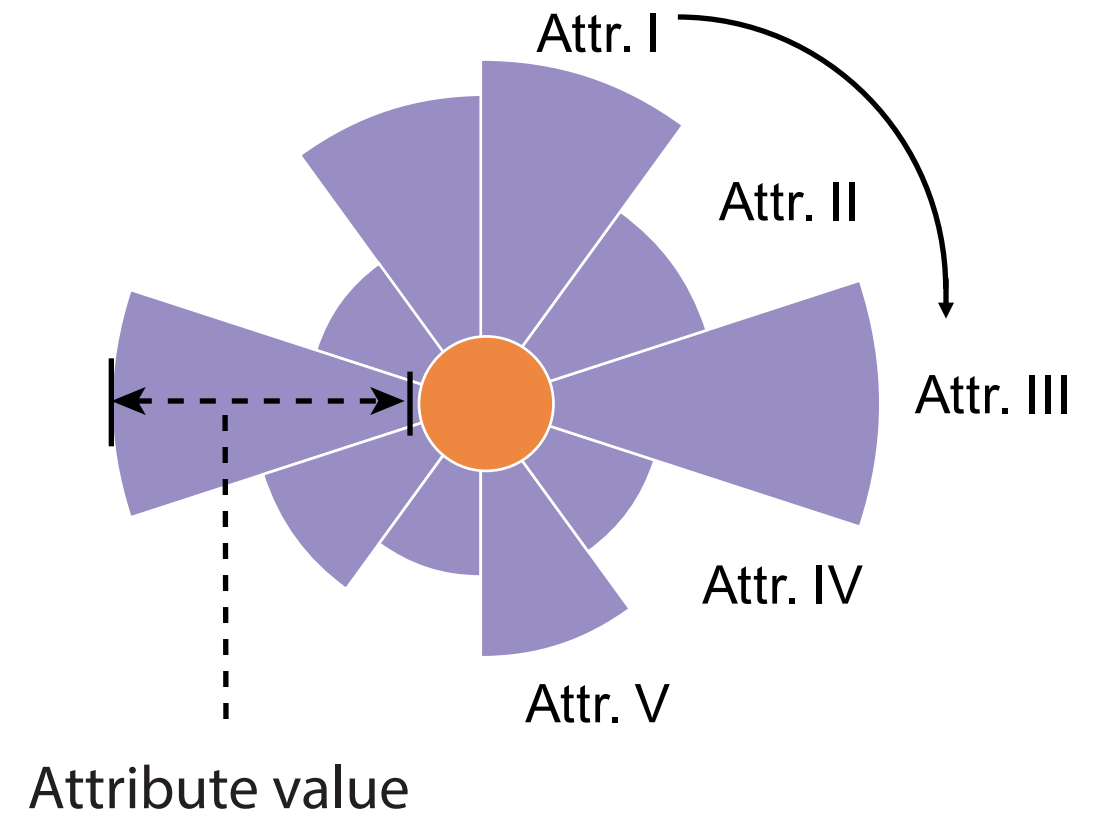


Focus mode

Projection View – Skyline Glyph

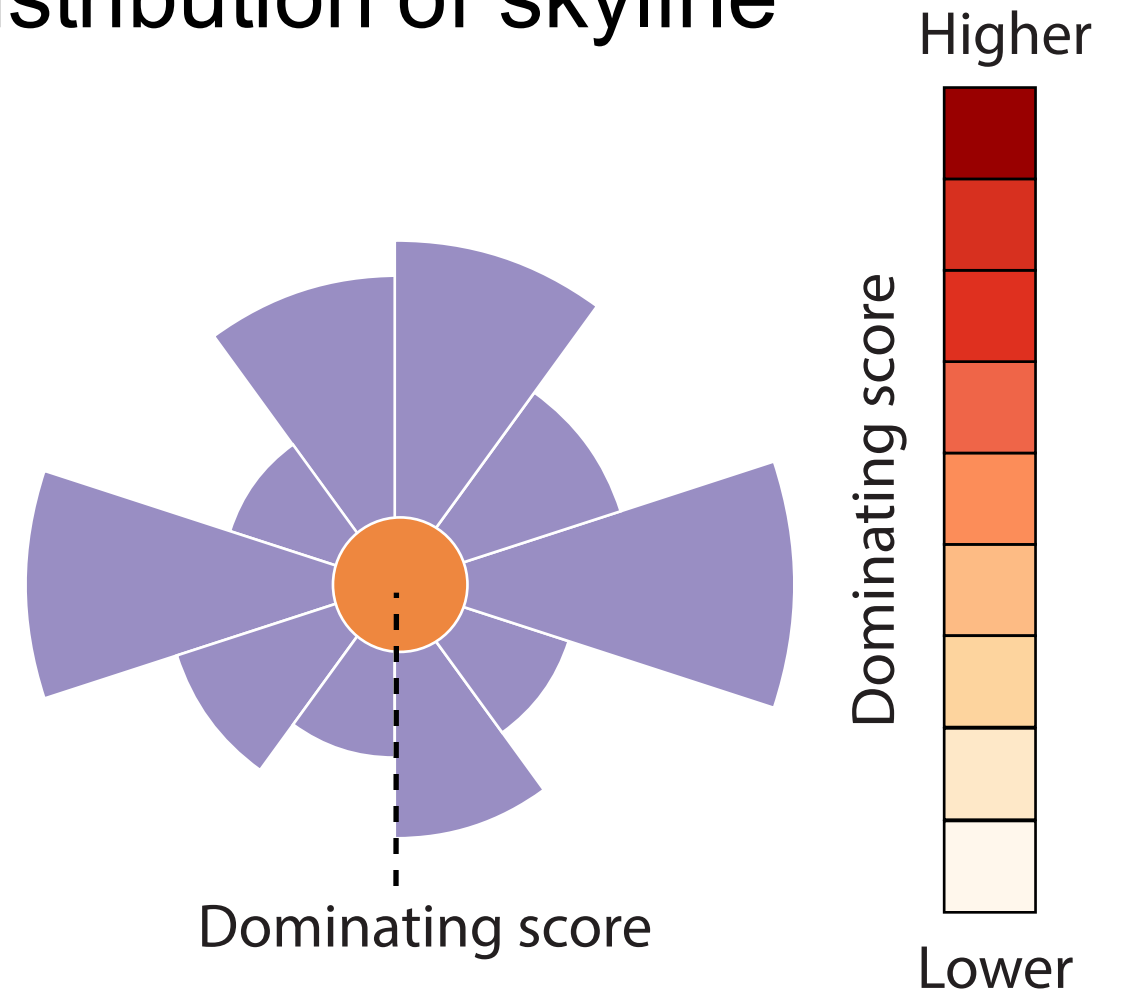
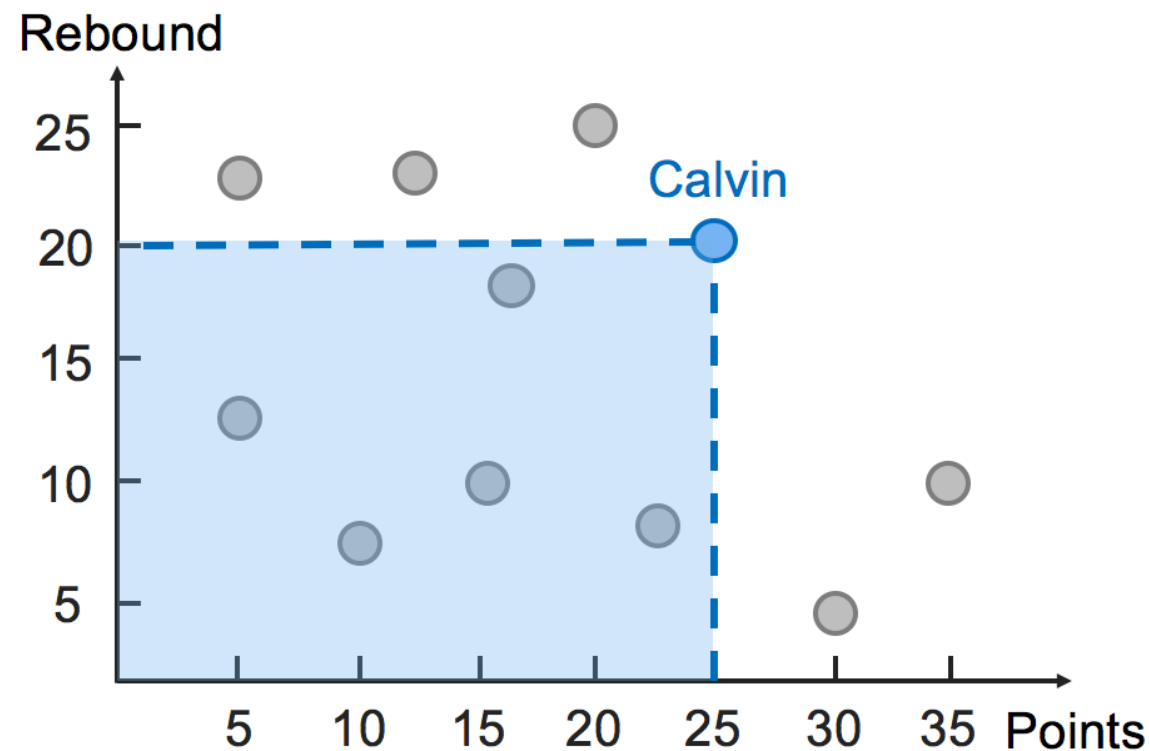
- **Normal mode:** show the attribute value distribution of skyline

Attribute	Value
Attr. I	5
Attr. II	3
Attr. III	7
Attr. IV	1
Attr. V	3
Attr. VI	1
...	



Projection View – Skyline Glyph

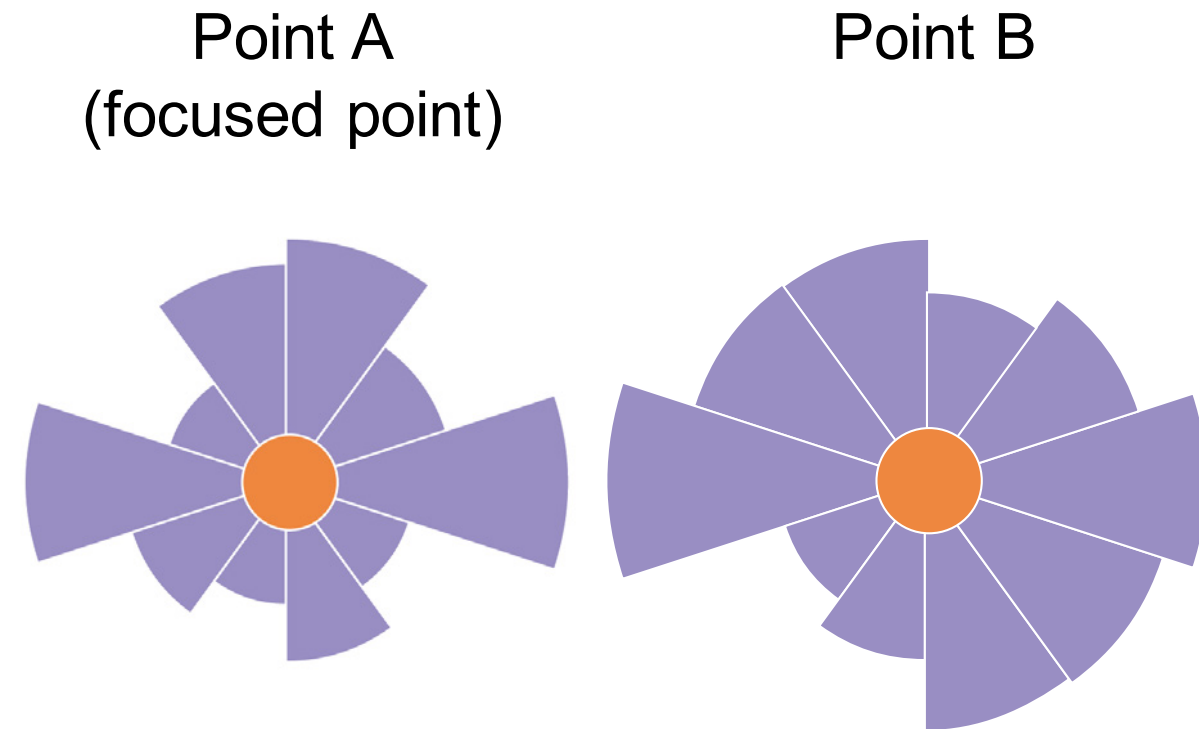
- Normal mode: show the attribute value distribution of skyline
- Dominating score (superiority metric):
 - # of points dominated by this point



Projection View – Skyline Glyph

- **Focus mode:** highlight how other points differ from a focused one

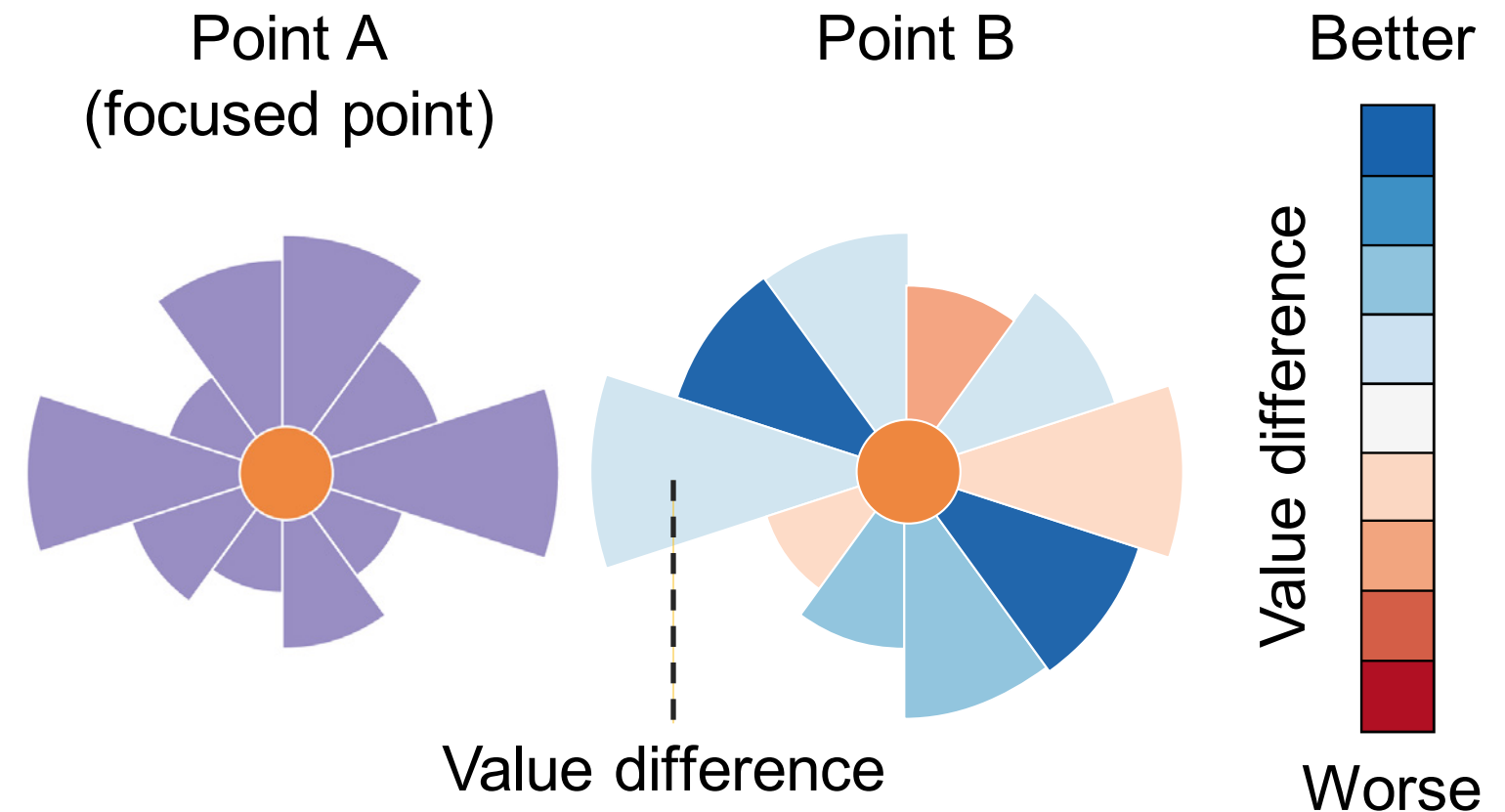
Attribute	Point A	Point B
Attr. I	5	3
Attr. II	3	4
Attr. III	7	6
Attr. IV	1	5
Attr. V	3	5
Attr. VI	1	3
...		



