



Санкт-Петербургский государственный университет
Кафедра системного программирования

Single, Multiple source BFS-parents с SPLA и Apache Giraph

Ахмедов Давид

Санкт-Петербург
2025

Определение (BFS-parents)

Пусть $G = (V, E)$ — ориентированный или неориентированный граф, $S \subseteq V$ — множество стартовых вершин. Алгоритм «поиска в ширину» строит массив

$$parent: V \longrightarrow V \cup \{\perp\},$$

где для каждой вершины v :

$$parent[v] = \begin{cases} \perp, & v \in S \text{ или } v \text{ недостижима из } S \\ u, & u \text{ — родитель вершины } v \end{cases}$$

BFS-parents с ЛА

semiring	set	\oplus	\otimes	0
min-select 1st	$a \in \mathbb{N} \cup \{+\infty\}$	min	sel1st	$+\infty$

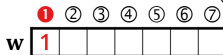
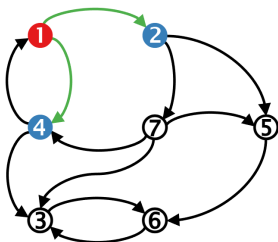
$\text{sel1st}(x, y) = x$

- 1: **Input:** A (adjacency matrix), s (sources)
- 2: **Output:** p (parent vector)
- 3: $x \leftarrow [1, 2, \dots, n]$
- 4: $w \leftarrow s$
- 5: $p \leftarrow \mathbf{0}_n$
- 6: **for** isUpdates **do**
- 7: $u \leftarrow \bar{p}$
- 8: $p\langle u \rangle \leftarrow w \cdot A$
- 9: $w\langle u \rangle \leftarrow p \odot x$
- 10: **end for**
- 11: **return** p

BFS – PARENTS

semiring	set	\oplus	\otimes	0
min-select 1st	$a \in \mathbb{N} \cup \{+\infty\}$	min	sel1st	$+\infty$

$$\text{sel1st}(x, y) = x$$



A	①	②	③	④	⑤	⑥	⑦
①		●		●			
②					●		●
③						●	
④	●		●				
⑤						●	
⑥			●				
⑦			●	●	●		

w	①	②	③	④	⑤	⑥	⑦
	1						

$$p(\bar{p}) = w \min. \text{sel1st } A$$

↓ index

w'	①	②	③	④	⑤	⑥	⑦
		2		4			

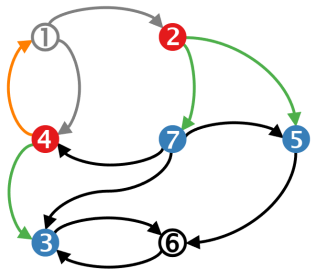
Рис.: Иллюстрация алгоритма¹

¹GraphBLAS: A linear algebraic approach for high-performance graph algorithms –

BFS – PARENTS

semiring	set	\oplus	\otimes	0
min-select 1st	$a \in \mathbb{N} \cup \{+\infty\}$	min	sel1st	$+\infty$

$\text{sel1st}(x,y) = x$



① ② ③ ④ ⑤ ⑥ ⑦

w	2		4			
---	---	--	---	--	--	--

A

	①	②	③	④	⑤	⑥	⑦
①		●		●			
②					●		●
③						●	
④	●		●				
⑤						●	
⑥			●				
⑦			●	●	●		

	①	②	③	④	⑤	⑥	⑦
w	0	1	4	1	2		2

$p(\bar{p}) = w \text{ min. sel1st } A$

↓ index

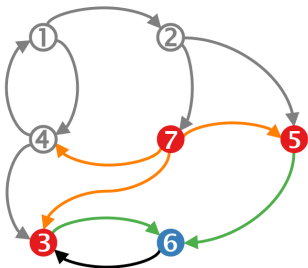
w'

