



On Skyline Groups

Chengkai Li
Naeemul Hassan
Gautam Das

University of Texas at Arlington

Nan Zhang
Sundaresan Rajasekaran

George Washington University



UNIVERSITY OF
TEXAS
ARLINGTON



معهد قطر لبحوث الحوسبة
Qatar Computing Research Institute

Motivation

Question-Answer Platforms

[Questions](#)[Tags](#)[Users](#)[Badges](#)[Unanswered](#)

Question

What is the difference between an int and an Integer in Java and C#?

92

I was just sitting at my local [Borders](#) sipping coffee and reading [More Joel on Software](#) (for free) when I came across [Joel Spolsky](#) saying something about a particular type of programmer knowing the difference between an int and an Integer in Java/C# (Object Oriented Programming Languages).

After a quick 'brain check,' I realized, to my dismay, that I didn't know the answer.

Skills

31

[c#](#) [java](#) [integer](#) [int](#)

Expert's Name	c#	java	int	integer	Total Weights
Eric Lippert	68	3	43	0	114
Bozho	0	51	52	0	103
Marc Gravell	100	2	0	0	102
BalusC	0	100	0	0	100
Ben Hoffstein	0	0	0	100	100
DannyT	0	0	100	0	100
Matthieu M.	0	0	87	0	87
Rohit Jain	0	18	66	0	84
dasblinkenlight	11	15	46	0	72
arshajii	0	10	56	0	66

Showing 1 to 10 of 4,654 entries

Previous Next

Goal: Find a group of experts who can answer this question

Motivation

Journal/Paper Review



Keywords (7)

Data Exploration
Experimental Evaluation
Faceted Search
Information Discovery
Internal Structure
Semantic Information
User Study

Skills

Academic > Publications > Facetedpedia: dynamic generation of query-dependent faceted interfaces for wikipedia

 Subscribe

 **Facetedpedia: dynamic generation of query-dependent faceted interfaces for wikipedia** (Citations: 3)  Export

Task

 Edit

Chengkai Li, Ning Yan, Senjuti Basu Roy, Lekhendo Lisham, Gautam Das

This paper proposes Facetedpedia, a faceted retrieval system for **information discovery** and exploration in Wikipedia. Given the set of Wikipedia articles resulting from a keyword query, Facetedpedia generates a faceted interface for navigating the result articles.

Compared with other faceted retrieval systems, Facetedpedia is fully automatic and dynamic in both facet generation and hierarchy

construction, and the facets are based on the collaborative vocabulary in Wikipedia (category system). Given the sheer size and

Expert's Name	Experimental Evaluation	Semantic Information	User Study	Total Weights
Wolfram Burgard	4	100	0	104
Patrick Baudisch	2	0	100	102
Wei-ying Ma	11	33	58	102
Hans-Peter Kriegel	100	0	0	100
Martha Palmer	0	100	0	100
Juan-zi Li	0	100	0	100
Mary Czerwinski	0	0	88	88
Yong Yu	4	67	12	83
Wolfgang Nejdl	13	50	12	75
Pearl Pu	0	17	58	75

Showing 1 to 10 of 14,522 entries

Goal: Find a group of experts who can review this paper

Motivation Fantasy Games

ESPN FANTASY BASKETBALL STANDARD EDITION

my ESPN NFL MLB NBA NHL NCAAF NCAAM NASCAR SOCCER MORE SPORTS WATCH FANTASY & GAMES espnW & X GAMES RADIO & MORE

Fantasy Basketball Create or Join a League Live Draft Lobby Mock Draft Lobby Player News Player Rater Daily Leaders FAQ More