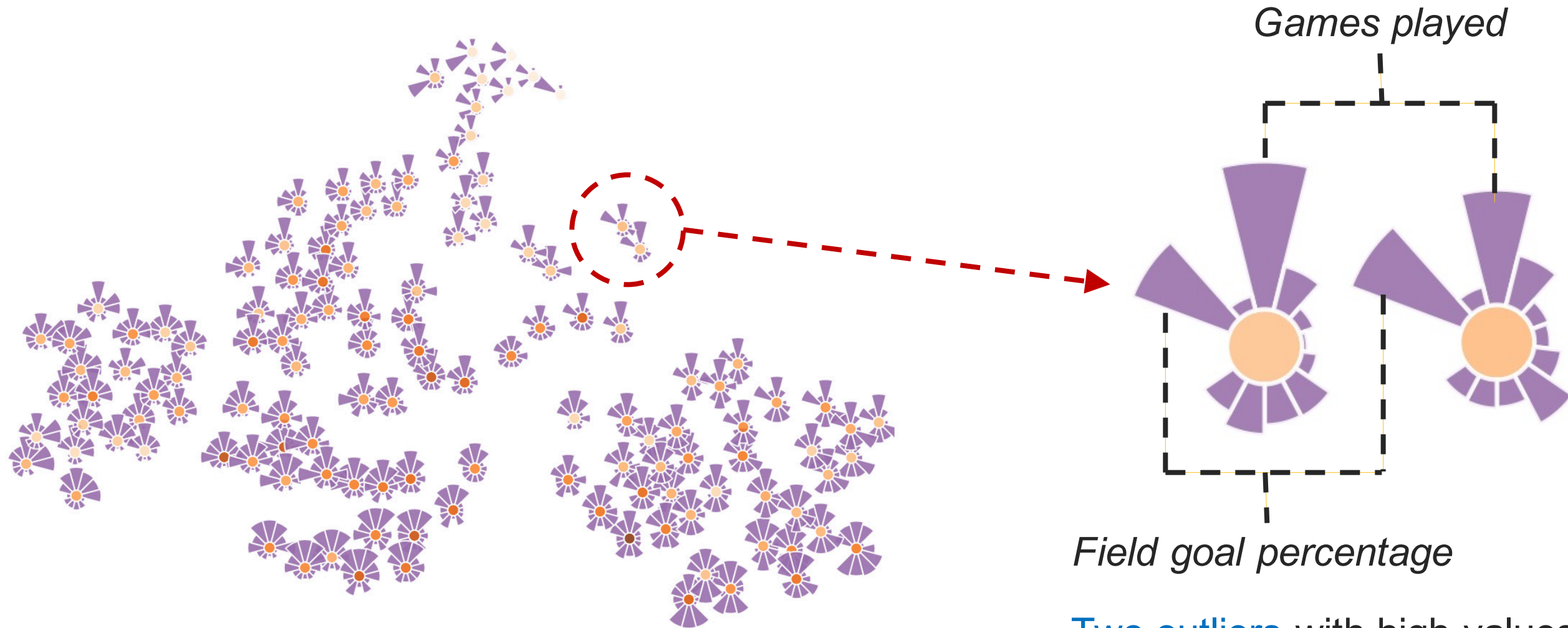
Projection View – Skyline Glyph

- **Focus mode:** highlight how other points differ from a focused one using color map

Attribute	Point A	Point B
Attr. I	5	3 (diff. = -2)
Attr. II	3	4 (diff. = 1)
Attr. III	7	6 (diff. = -1)
Attr. IV	1	5 (diff. = 4)
Attr. V	3	5 (diff. = 2)
Attr. VI	1	3 (diff. = 2)
	...	

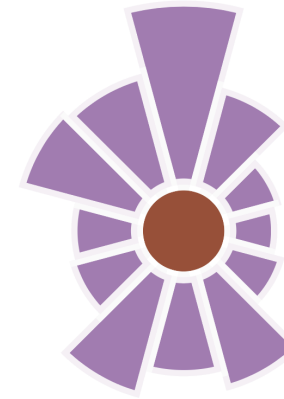
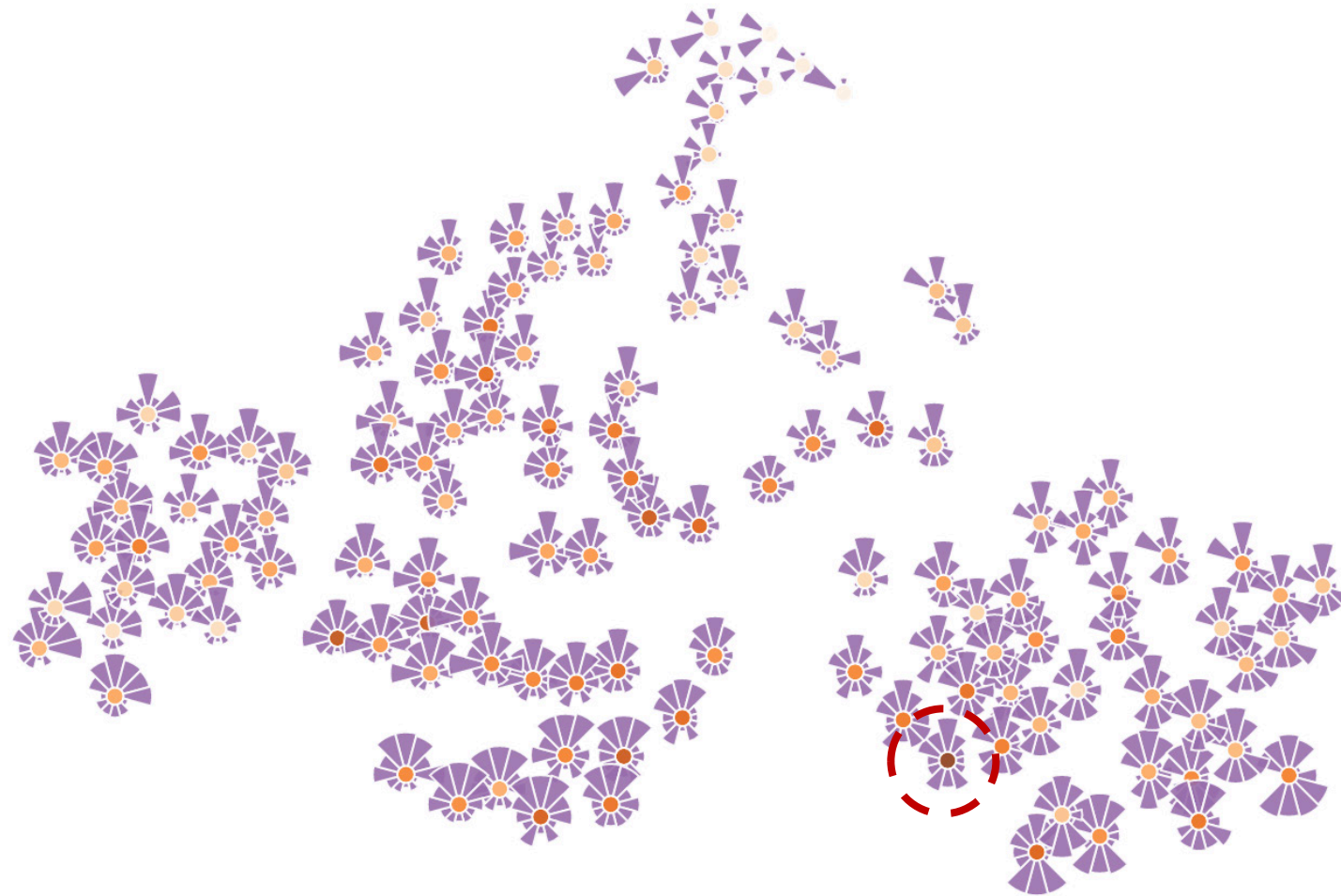


Projection View – Case Studies



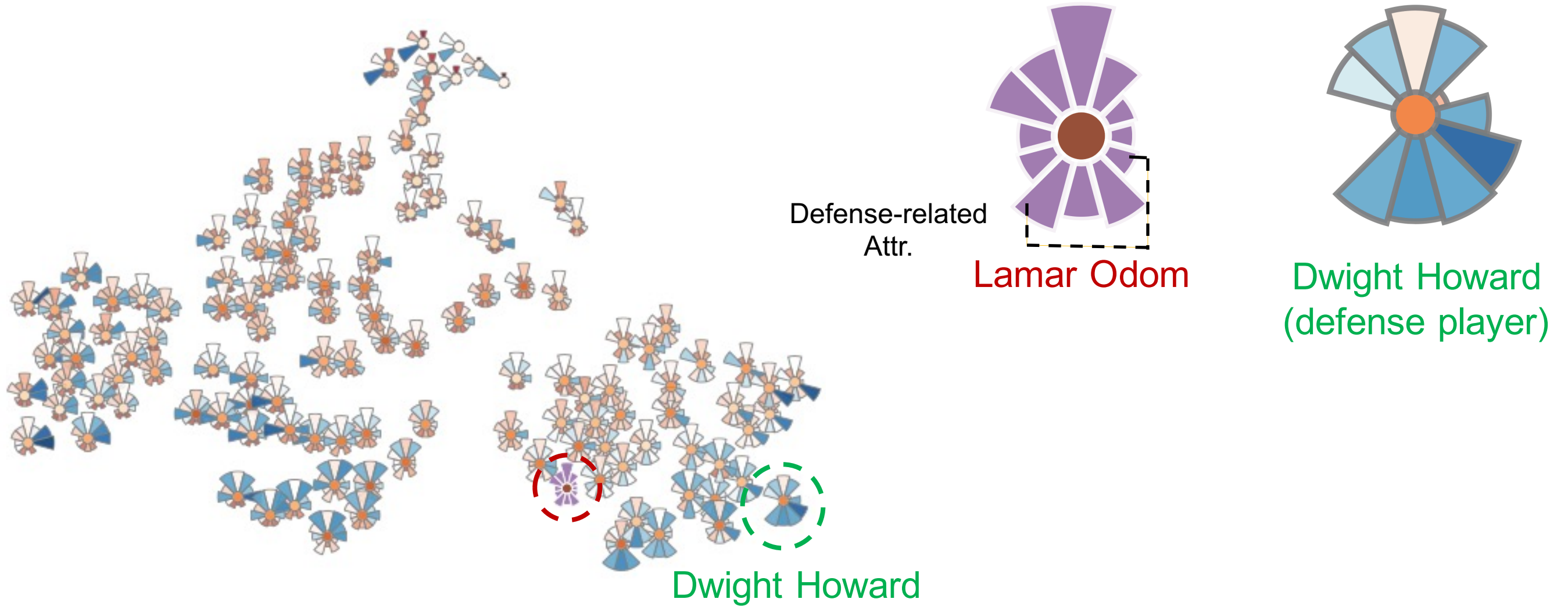
Two outliers with high values on *games played* and *FG%*

Projection View – Case Studies



Lamar Odom has the largest dominating score (central circle color)

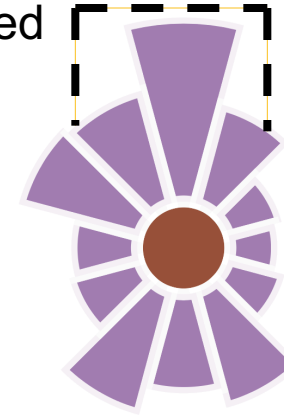
Projection View – Case Studies



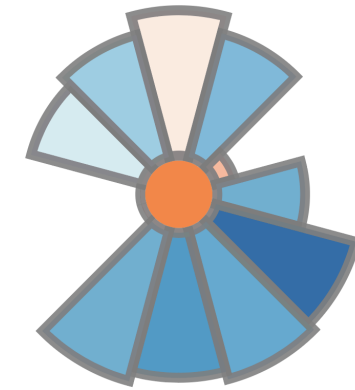
Switching to *focus mode* : three clusters can be found

Projection View – Case Studies

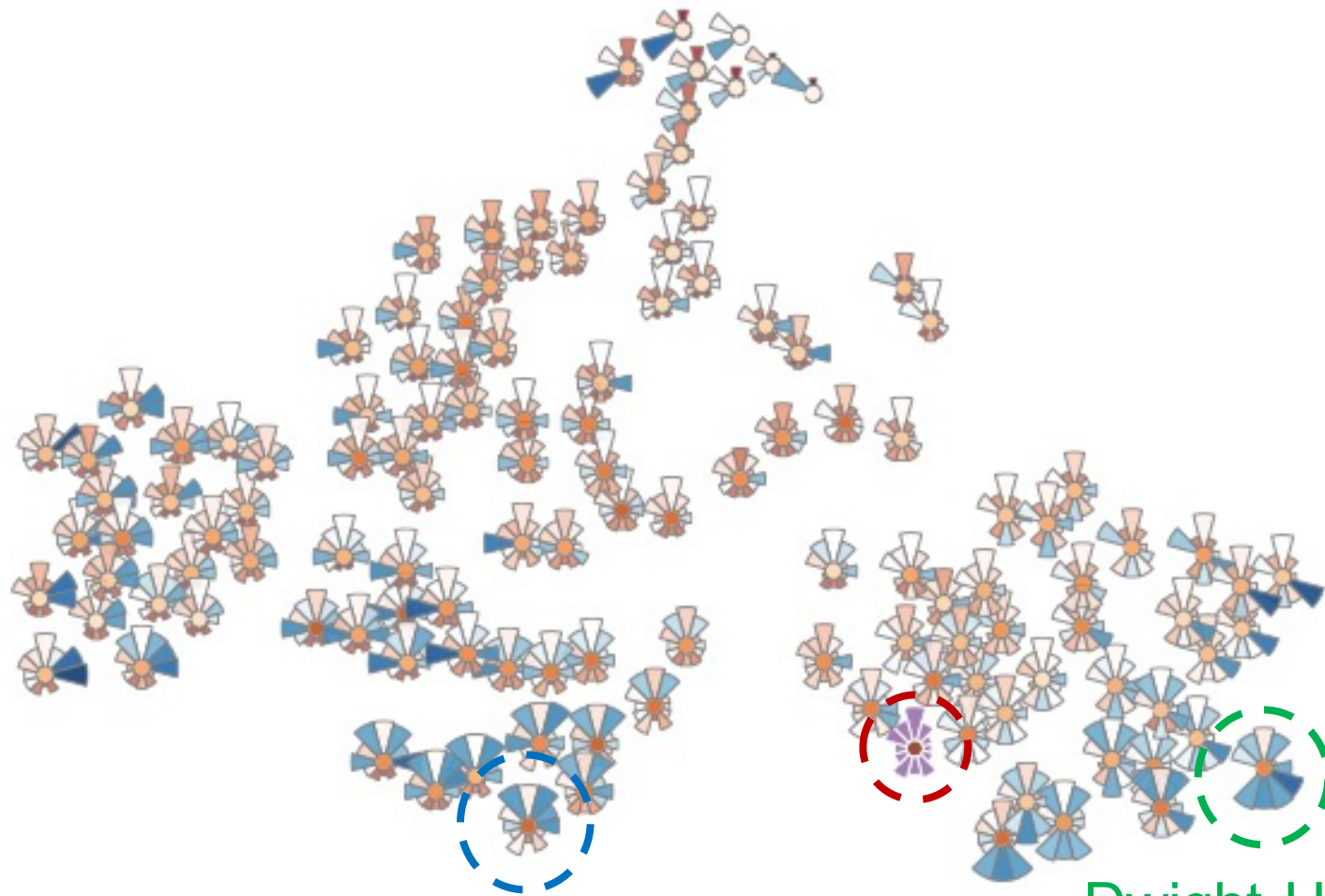
Point made-related
Attr.



