

Реализация и анализ алгоритма разбиения гиперграфа с использованием алгоритма Фидучча-Матейсис

<https://github.com/pr0kuk/F-M-Alg>

Таблица результатов работы (см. res.xlsx)

Абсолютные характеристики

Graph	Edges	Vertices	Cutset	Time, ms	Iterations	Cutset	Time, ms	Iterations
ISPD98_ibm01.hgr	14112	12753	489	548	10	2065	531	10
ISPD98_ibm02.hgr	19585	19602	572	972	10	1328	1067	11
ISPD98_ibm03.hgr	27402	23137	2298	2038	20	4459	2280	23
ISPD98_ibm04.hgr	31971	27508	950	1927	16	5116	3772	32
ISPD98_ibm05.hgr	28447	29348	3610	3665	28	6948	1707	14
ISPD98_ibm06.hgr	34827	32499	1482	3511	25	5727	4494	33
ISPD98_ibm07.hgr	48118	45927	2161	3987	20	8394	7113	34
ISPD98_ibm08.hgr	50514	51310	4241	3906	16	9280	5204	21
ISPD98_ibm09.hgr	60903	53396	2853	4321	17	10092	10080	40
ISPD98_ibm10.hgr	75197	69430	2688	6537	19	13093	12433	38
ISPD98_ibm11.hgr	81455	70559	7993	6088	19	14952	5561	18
ISPD98_ibm12.hgr	77241	71077	4271	6463	18	15707	9182	27
ISPD98_ibm13.hgr	99667	84200	3870	5751	14	17424	7077	18
ISPD98_ibm14.hgr	152773	147606	10394	18521	27	24712	22657	33
ISPD98_ibm15.hgr	186609	161571	10063	13212	15	31863	41869	48
ISPD98_ibm16.hgr	190049	183485	5786	12927	13	36399	20593	21
ISPD98_ibm17.hgr	189582	185496	9378	35341	32	42347	37974	36
ISPD98_ibm18.hgr	201921	210614	4577	43818	41	35358	42306	40
dac2012_superblue2.h	990900	1010322	29852	133561	18	49154	151875	21
dac2012_superblue3.h	898002	917945	17913	170047	18	22142	93244	10
dac2012_superblue6.h	1006630	1011663	21621	121958	13	22153	160335	17
dac2012_superblue7.h	1340419	1360218	48057	269748	17	58706	236880	15
dac2012_superblue9.h	833809	844333	15522	158568	10	30608	126735	8
dac2012_superblue11.	935732	952508	14726	71649	11	25797	131312	20
dac2012_superblue12.	1293437	1291932	22786	587683	30	38189	409555	20
dac2012_superblue14.	619816	630803	15579	144896	21	25174	87642	13
dac2012_superblue16.	697459	698340	20590	90750	21	34093	48975	11
dac2012_superblue19.	511686	522483	11904	266974	22	18646	178417	15
			Base algorithm			Modified algorithm		