ESPN » Free Fantasy Basketball

Daily Leaders

Position: ALL | PG | SG | SF | PF | C | G | F | UTIL

Last Name: Go Team: All

View: 2015 Season

PLAYERS	2015 SEASON STATS										
	MIN	FGM/FGA	FG%	FTM/FTA	FT%	3PM	REB	AST	STL	BLK	PTS
Kobe Bryant, LAL SG, SF											
Carmelo Anthony, NY SF, PF	35.5	8.9/18.8	.473	4.4/5.8	.760	1.5	5.6	3.5	0.6	0.2	23.8
James Harden, Hou SG	37.0	6.5/16.8	.386	8.9/10.0	.892	1.8	6.3	6.9	1.7	1.0	23.8
LeBron James, Cle SF, PF	38.6	9.1/19.2	.474	5.8/7.3	.795	1.9	6.6	6.8	1.1	0.8	25.9
Anthony Davis, Nor PF, C	36.7	10.3/17.8	.579	4.9/6.4	.766	0.0	11.4	2.0	2.3	3.9	25.5
Al Jefferson, Cha C	33.4	9.0/18.5	.486	2.7/4.3	.627	0.0	6.8	1.9	0.6	1.3	20.7
Stephen Curry, GS PG, SG	33.5	8.6/17.5	.491	4.5/4.7	.957	3.1	5.5	7.7	2.3	0.0	24.8
DeMarcus Cousins, Sac C, PF DTD	30.5	8.1/16.0	.506	6.4/8.2	.778	0.0	11.6	1.7	1.1	1.5	22.5
Tobias Harris, Ori PF, SF	36.3	7.2/14.8	.482	3.5/4.4	.807	1.1	7.8	1.8	1.2	0.6	18.9
Nikola Vucevic, Ori C, PF	36.3	8.2/15.3	.538	2.2/2.8	.757	0.2	11.9	2.6	0.8	1.1	18.8
Rudy Gay, Sac SF DTD	36.9	7.3/16.5	.440	6.3/7.3	.863	1.0	6.5	3.6	1.2	0.5	21.8
Monta Ellis, Dal SG, PG	32.1	7.6/15.8	.481	3.6/4.4	.811	0.9	2.4	4.2	1.2	0.3	19.7
DeMar DeRozan, Tor SG, SF	34.4	7.2/17.2	.418	6.6/8.1	.820	0.2	4.3	2.6	1.9	0.1	21.2

Skills

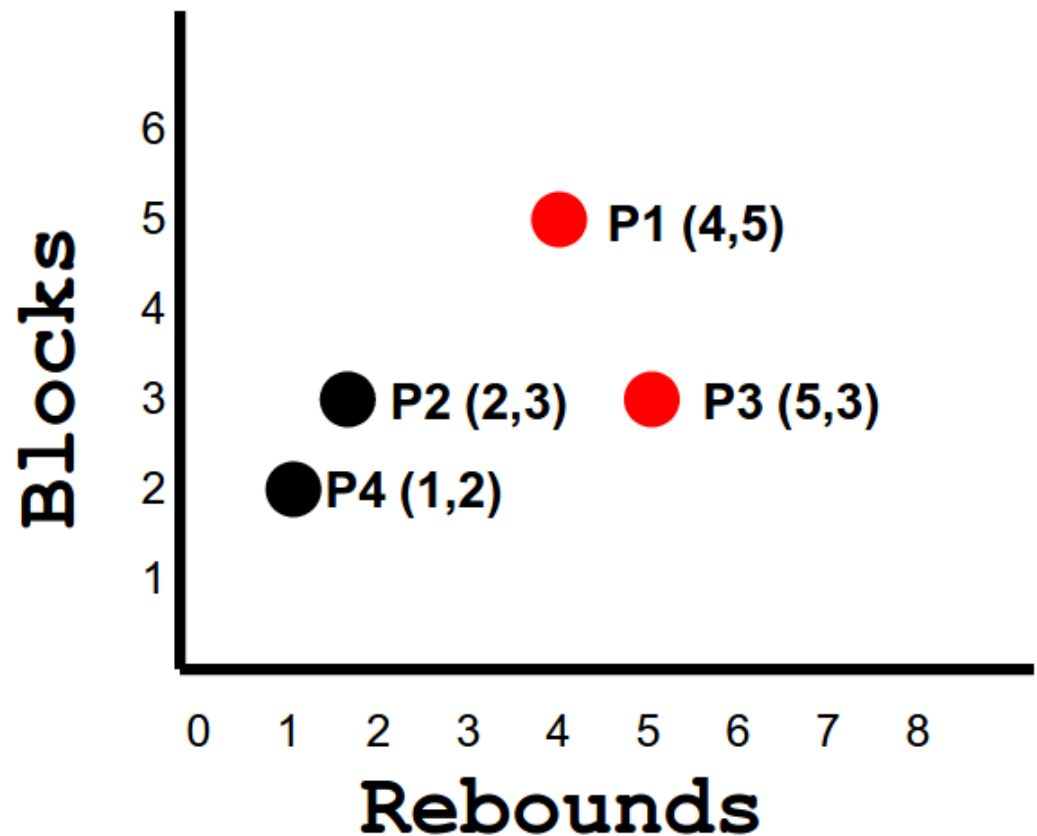
Goal: Find a group of players for Fantasy Basketball