Lamar Odom

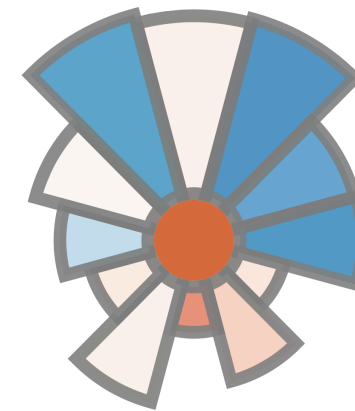


Dwight Howard
(defense player)



LeBron James

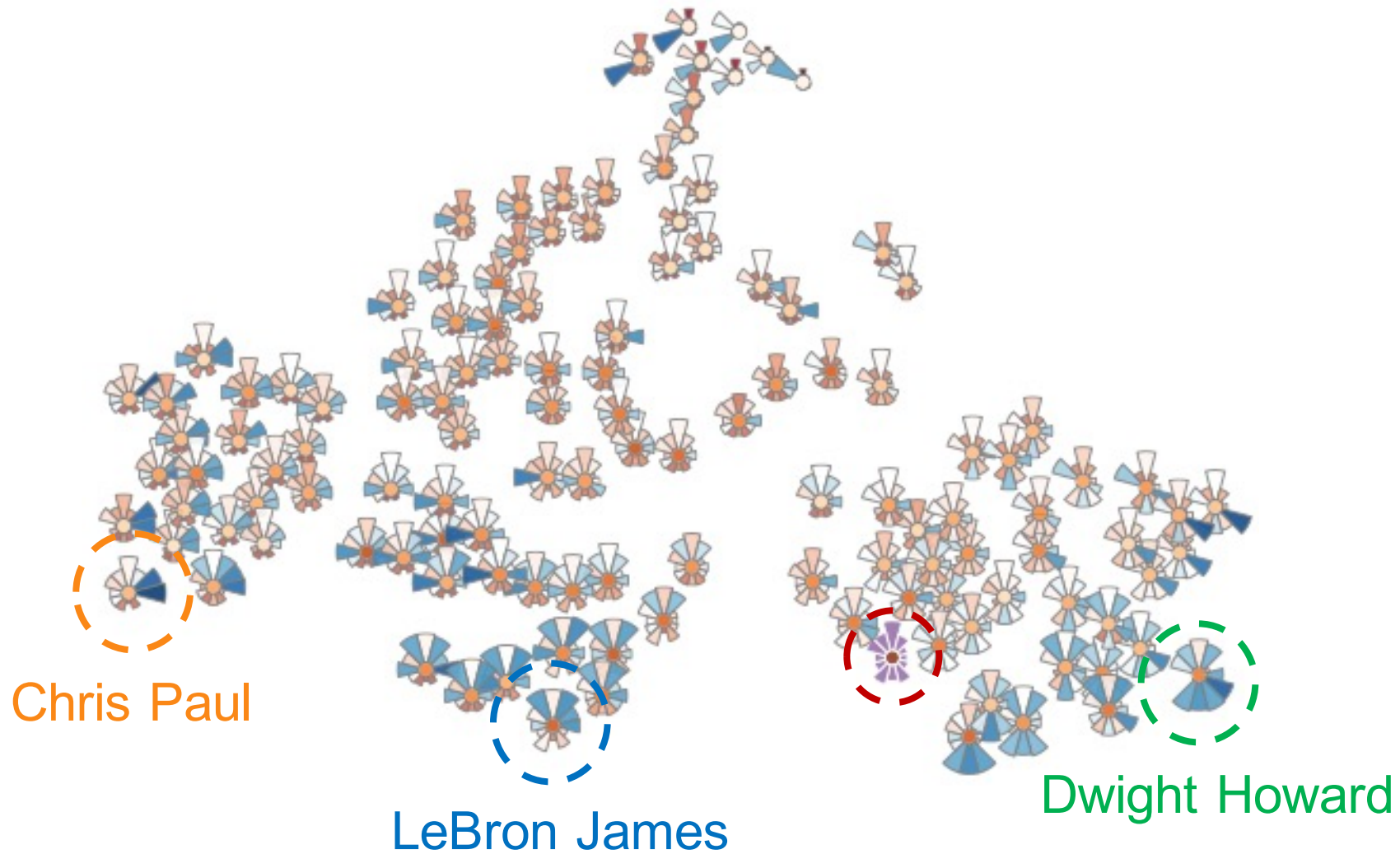
Dwight Howard



LeBron James
(pointer player)

Switching to *focus mode* : three clusters can be found

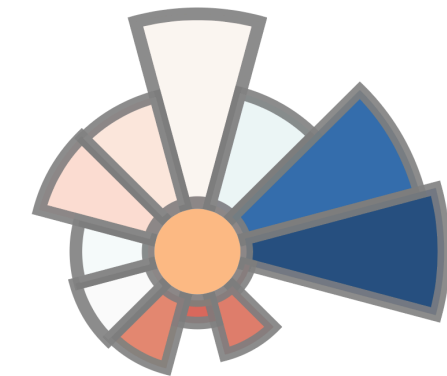
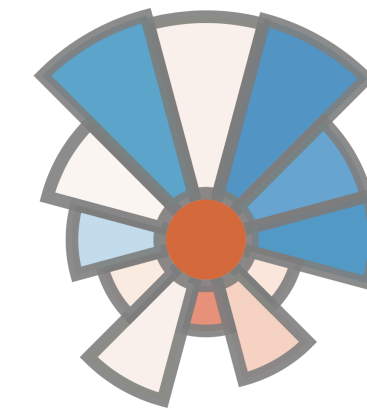
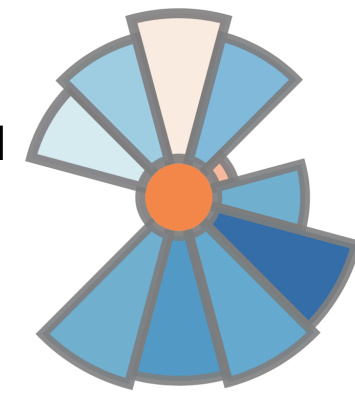
Projection View – Case Studies



Switching to *focus mode* : three clusters can be found

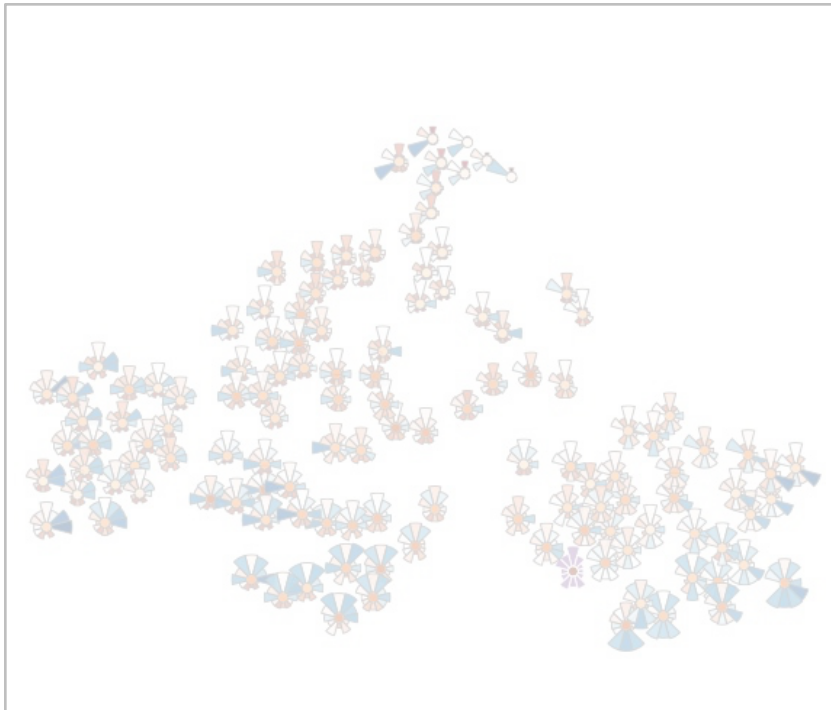


Lamar Odom

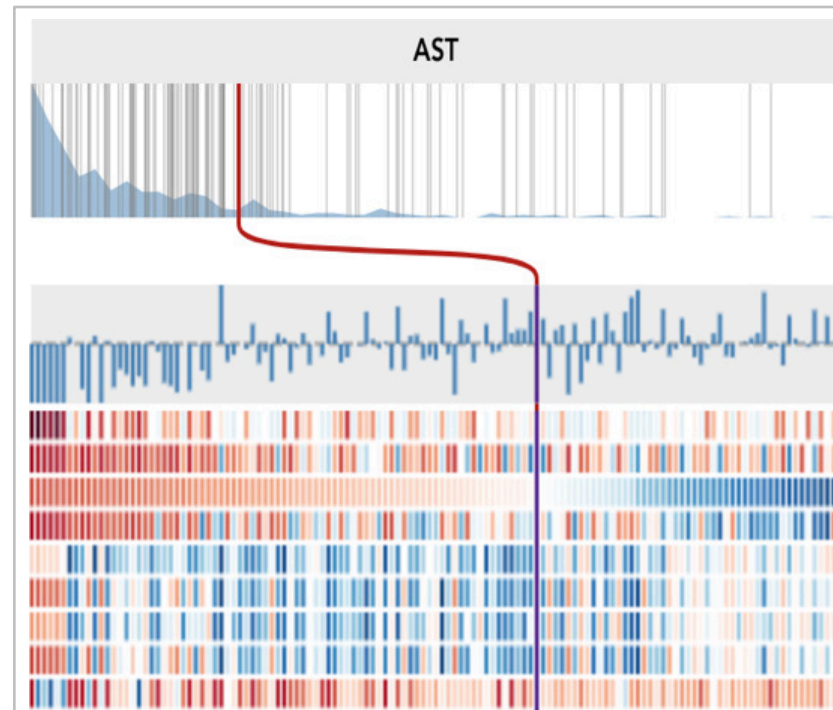


SkyLens – Tabular View

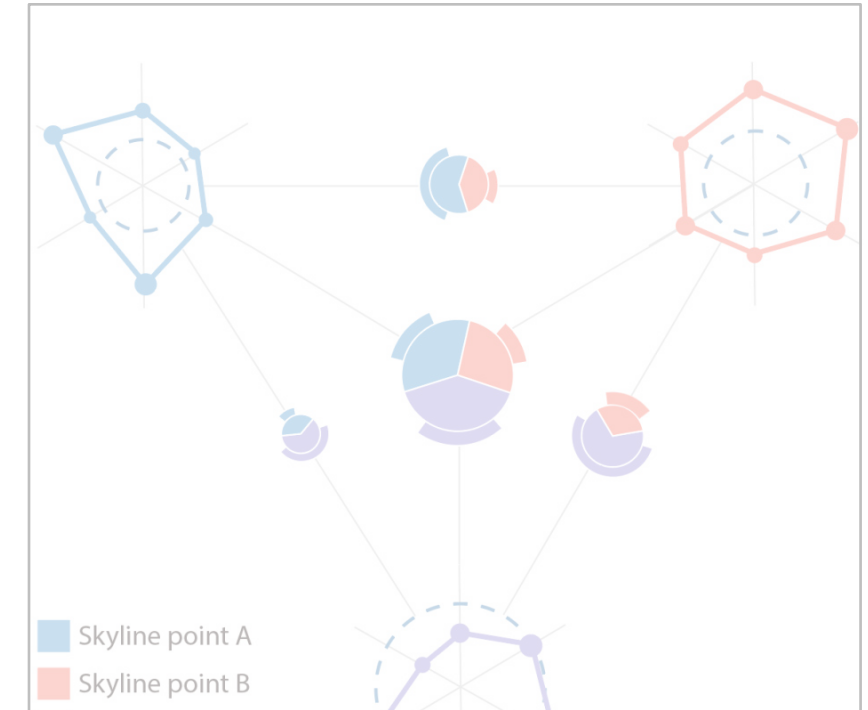
Projection View



Tabular View



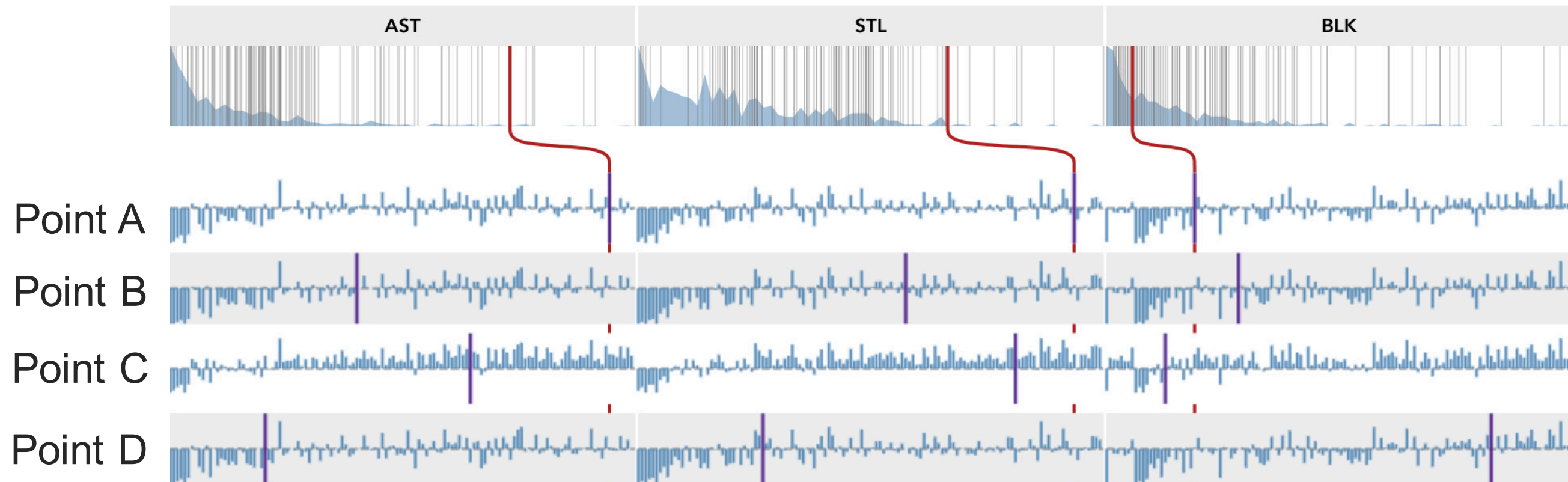
Comparison View



Tabular View: infer the underlying reasons that make a point in skyline

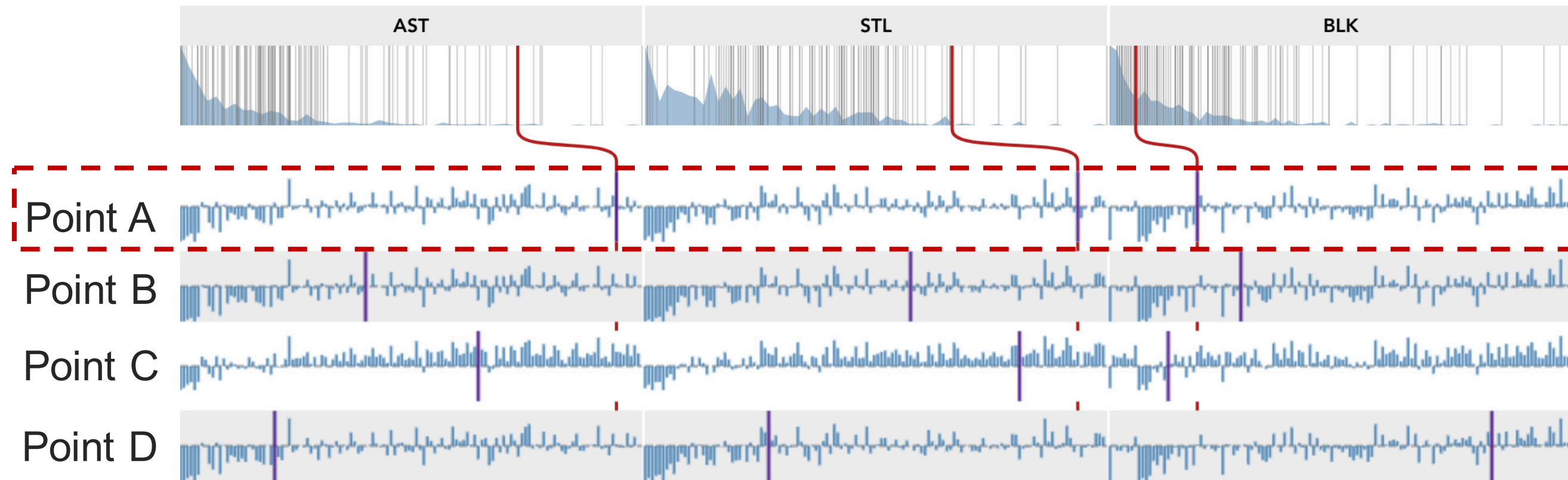
Tabular View

- Methods: matrix representation & in-cell bar chart visualization



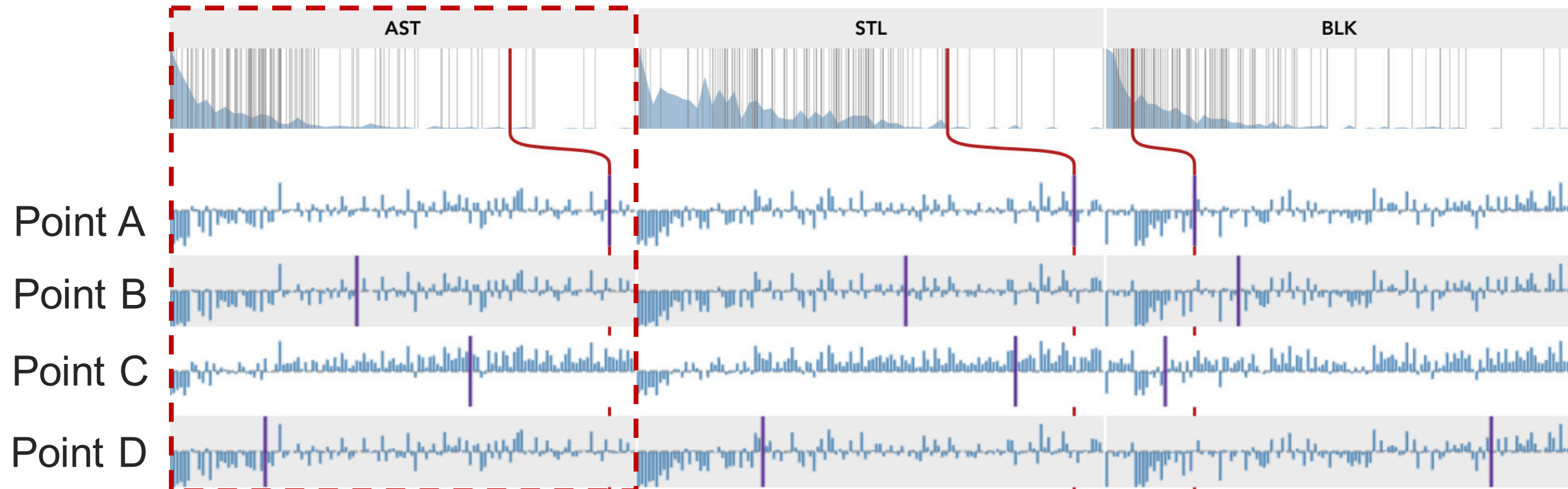
Tabular View

- Methods: matrix representation & in-cell bar chart visualization
 - Each row represents a skyline point