Отношения характеристик модифицированного / базового

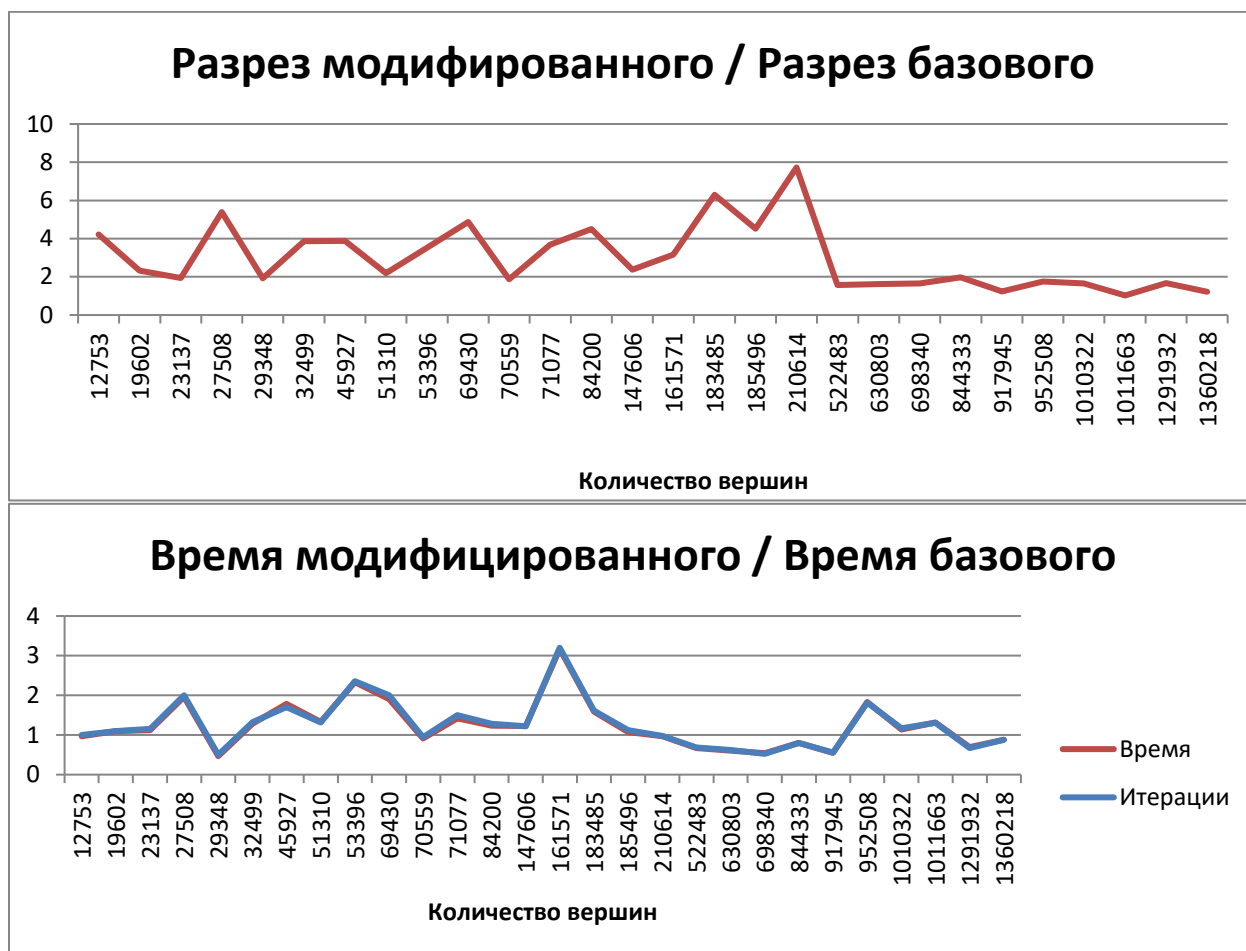
Вершины	Разрез	Время	Итерации		Разрез	Время	Итерации
12753	4,222904	0,968978	1	Лучшее	1,024606	0,465757	0,5
19602	2,321678	1,097737	1,1	Худшее	7,725147	3,169013	3,2
23137	1,940383	1,118744	1,15	Среднее	2,986879	1,244634	1,261733
27508	5,385263	1,957447	2	СтдОткл	1,698705	0,606609	0,612556
29348	1,924654	0,465757	0,5				
32499	3,864372	1,279977	1,32				
45927	3,884313	1,784048	1,7				
51310	2,188163	1,332309	1,3125				
53396	3,537329	2,332793	2,352941				
69430	4,870908	1,901943	2				
70559	1,870637	0,913436	0,947368				
71077	3,677593	1,420702	1,5				
84200	4,502326	1,230569	1,285714				
147606	2,377525	1,223314	1,222222				
161571	3,166352	3,169013	3,2				
183485	6,290875	1,593022	1,615385				
185496	4,515568	1,074503	1,125				
210614	7,725147	0,965494	0,97561				
522483	1,566364	0,668294	0,681818				
630803	1,615893	0,604861	0,619048				
698340	1,655804	0,539669	0,52381				
844333	1,971911	0,799247	0,8				
917945	1,236086	0,548343	0,555556				
952508	1,7518	1,832712	1,818182				
1010322	1,64659	1,137121	1,166667				
1011663	1,024606	1,314674	1,307692				
1291932	1,675985	0,696898	0,666667				
1360218	1,221591	0,878153	0,882353				

Визуализация результатов работы

Абсолютные характеристики



Отношения характеристик модифицированного / базового



Начальное распределение

Если $v < |V|/2$: $F(v) = 0$

Если $v \geq |V|/2$: $F(v) = 1$

Алгоритм очень чувствителен к начальному распределению, поэтому для корректных исследований требуется более тщательно подходить к его выбору, используя градиентный спуск или среднее по выборке из нескольких случайных инициализаций, однако в данной работе исследуется лишь корректность работы алгоритма и влияние незначительных модификаций, поэтому для удобства отладки было выбрано тривиальное распределение.

Описание модифицированной версии алгоритма.

В качестве модификации исследуется влияние на итоговый результат алгоритм выбора вершины из элементов `gain_container` с одинаковым гейном. А именно после перемещения одной вершины, все вершины с обновлённым гейном помещаются в конец соответствующих списков, а не в начало.

Сравнительный анализ алгоритмов

Добавление вершины в начало списка вершин с данным гейном концептуально означает, что преимущественно алгоритм будет работать с вершинами, которые принадлежат рёбрам, раскраска которых уже была изменена алгоритмом. Добавление вершин в конец, напротив приводит к преимущественному изменению раскрасок различных рёбер. Считая, что распределение вершин в рёбрах равновероятное (в среднем для любого ребра число вершин с номером <

$|V|/2$ равно числу вершин с номером >

$|V|/2$) и используя вышеописанное

начальное распределение, можно

предположить, что в случае

преимущественного изменения раскрасок

одних и тех же рёбер финальный разрез

будет меняться скачкообразно, что позволит

избежать локальных максимумов.

Изменение раскрасок большего числа рёбер

потенциально должно приводить к более

плавному и долгому схождению, что в

большинстве случаев приводит к ложному

локальному минимуму. Плавное изменение

разреза означает увеличение требуемого

числа итераций для сходимости, что в свою

очередь приводит к увеличению времени

работы. На описанных выше бенчмарках

использование модификации приводит к

увеличению времени работы в среднем в

1.24 раза и увеличение разреза в среднем в

2.99 раза.

Кроме того, можно привести некоторые

эмпирические закономерности, такие как

увеличение времени работы и финального

разреза при увеличении размера графа.

Количество требуемых для схождения

итераций при этом можно считать

независимым или слабо зависимым от

размера графа и колеблющимся в пределах

10-50.