Problem Definition

What is Skyline?

NBA Players Score

	Rebounds	Blocks
P1	4	5
P2	2	3
P3	5	3
P4	1	2
P5	1	2



Skyline tuple: A tuple which is not dominated by any other tuple.

Problem Definition

What is Skyline Group?

NBA Players Score

	Points	Rebounds	Blocks
P1	3	4	5
P2	4	2	3
P3	4	5	3
P4	2	1	2
P5	4	1	2

Skyline Players

Find a group of 3 players

5 Choose 3 = 10 possible groups

Skyline Groups

	SUM			MIN			MAX		
	P	R	B	P	R	B	P	R	B
P1, P2, P3	11	11	11	3	2	3	4	5	5
P1, P2, P4	9	7	10	2	1	2	4	4	5
P1, P2, P5	11	7	10	3	1	2	4	4	5
P1, P3, P4	9	10	10	2	1	2	4	5	5
P1, P3, P5	11	10	10	3	1	2	4	5	5
P1, P4, P5	9	6	9	2	1	2	4	4	5
P2, P3, P4	10	8	8	2	1	2	4	5	3
P2, P3, P5	12	8	8	4	1	2	4	5	3
P2, P4, P5	10	4	7	2	1	2	4	2	3
P3, P4, P5	10	7	7	2	1	2	4	5	3

Problem Definition

Why Skyline Group?

NBA Players Score

	Points	Rebounds	Blocks
P1	3	4	5
P2	4	2	3
P3	4	5	3
P4	2	1	2
P5	4	1	2

What's wrong with taking
most expert in each field?

Any other group is
dominated by a Skyline

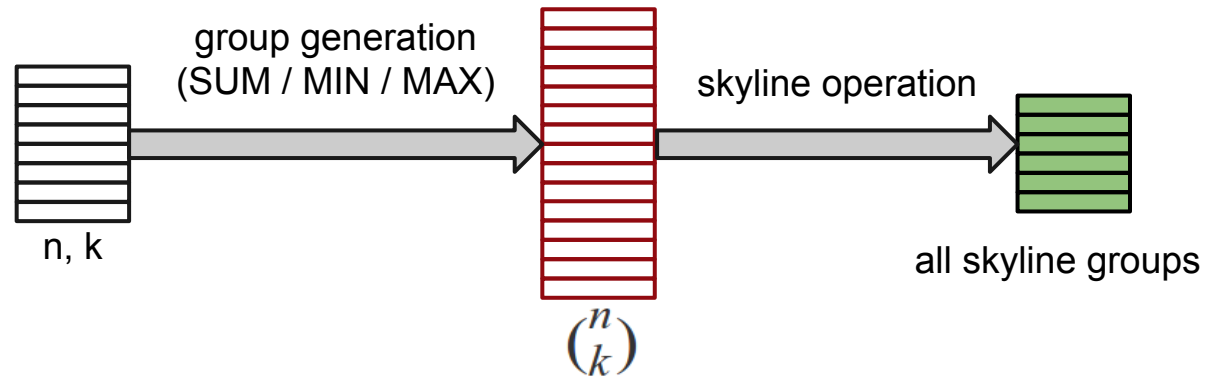
	SUM			MIN			MAX		
	P	R	B	P	R	B	P	R	B
P1, P2, P3	11	11	11	3	2	3	4	5	5
P1, P2, P4	9	7	10	2	1	2	4	4	5
P1, P2, P5	11	7	10	3	1	2	4	4	5
P1, P3, P4	9	10	10	2	1	2	4	5	5
P1, P3, P5	11	10	10	3	1	2	4	5	5
P1, P4, P5	9	6	9	2	1	2	4	4	5
P2, P3, P4	10	8	8	2	1	2	4	5	3
P2, P3, P5	12	8	8	4	1	2	4	5	3
P2, P4, P5	10	4	7	2	1	2	4	2	3
P3, P4, P5	10	7	7	2	1	2	4	5	3