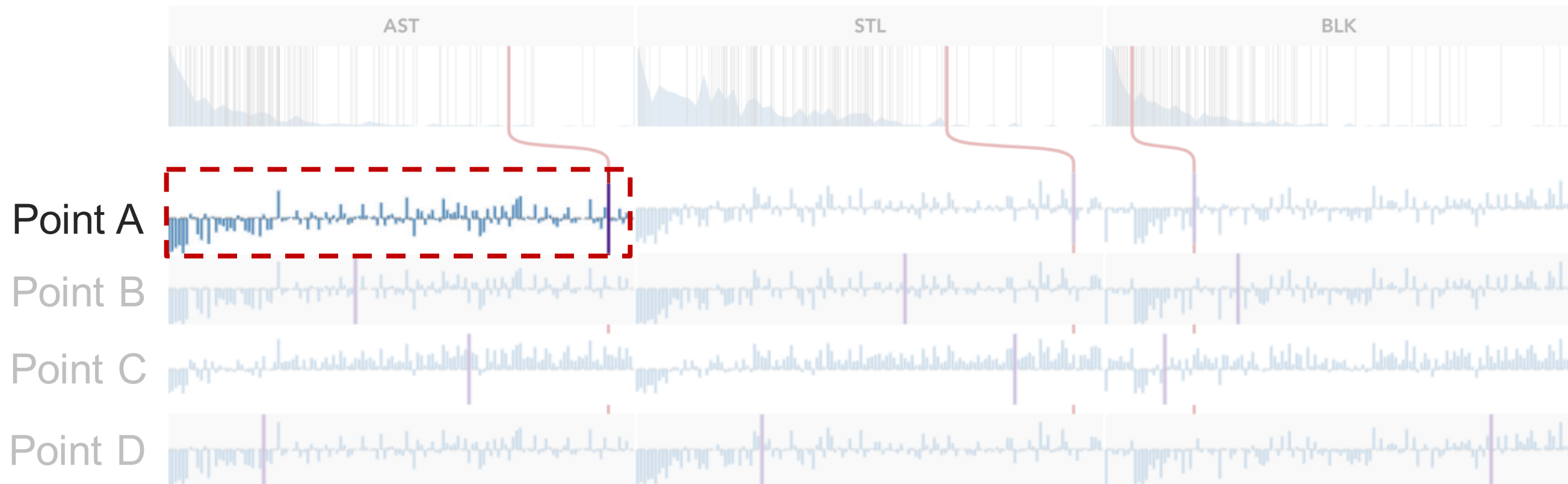
Tabular View

- Methods: matrix representation & in-cell bar chart visualization
 - Each row represents a skyline point
 - Each column represents an attribute



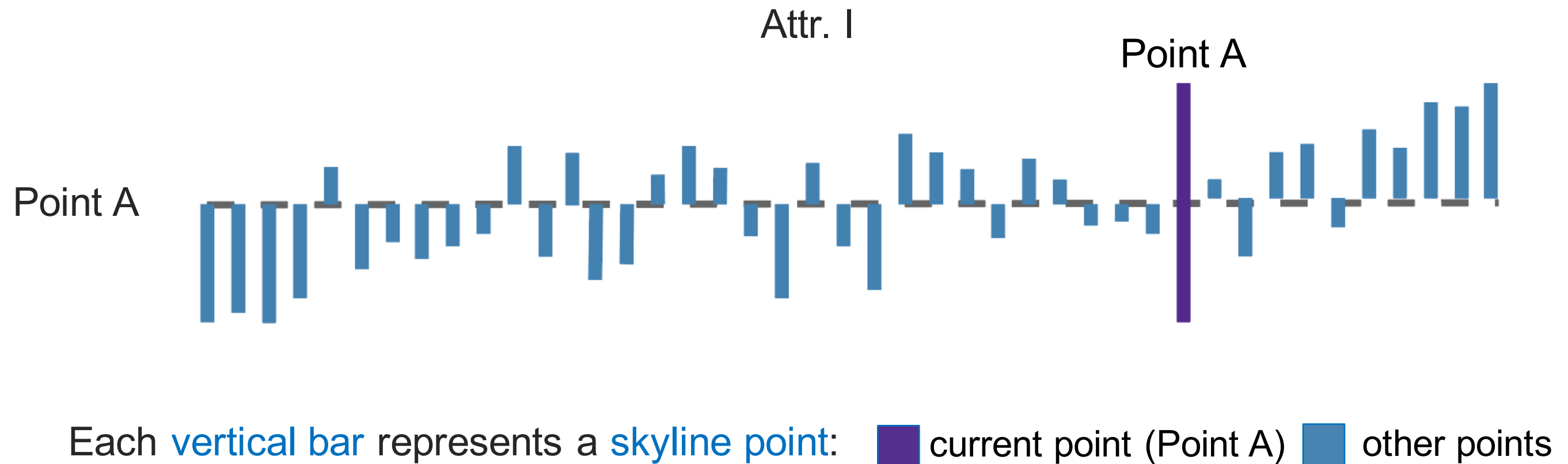
Tabular View

- Table cell – divergent bar chart visualization
 - Goal: summarize the overall differences among skyline points



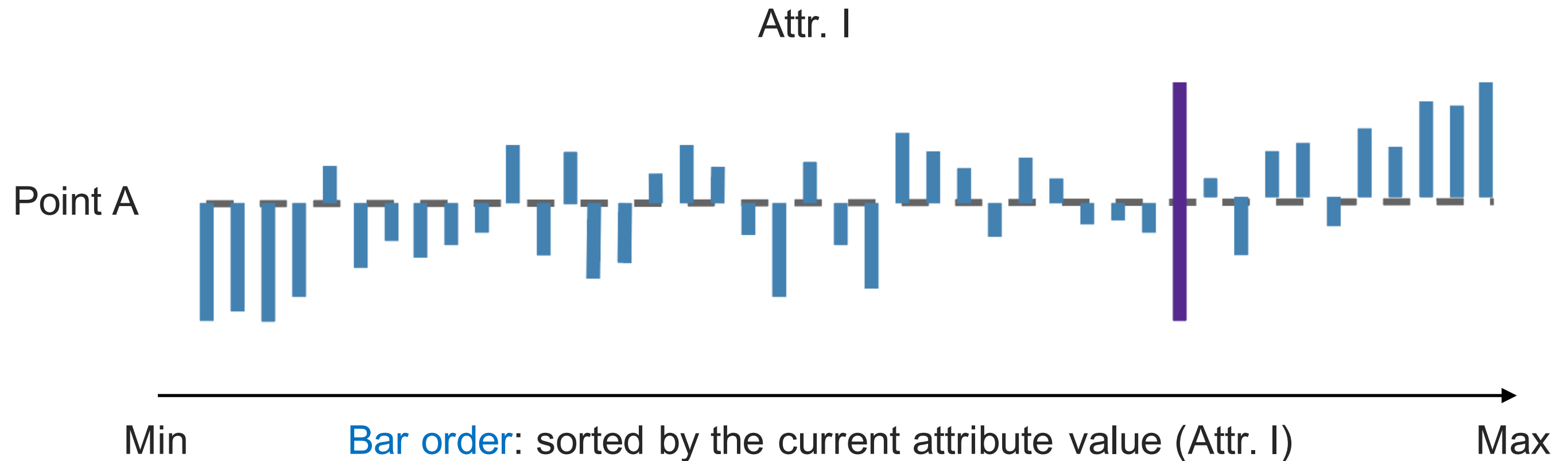
Tabular View

- Table cell – divergent bar chart visualization
 - Goal: summarize the overall differences between skyline points



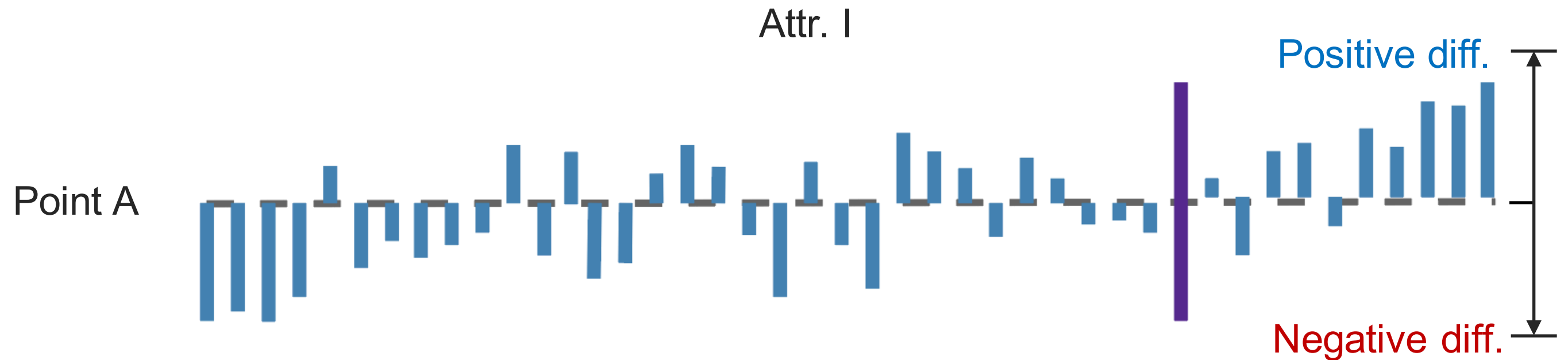
Tabular View

- Table cell – divergent bar chart visualization
 - Goal: summarize the overall differences between skyline points



Tabular View

- Table cell – divergent bar chart visualization
 - Goal: summarize the overall differences between skyline points

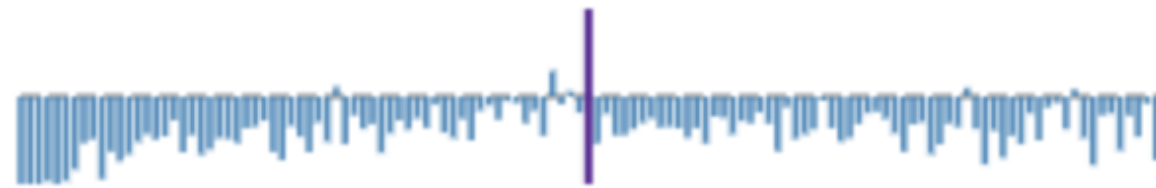


Bar length: other skyline points' average value differences compared with point A

Tabular View – Case Studies

- Table cell – divergent bar chart visualization
 - Goal: summarize the overall differences between skyline points

Attr. I: games played



Attr. II: points



Attr. III: assist



Attr. IV: block



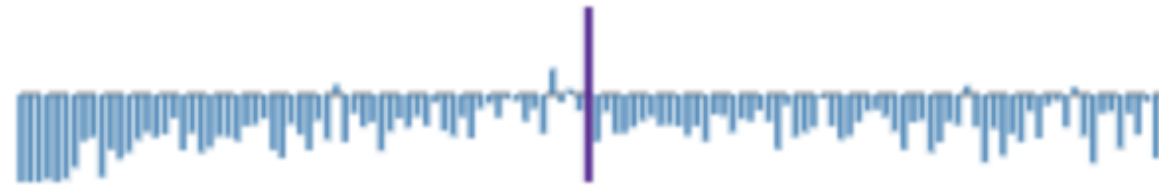
LeBron James:

- High ranking on **points & assist**

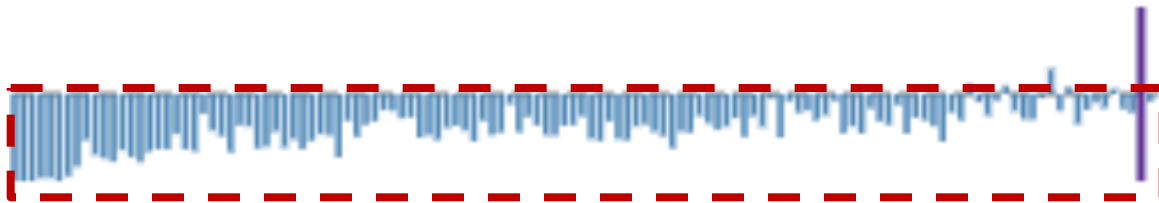
Tabular View – Case Studies

- Table cell – divergent bar chart visualization
 - Goal: summarize the overall differences between skyline points

Attr. I: games played



Attr. II: points



Attr. III: assist



Attr. IV: block



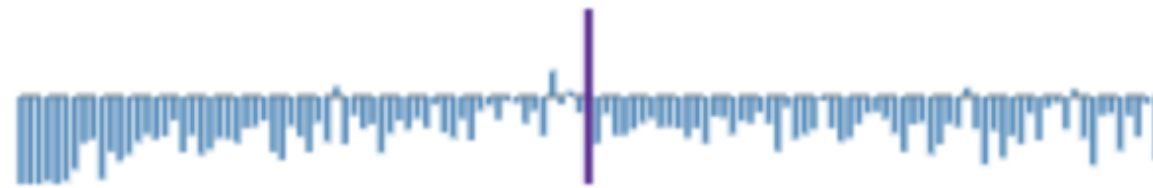
LeBron James:

- High ranking in points & assist
- **Better overall performance** than most skyline points

Tabular View – Case Studies

- Table cell – divergent bar chart visualization
 - Goal: summarize the overall differences between skyline points

Attr. I: game played



Dwight Howard

Attr. II: points



Attr. III: assist



Attr. IV: block

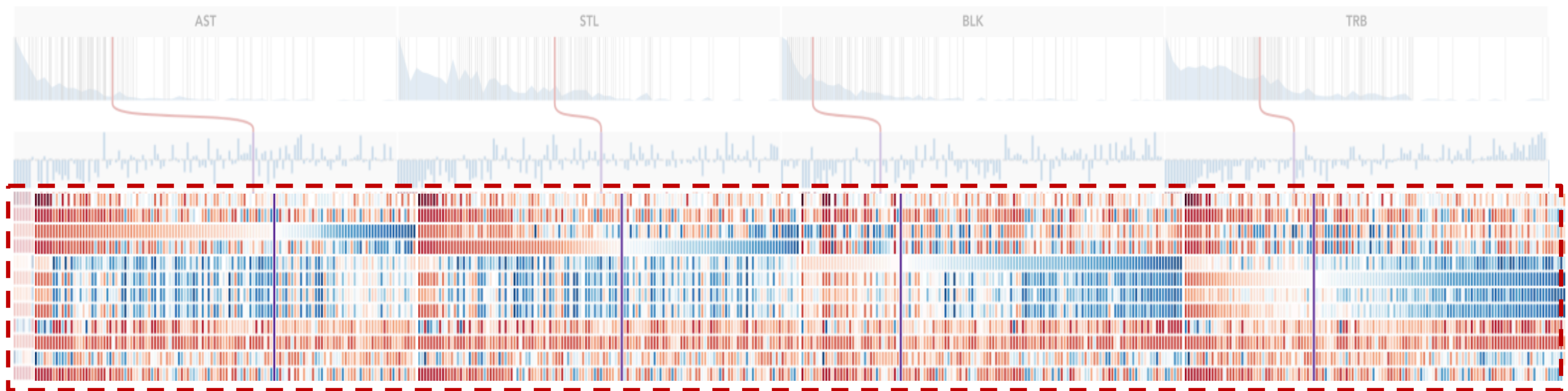


LeBron James:

- High ranking in points & assist
- Better overall performance than most skyline points
- Dwight has an overall **comparable performance** with LeBron