	①	②	③	④	⑤	⑥	⑦
			3		5		7

BFS – PARENTS

semiring	set	\oplus	\otimes	0
min-select 1st	$a \in \mathbb{N} \cup \{+\infty\}$	min	sel1st	$+\infty$

$$\text{sel1st}(x, y) = x$$



A	①	②	③	④	⑤	⑥	⑦
①		●		●			
②					●		●
③						●	
④	●		●				
⑤						●	
⑥			●				
⑦			●	●	●		

⊗	1	4	1	2	3	2
---	---	---	---	---	---	---

$$p(\bar{p}) = w \text{ min. sel1st } A$$

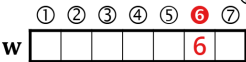
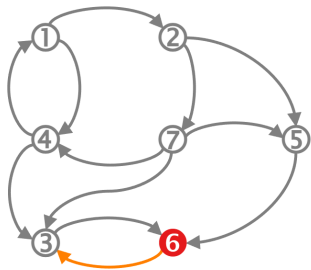
↓ index

①	②	③	④	⑤	⑥	⑦
w'					6	

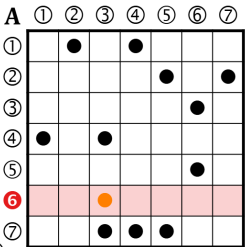
BFS – PARENTS

semiring	set	\oplus	\otimes	0
min-select 1st	$a \in \mathbb{N} \cup \{+\infty\}$	min	sel1st	$+\infty$

$\text{sel1st}(x, y) = x$

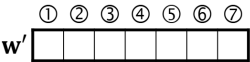


w' is empty
→ terminate



$p(\bar{p}) = w \text{ min. sel1st } A$

↓ index



Экспериментальный набор данных

Таблица: Статистика графов SNAP²

Граф (название в SNAP К.К.С)	Вершин	Ребер	К.С.	Средн. степ. в К.К.С	Макс. степ. в К.К.С	Крупнейшая К.С.
Twitch gamers network (large_twitch_edges)	168.1K	6.8M	1	80.86	35.2K	168.1K (100%)
Gemsec Facebook dataset (gemsec_facebook_artist_edges)	134.8K	1.4M	8	32.43	1.4K	50.5K (37.4 %)
Gemsec Deezer dataset (gemsec_deezer_HR_edges)	143.9K	846.9K	3	18.25	420	54.5K (37.9%)
Twitch social networks (musae_twitch_DE_edges)	34.1K	429.1K	6	32.24	4.2K	9.4K (27.5%)
Github developer network (musae_git_edges)	37.7K	289.0K	1	15.33	9.5K	37.7K (100%)
Facebook page-page network (musae_facebook_edges)	22.5K	171.0K	1	15.22	709	22.5K (100%)
Facebook social circles (facebook_combined)	4.0K	88.2K	1	43.69	1.0K	4.0K (100%)
LastFM Asia social network (lastfm_asia_edges)	7.6K	27.8K	1	7.29	216	7.6K (100%)

²<http://snap.stanford.edu/data> – Stanford Large Network Dataset Collection

RQ:

- (ms) Какой из алгоритмов лучше масштабируется при увеличении числа стартовых вершин? Выбор 3, 30, 100 стартовых вершин (вершины выбираются случайно с фиксированным сидом, из K.K.C.).
- (ss) Сравнение алгоритмов с использованием SPLA vs SuiteSparse:GraphBLAS: время выполнения
- Сравнение алгоритмов с использованием SPLA vs SuiteSparse:GraphBLAS: потребление памяти (мониторинг пикового потребления за все итерации)

Ход эксперимента: Метод подынтервалов для сбора статистики (N вершин, M запусков), IQR для борьбы с выбросами

Характеристики тестового оборудования

- AMD Ryzen 5 5500U 2.1MHz @ AMD Radeon RX Vega 7 @ ROCm
 - ▶ 6 ядер
 - ▶ все ядра равнозначны
 - ▶ без гипертрединга
- RAM 16 GB @ 3200 MHz
 - ▶ без свопа
- Ubuntu 22.04.5 x86_64