Solution Framework

Baseline Method

Input

- n players/tuples
- group size k
- aggregate function (sum/min/max)



Problems

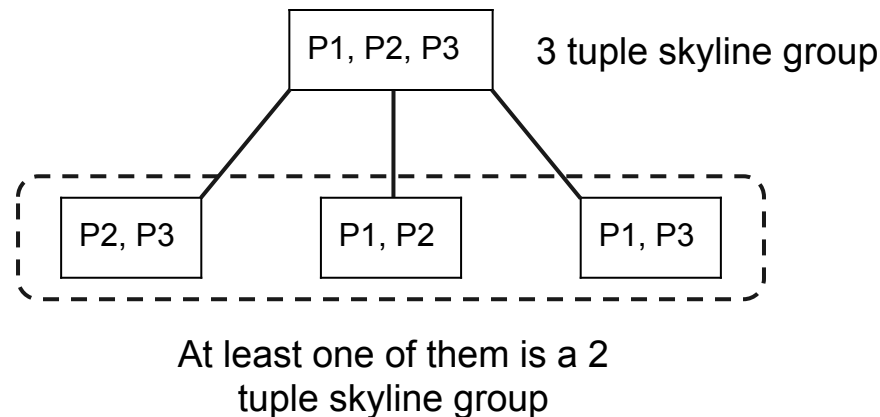
- Exponential group generation. We may not afford to compute or store them.
 - Example: For $n = 2000$, $k = 3$.
 - 1331334000 groups
 - 30 GB space [assuming 24B for each group]
 - 15 days time [assuming 1 millisecond for each group]

Solution Framework

Advanced Method: WCM

Weak Candidate Generation Property: If G is a k tuple skyline group, then there is **at least one $(k-1)$ tuple subset of G** such that it is a $(k-1)$ tuple skyline group.