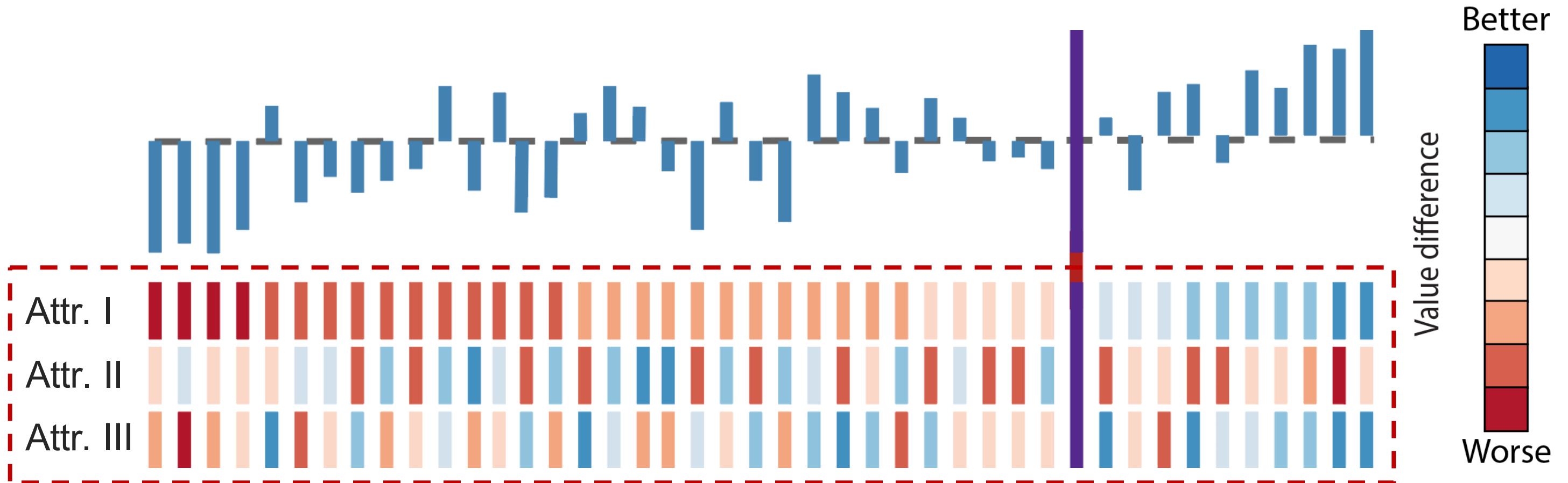
Tabular View

- Table cell interaction: expanding a row for detailed information



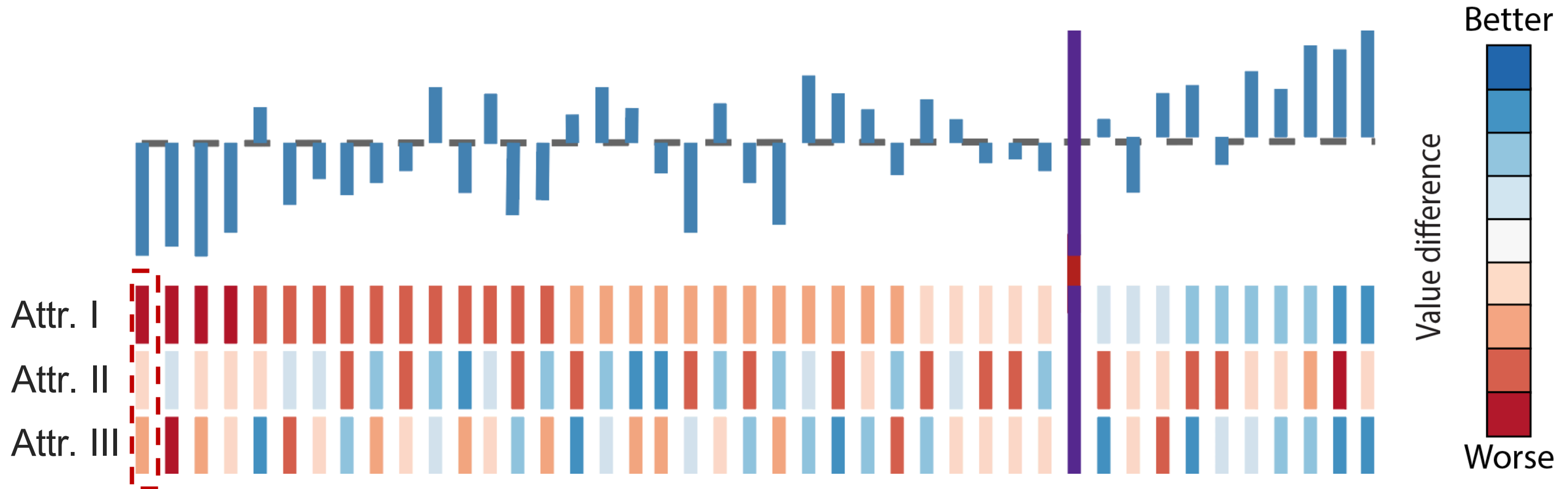
Tabular View

- Table cell interaction: expanding a row for detailed information



Tabular View

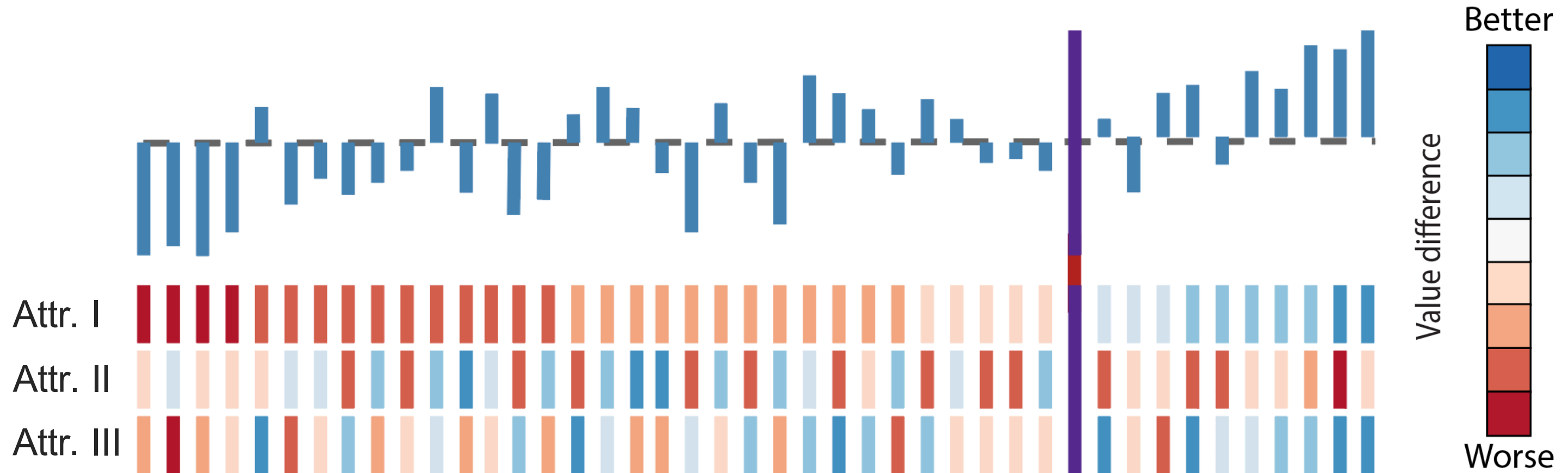
- Table cell interaction: expanding a row for detailed information



Each column is an extension of the corresponding vertical blue bar and represents the same skyline point

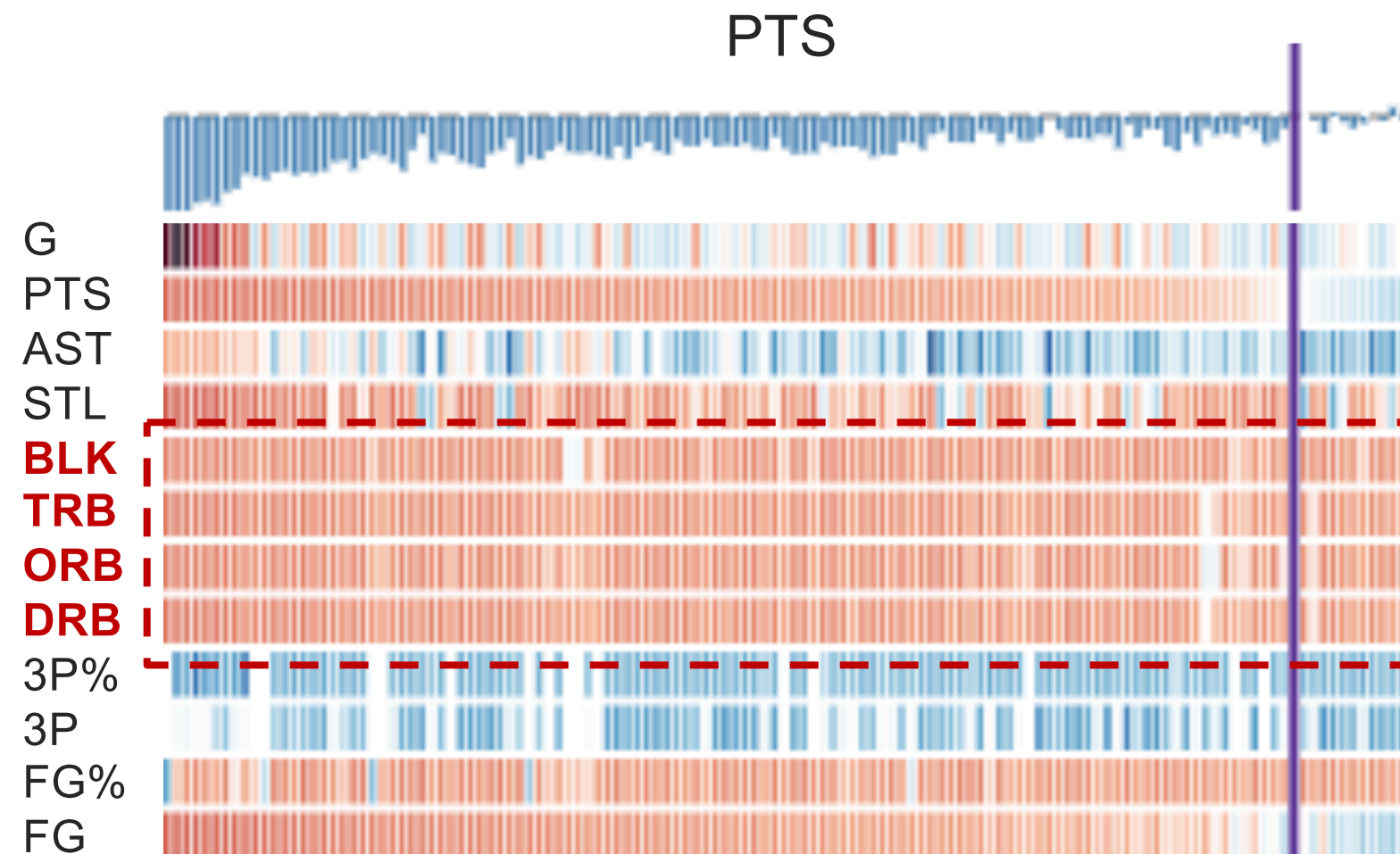
Tabular View

- Table cell interaction: expanding a row for detailed information



Matrix cell color: the value difference of the corresponding attribute (Attr. III)

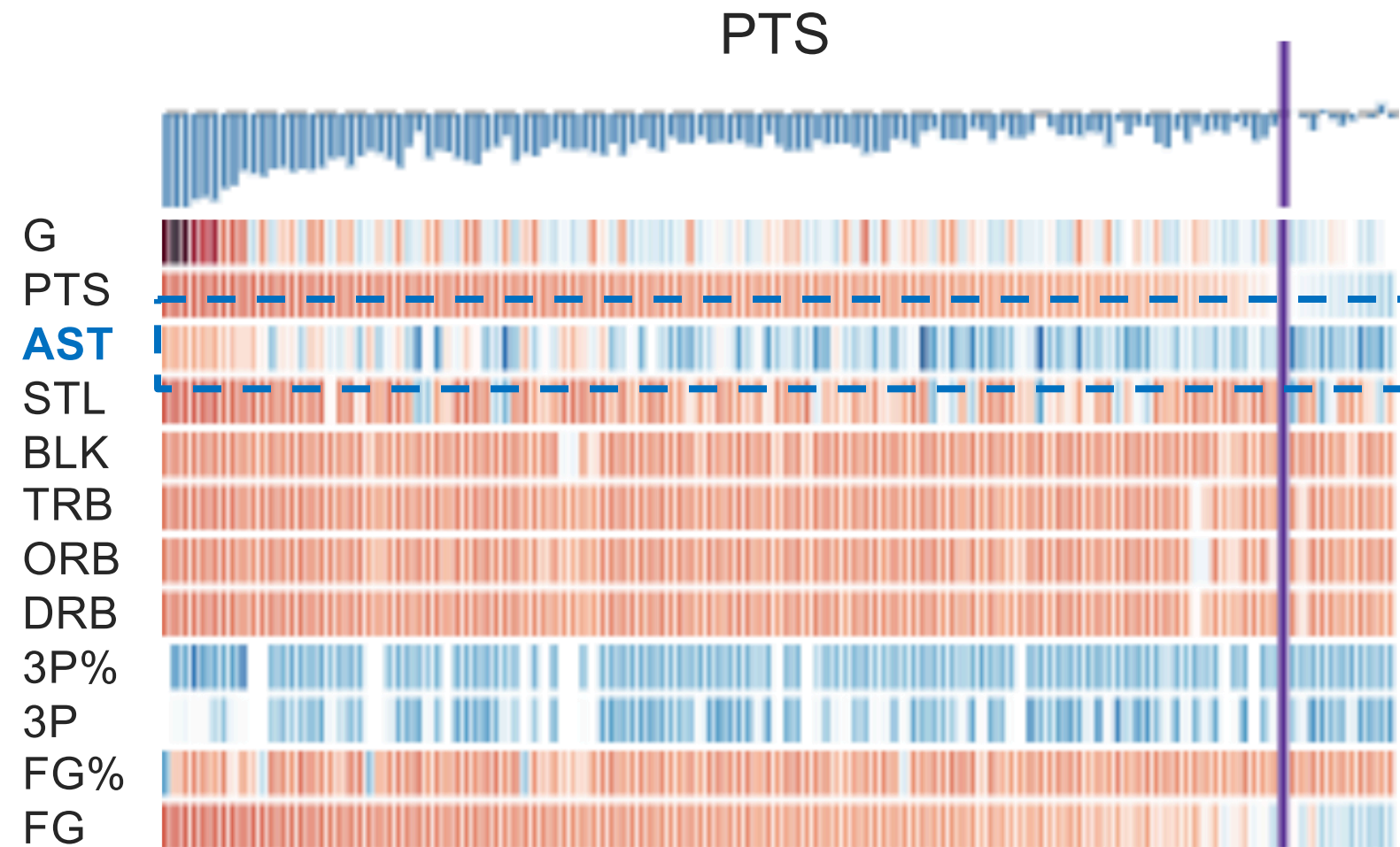
Tabular View – Case Studies



Dwight Howard:

- No players has better performance than him in **defense-related** attributes

Tabular View – Case Studies

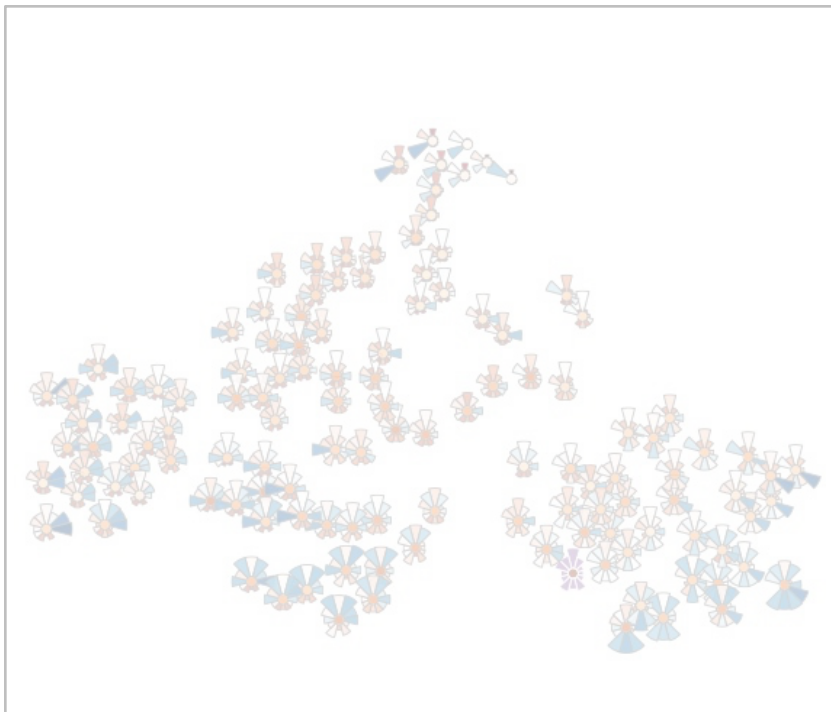


Dwight Howard:

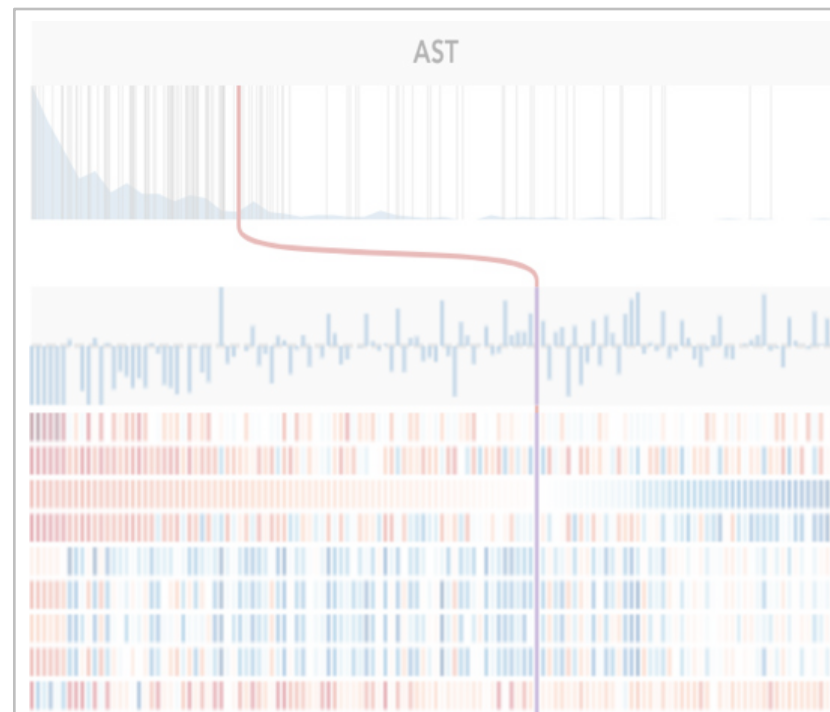
- No players has better performance than him in defense-related attributes
- Many players outperform him in **AST**