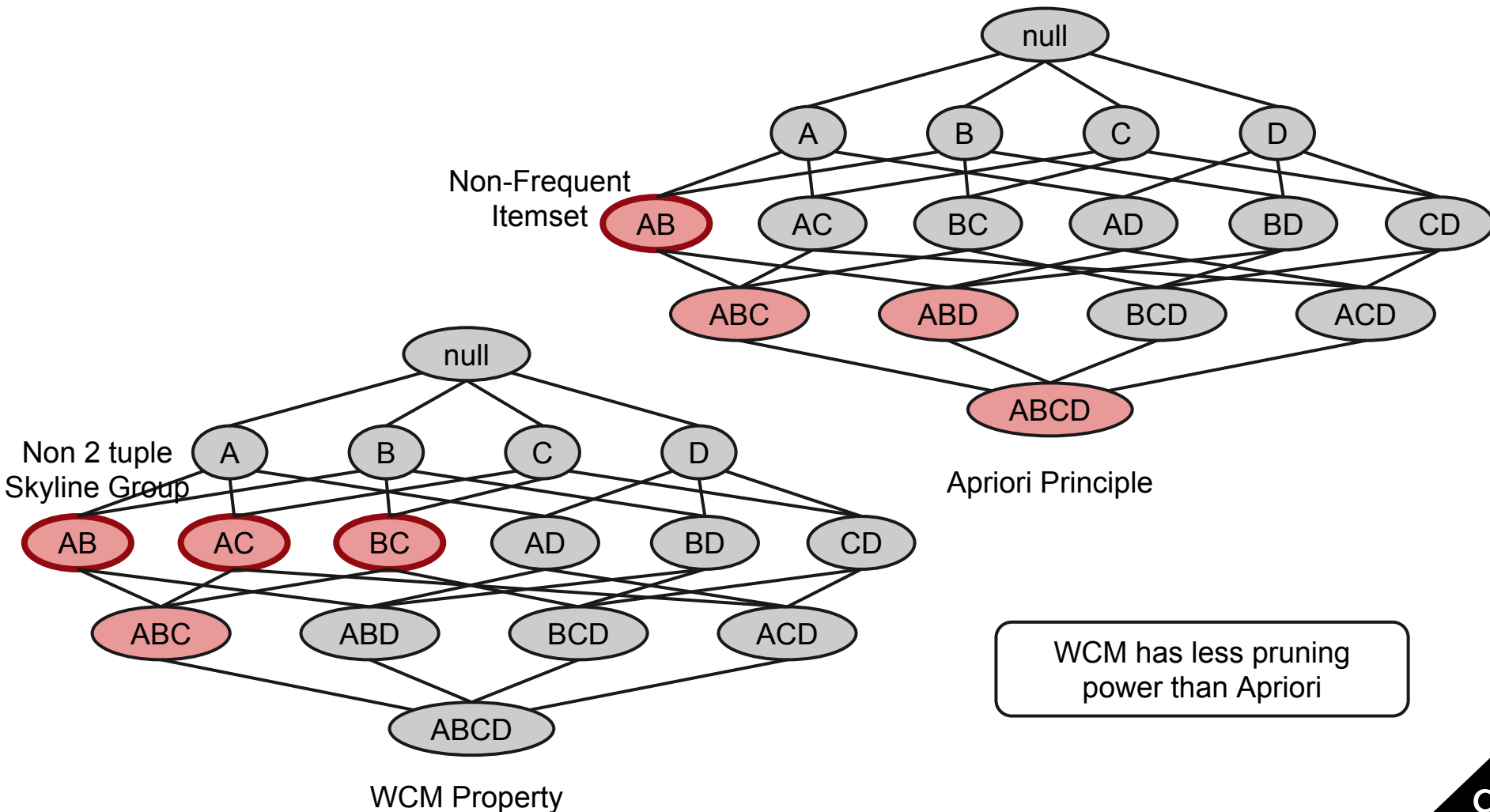
Example:



Does this property sound familiar?

Apriori Principle: If an itemset is frequent, then **all of its subsets** must also be frequent