SkyLens – Comparison View

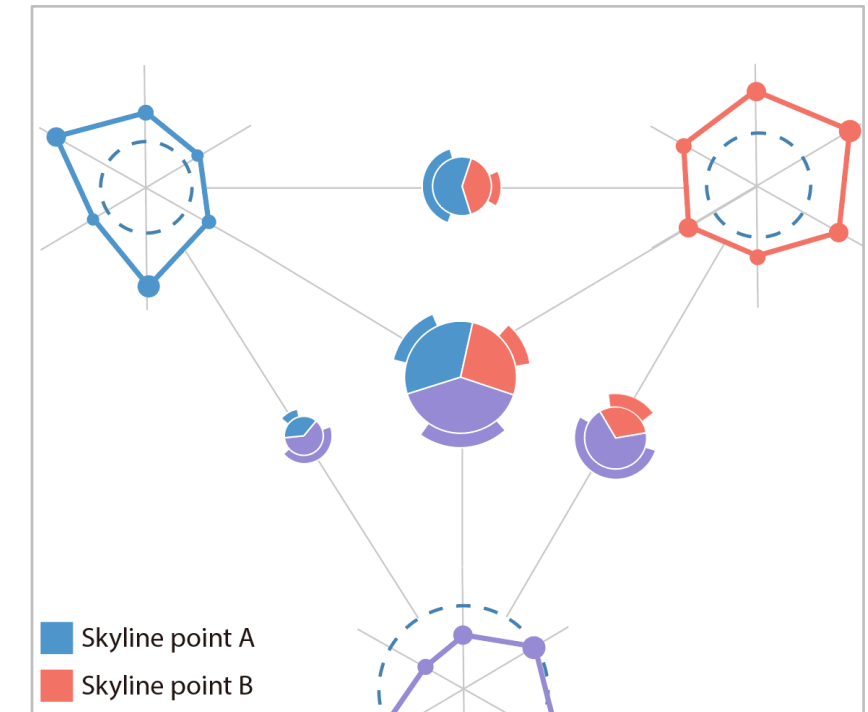
Projection View



Tabular View



Comparison View

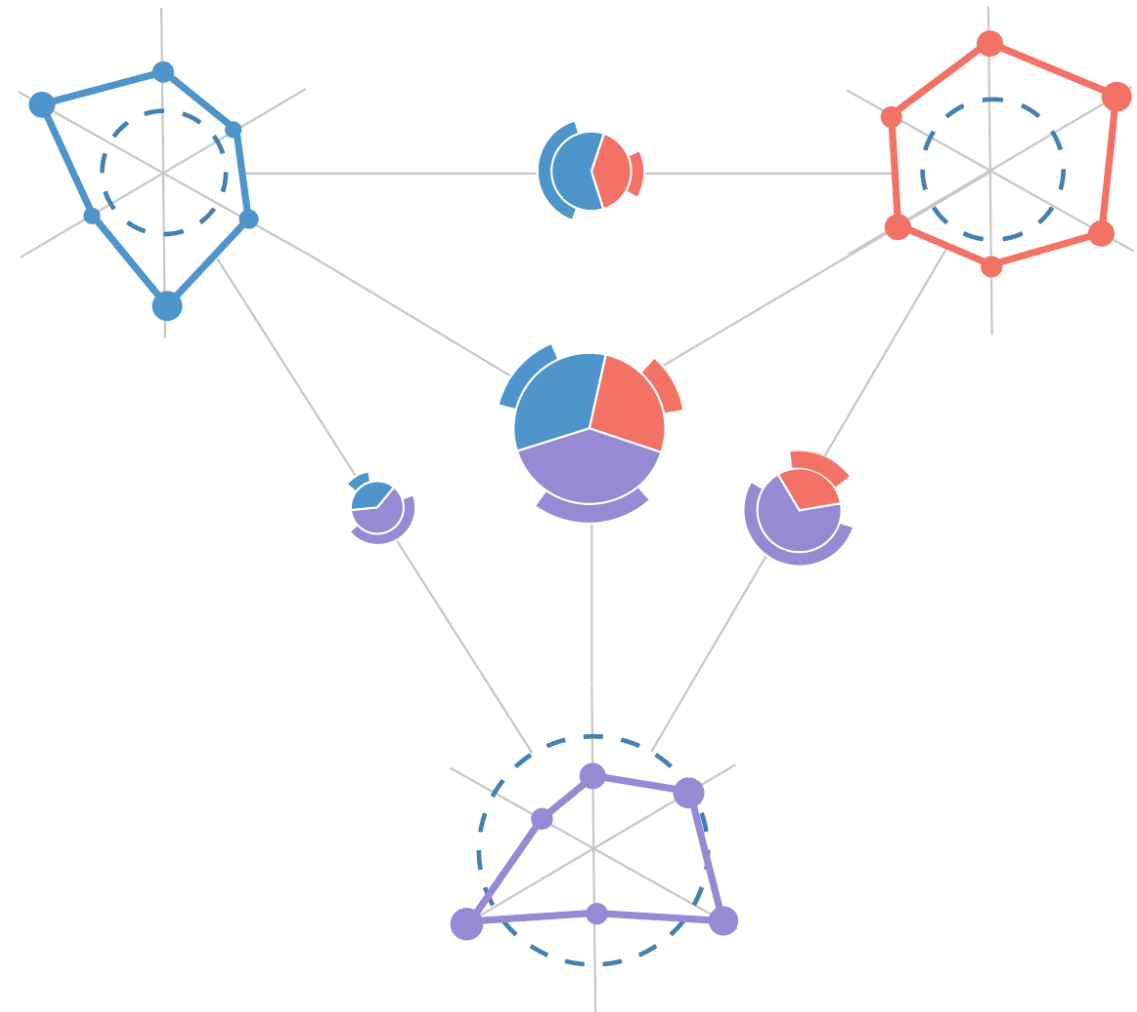


Comparison View: support a thorough comparison between skyline points

Comparison View

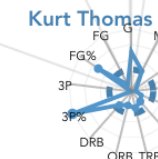
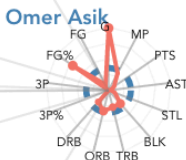
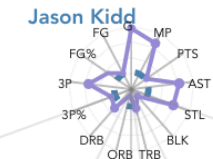
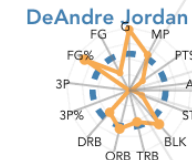
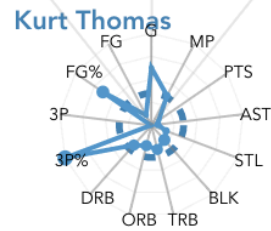
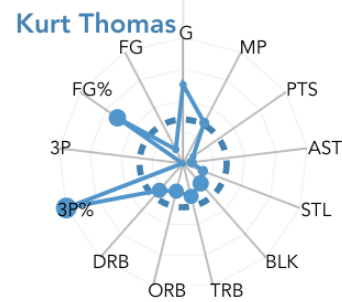
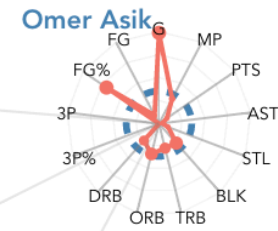
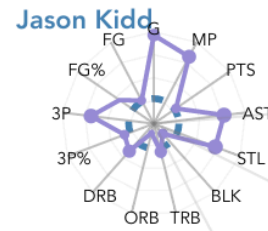
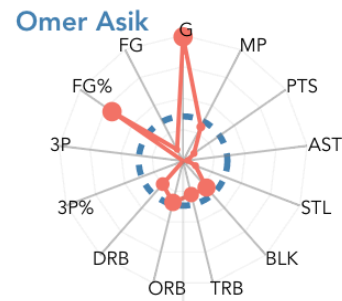
- Methods: radar charts & domination glyphs
 - Comparing attribute values
 - Examining dominating scores
 - Investigating dominated points

Goal: a thorough comparison
on 2 ~ 5 skyline points



Comparison View

- Radial layout for the radar charts & domination glyphs

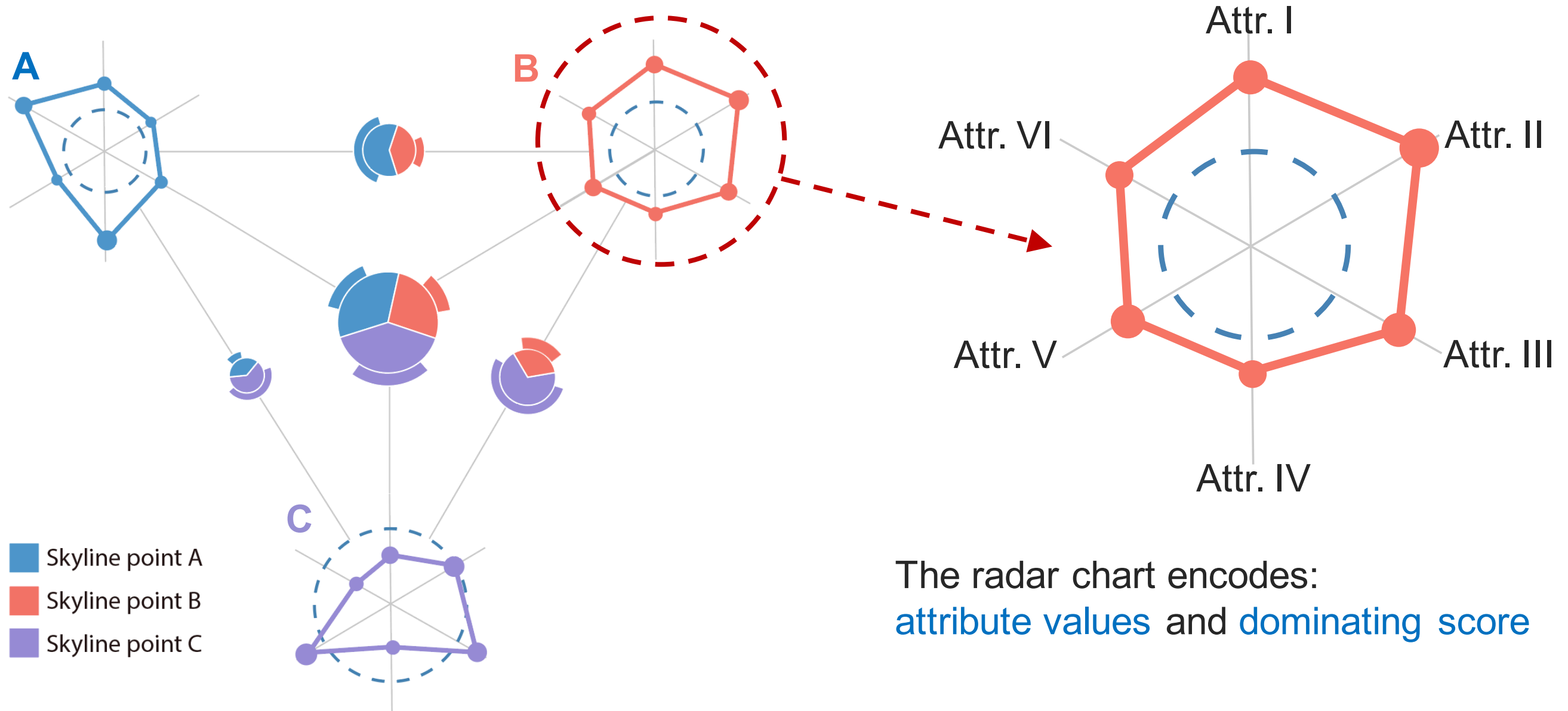


2-point comparison

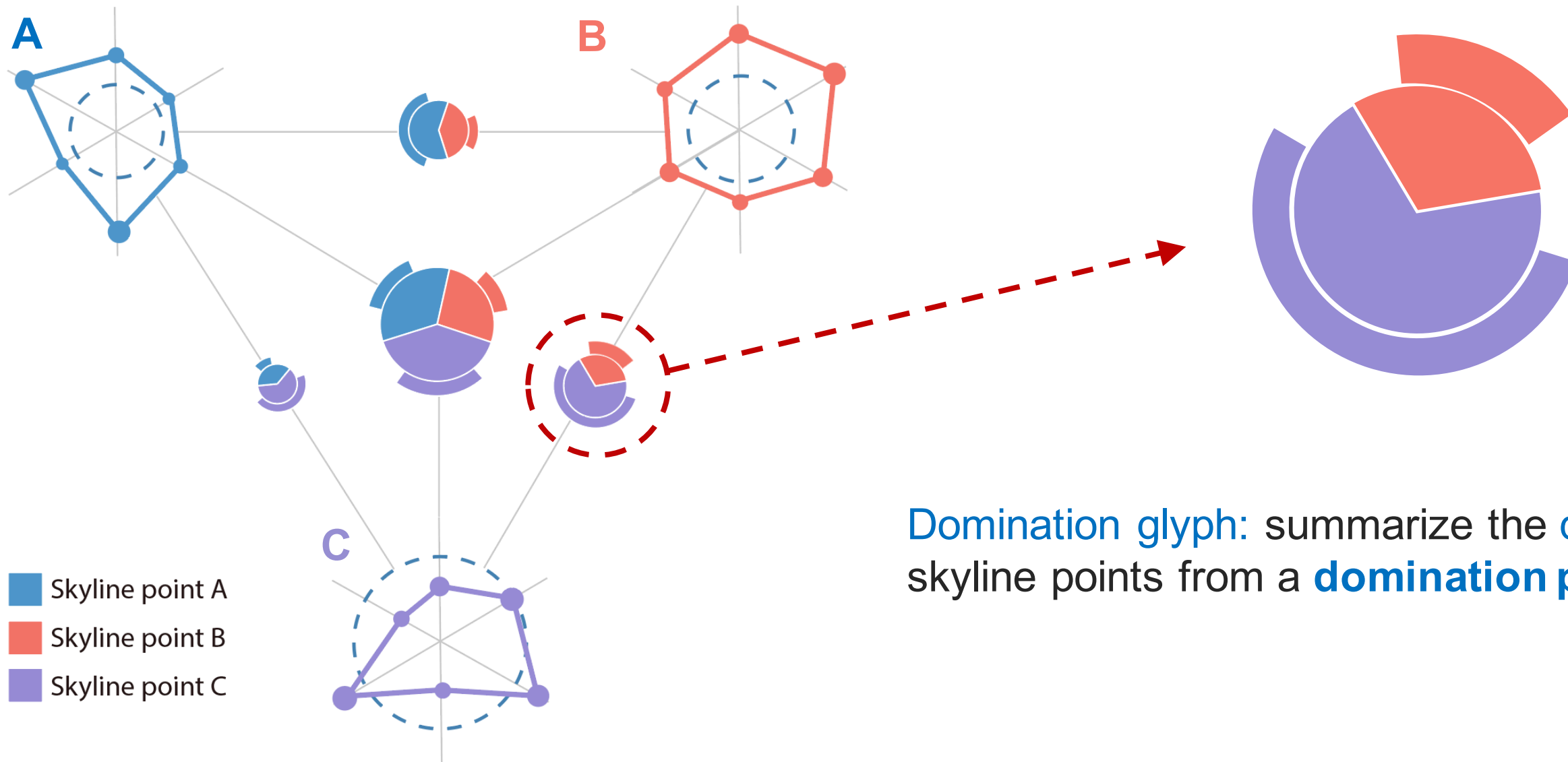
3-point comparison

4-point comparison

Comparison View – Radar Chart

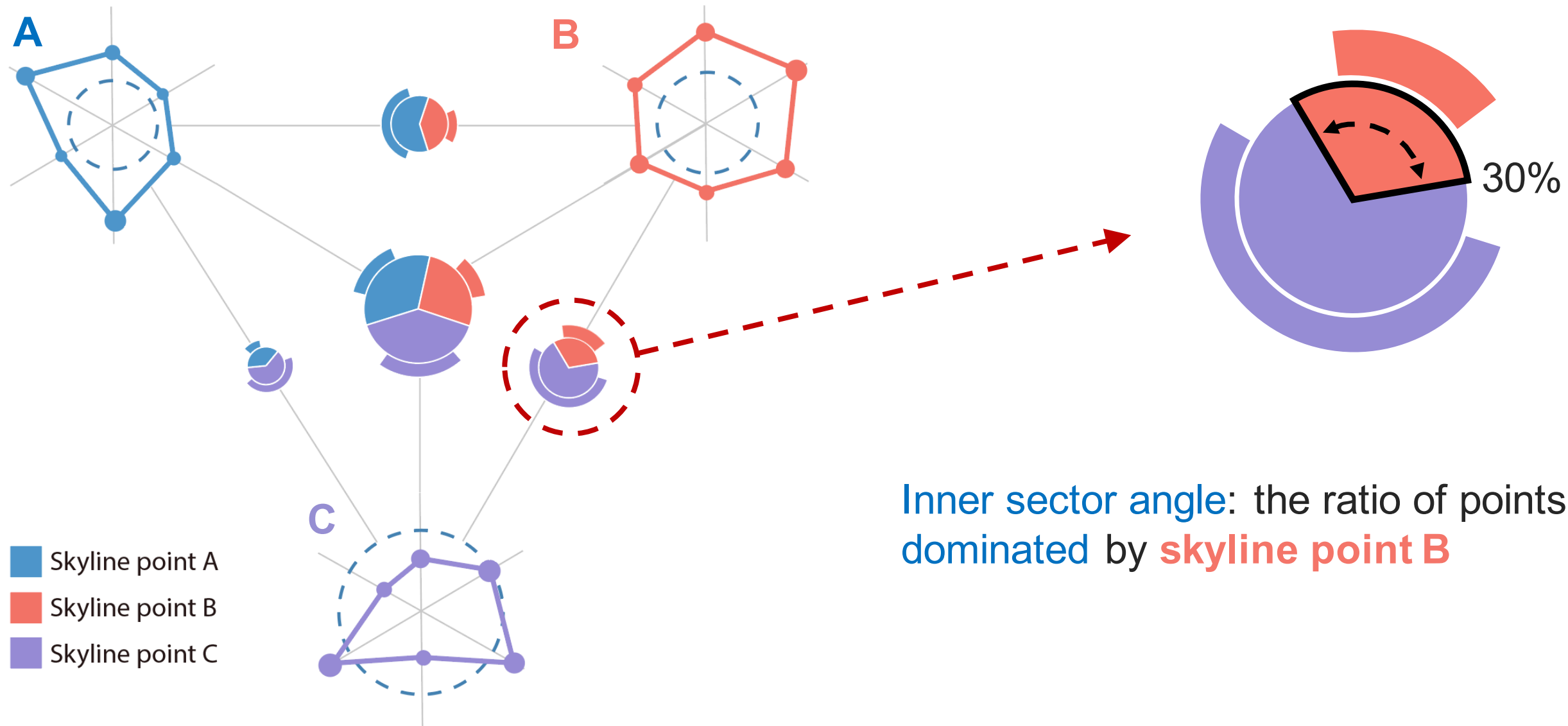


Comparison View – Domination Glyph

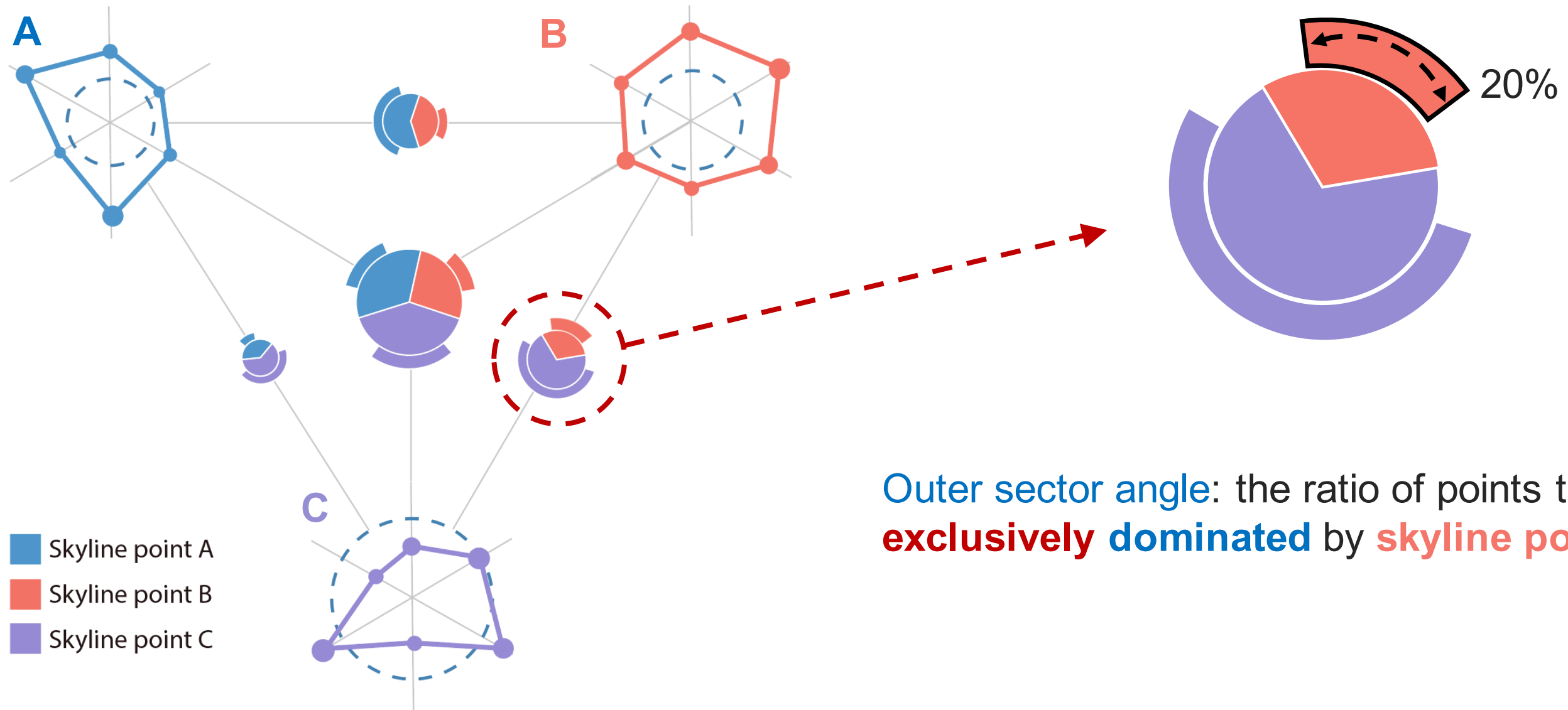


Domination glyph: summarize the differences among skyline points from a **domination perspective**.

Comparison View – Domination Glyph

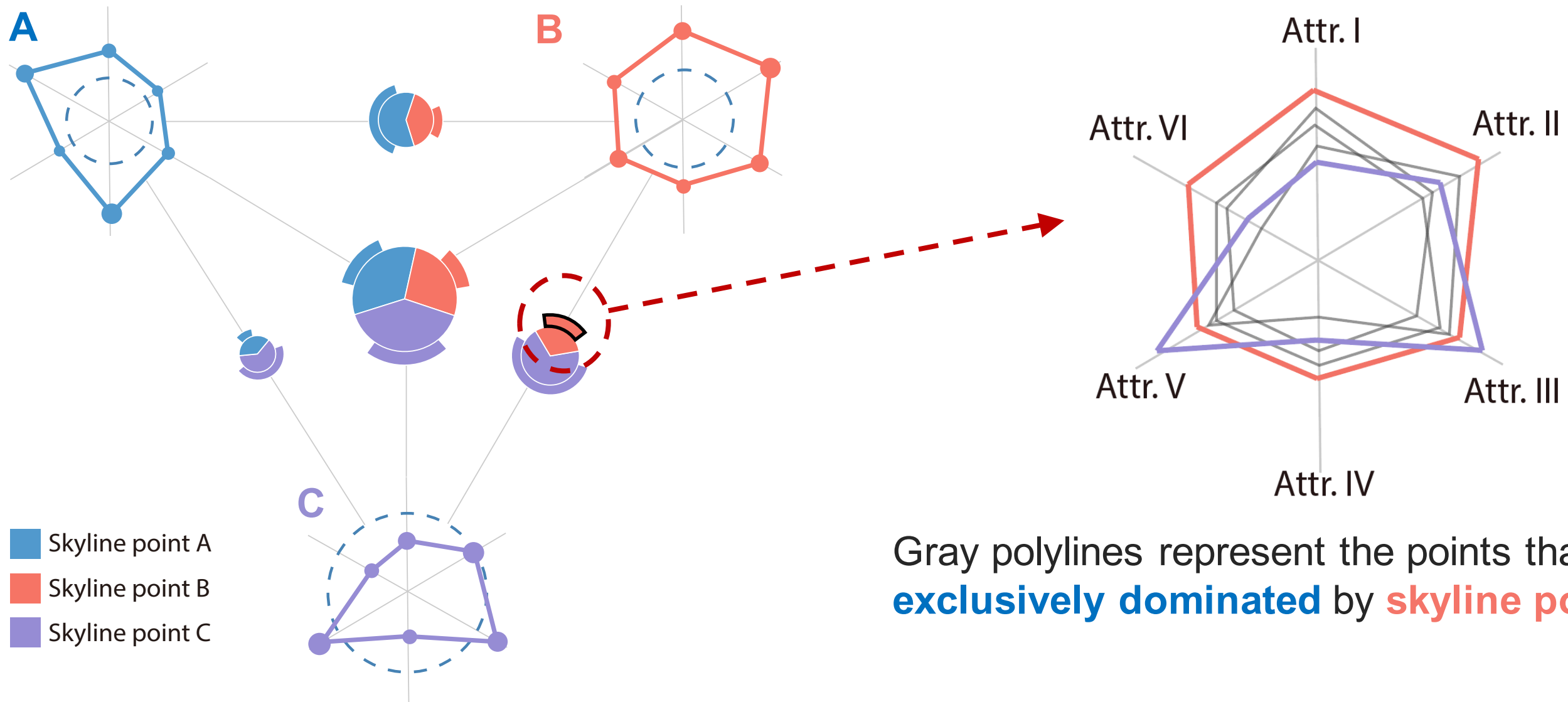


Comparison View – Domination Glyph



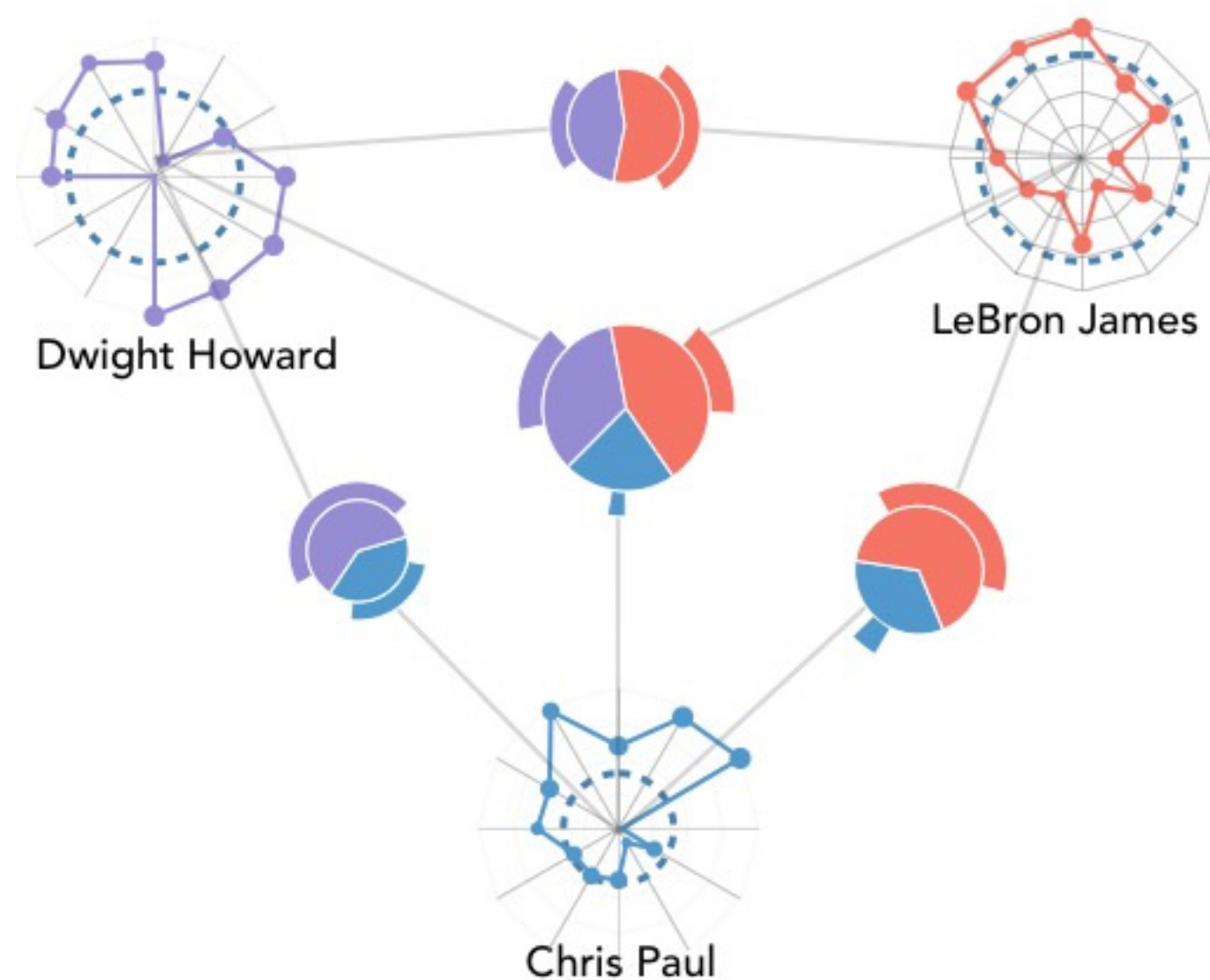
Comparison View – Domination Glyph Interaction