Comparison Between Apriori & WCM Property



WCM Algorithm

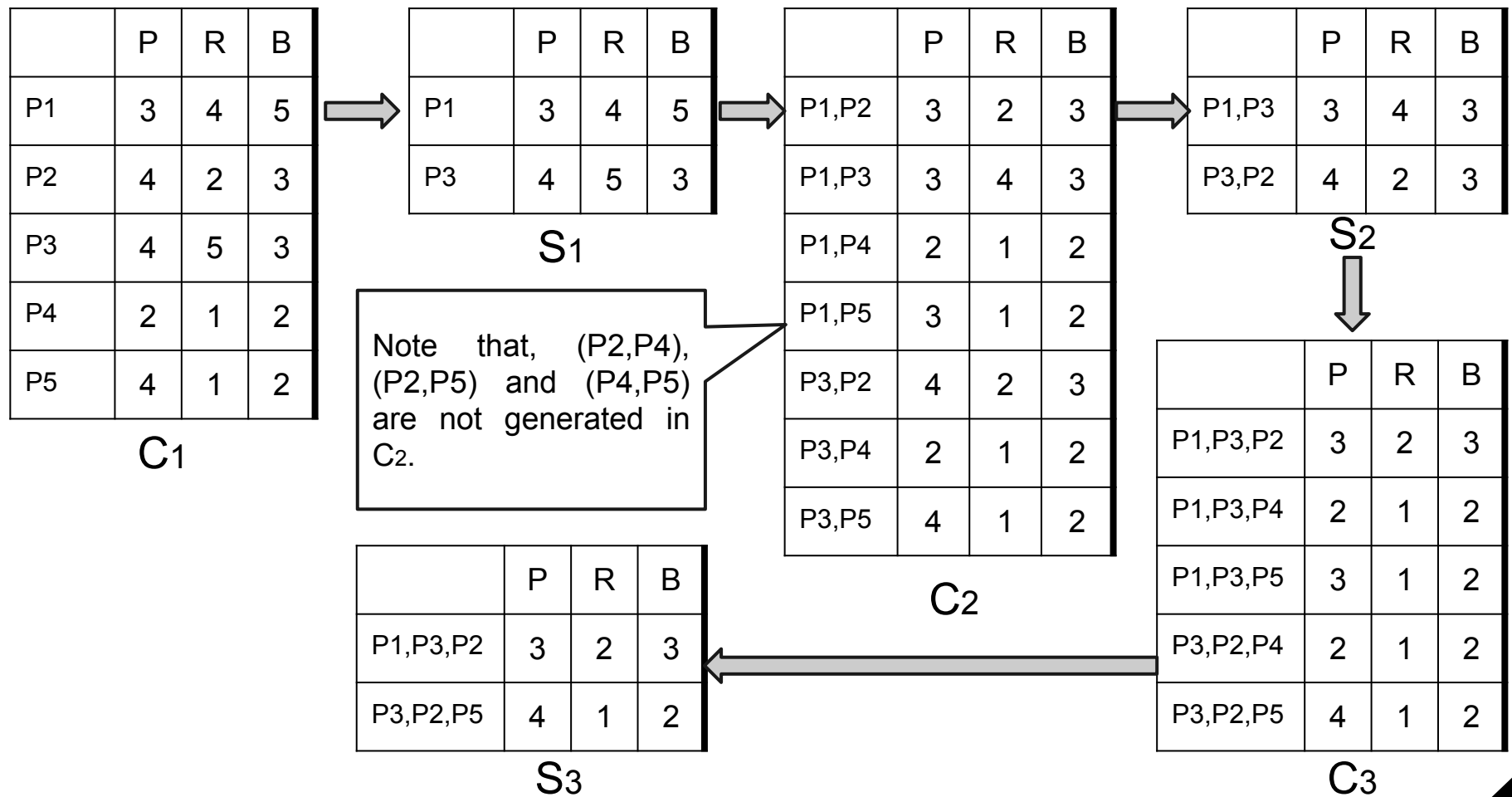
Input: n tuples, group size k , aggregate function = **min/max (not sum)**

1. Let, $i = 1$
2. Generate 1 tuple Candidate groups, $C_1 = \text{all } n \text{ tuples}$
3. Generate 1 tuple Skyline groups, $S_1 = \text{skyline_operation}(C_1)$
4. for $i = 2$ to k
 - a. Generate i tuple Candidate groups, C_i from S_{i-1}
 - b. Generate i tuple Skyline groups, $S_i = \text{skyline_operation}(C_i)$
5. Return S_k

WCM Algorithm

Explained with Example

Input: n tuple {P1, P2, P3, P4, P5}, group size k = 3, aggregate function = min





Question



CrewScout System

CrewScout

An Expert Team Finding Tool



Select Dataset: StackOverflow QA Go

Search

Total 5252 Available Tasks

Task ID:123. Java lib or app to convert CSV to XML file?

Task Detail

Task ID: 564

Title: What is the difference between an int and an Integer in Java and C#?

Description: 2008-08-02 21:47:34

Task ID:2158. Creating a custom button in Java

Task ID:2968. What are the different methods to parse strings in Java?

Required Skills

☒ c#
☒ int
☒ integer
☐ java

Show Experts

Aggregate Function

☒ AVG ☐ MIN ☐ MAX

Skyline Team Size: 3

Skyline Teams

Eligible Experts

Show 10 entries

Expert's Name	c#	int	integer	Total Weights
Eric Lippert	68	43	0	111
Marc Gravell	100	0	0	100
Ben Hoffstein	0	0	100	100
DannyT	0	100	0	100
Matthieu M.	0	87	0	87
Rohit Jain	0	66	0	66
dasblinkenlight	11	46	0	57
arshajii	0	56	0	56
Jean	0	53	0	53
Anthony Pegram	8	44	0	52

Showing 1 to 10 of 2,562 entries

Previous Next

[CONTACT](#)

<http://idir.uta.edu/crewscout>

Thank You!