Hovering interaction: pop-up window showing the **overlaid radar chart**



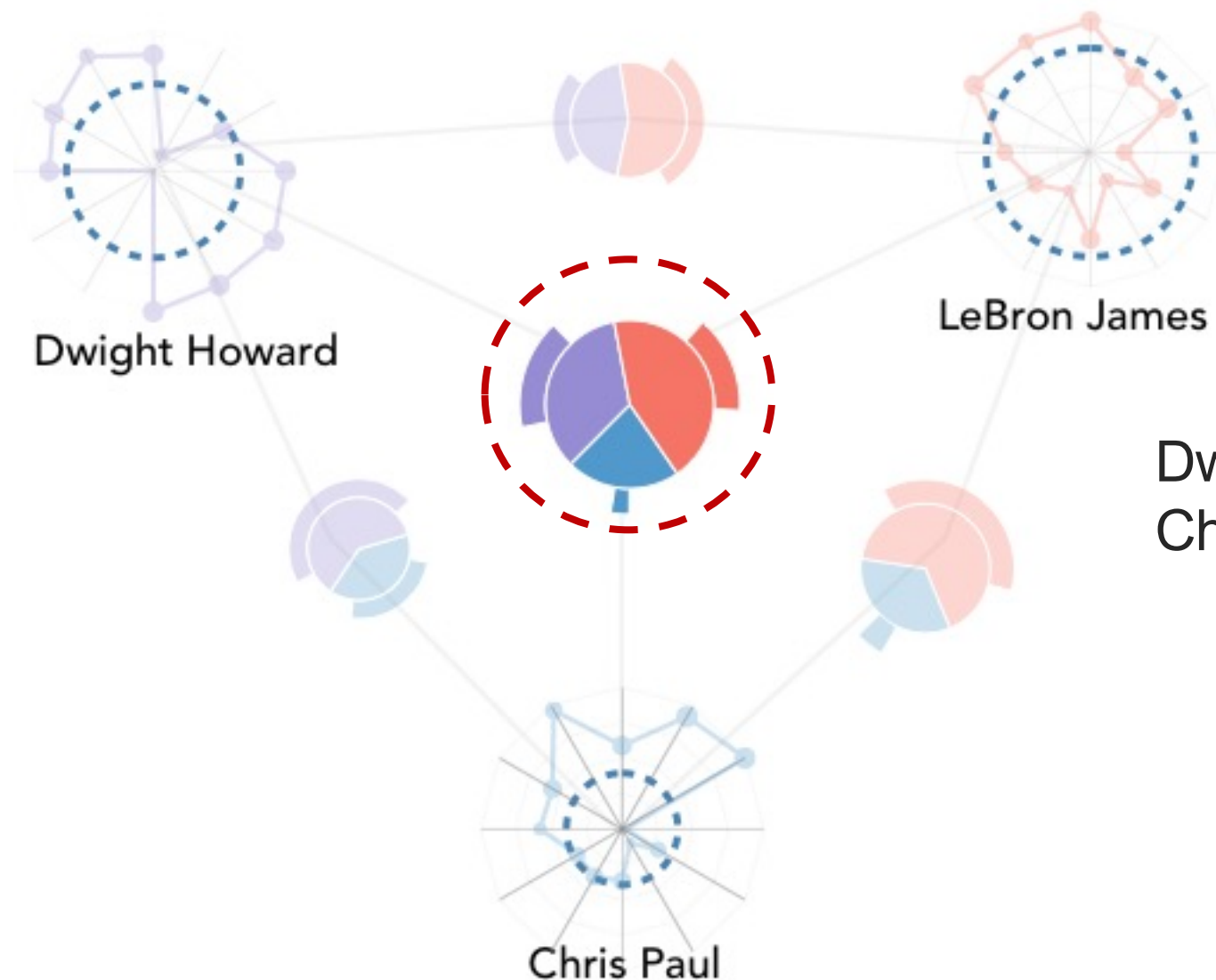
Comparison View – Case Studies

Comparing Dwight, LeBron, and Chris in the perspective of domination relation



Comparison View – Case Studies

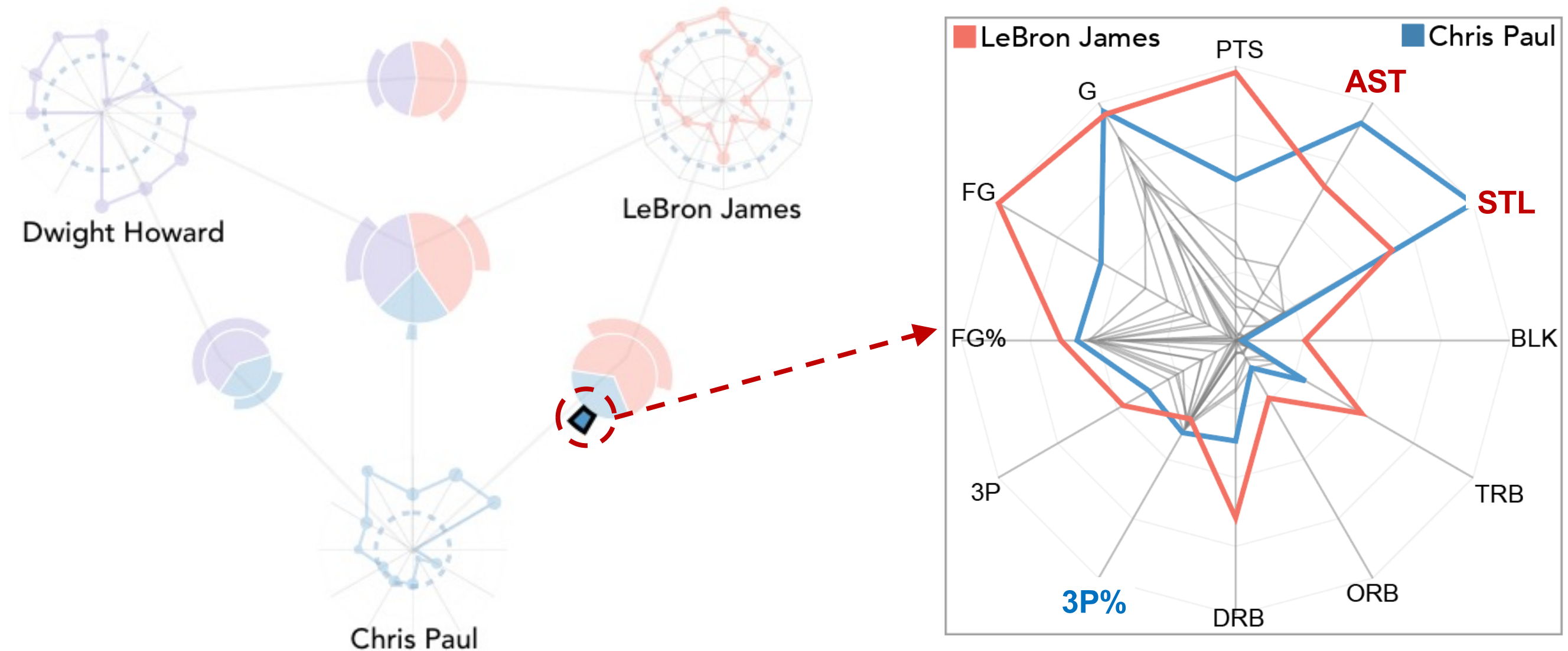
Comparing Dwight, LeBron, and Chris in the perspective of domination relation



Dwight and LeBron have **similar dominating scores**
Chris has a **smaller dominating score** compared with them

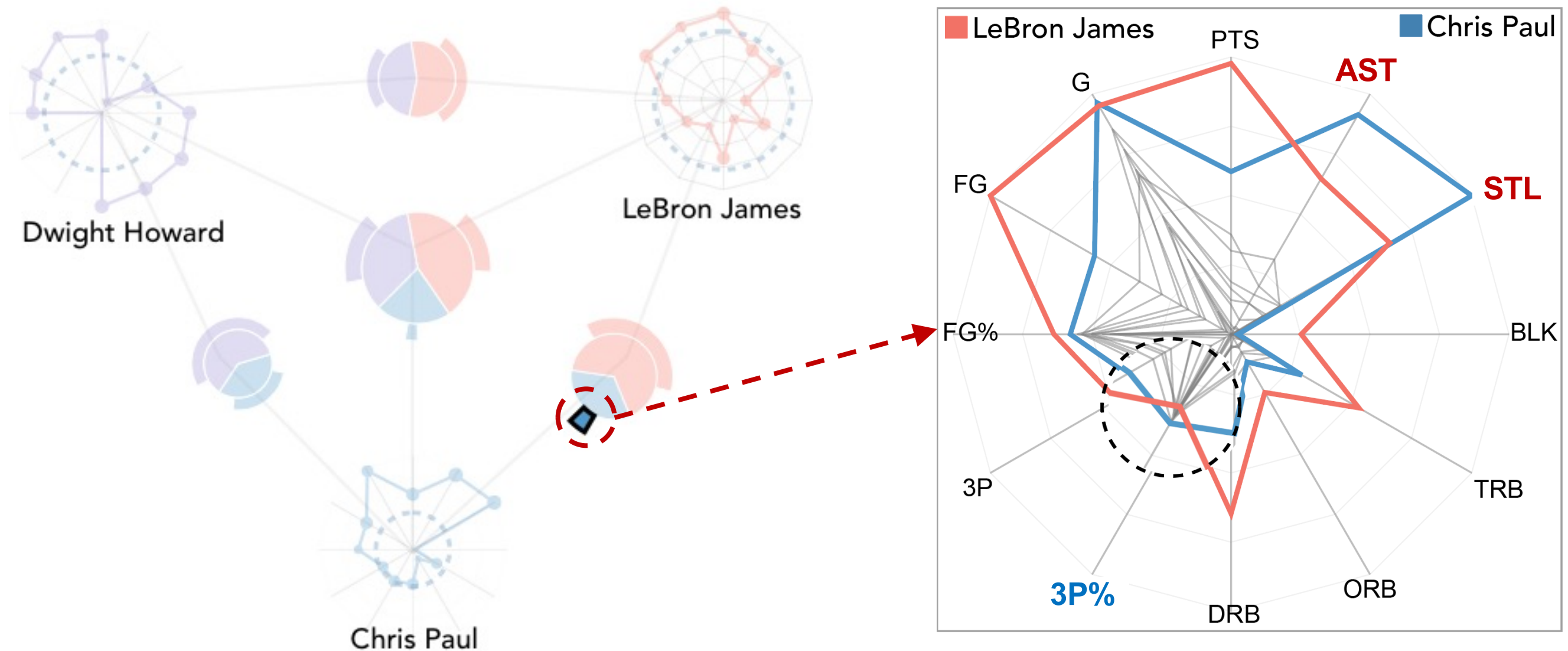
Comparison View – Case Studies

Hovering over the points that are **exclusively dominated** by Chris against LeBron



Comparison View – Case Studies

Hovering over the points that are **exclusively dominated** by Chris against LeBron

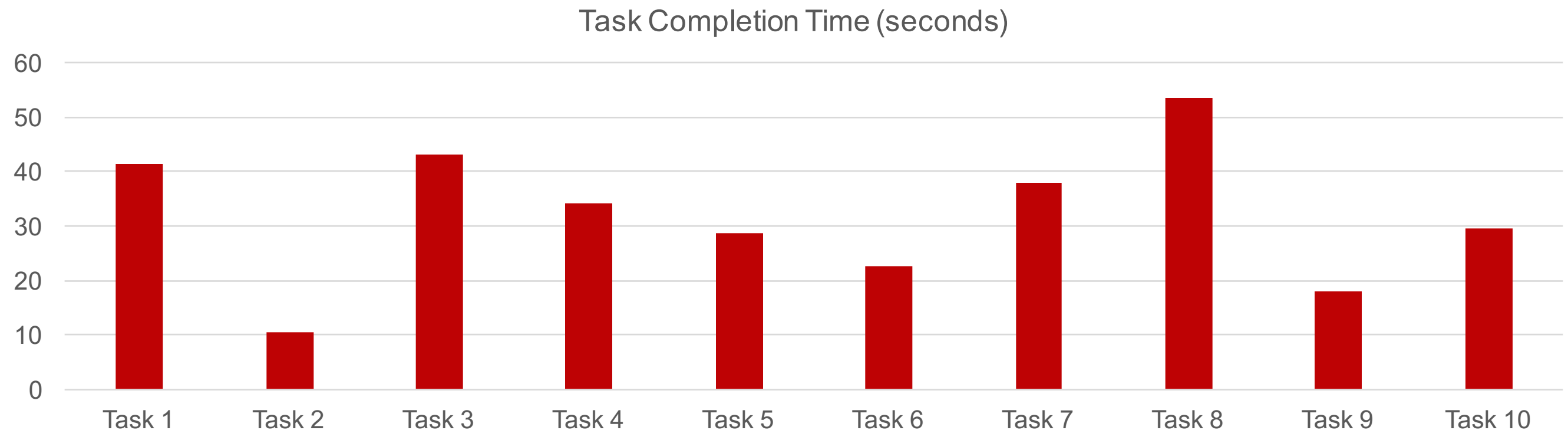


Evaluation – Case Studies

- Two case studies using the **NBA** and **Numbeo** quality-of-life data
- NBA 2010 - 2011 regular season statistics
 - 452 players and 12 numerical attributes
- Numbeo quality-of-life data
 - 176 cities and 8 numerical attributes

Evaluation – User Study

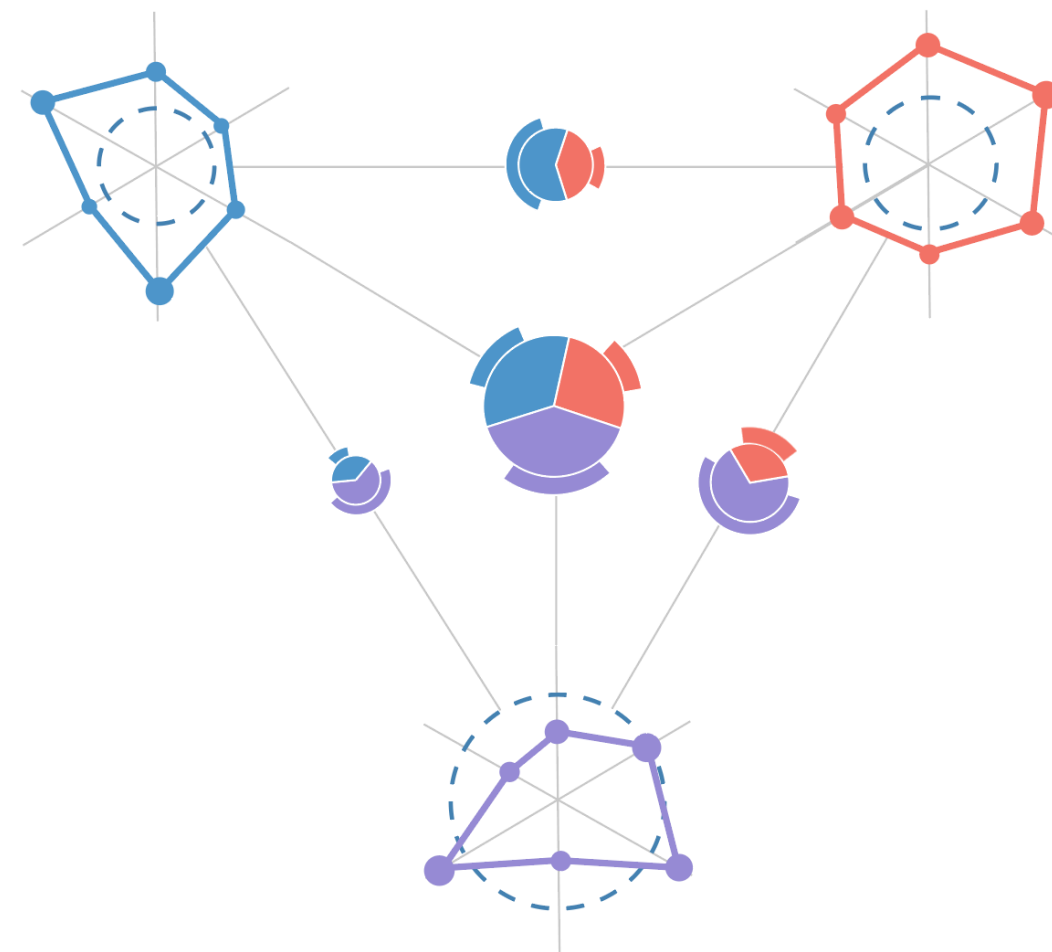
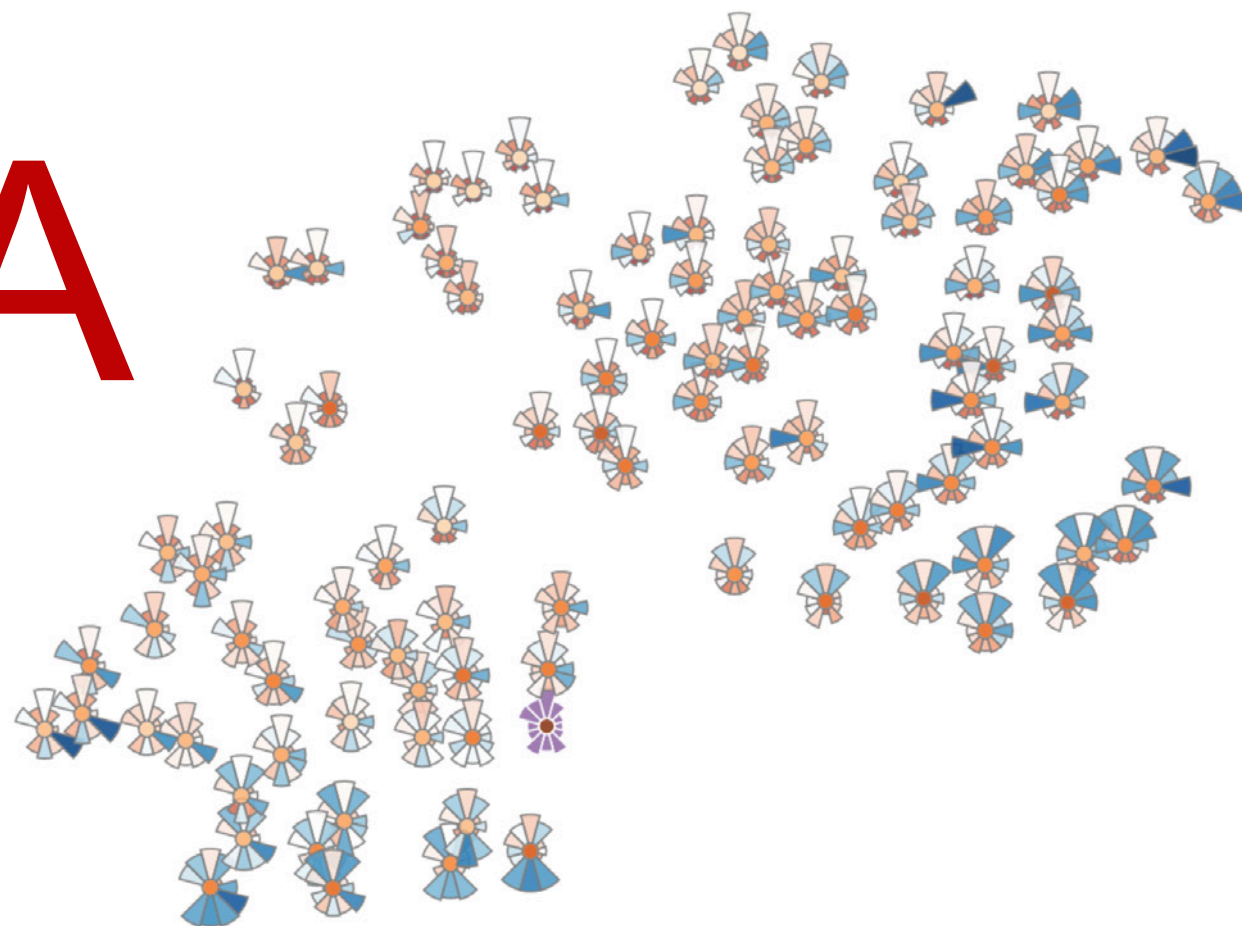
- Qualitative user study
 - 12 participants recruited from the local university
 - 10 tasks covering all important aspects in skyline analysis
 - 19 questions related with SkyLens usage in a post-session interview



Future Work

- Include **nominal attribute** analysis
- Support data with **uncertain values**
- Track **temporal changes** of skyline

Q&A



SkyLens: Visual Analysis of Skyline on Multi-dimensional Data

Xun Zhao: <http://zhaoxun.me/>

Contact: xzhaoag@ust.hk

Project page: <http://zhaoxun.me/skylens>