Turbo charging heuristics: adjusting the parameters for optimum performance. (Talk 2)

Faisal N. Abu-Khzam[1,2] Shaowei Cai[3] Judith Egan[1] **Peter Shaw[4]** Kai Wang[1]

Charles Darwin University AU

Lebanese American University, Beirut

Massey University, Manawatu

Chinese Academy of Sciences, Beijing

July 2018

Abstract

Turbo-charging is a recent algorithmic technique for hard problems that employs an FPT subroutine as part of a heuristic. We demonstrate the effectiveness of this technique and develop the turbo-charging idea further. In this talk we will explore how the performance can be improved through adjusting the parameters and moment-of-regret function.
We implement both the initially proposed “turbo-greedy” algorithm of Downey et al. and a new hybrid heuristic for the $W[2]$-hard Dominating Set problem and evaluated their performance for a range of parameters and datasets. Our algorithm often produced results that were either exact or better than all the other available heuristic algorithms. The results vary depending on the parameter, with the best results obtained for larger values of k and r.

Overview

Motivation

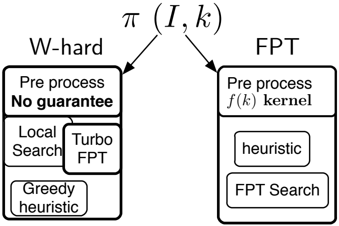


Figure: The Dynamic FPT formulations of some W-hard problems are in FPT.

- FPT
  - $T(n,k) = 2^{O(k)} \cdot n^c$
- W Hierarchy
  - FPT $\subseteq W[1] \subseteq W[2] \subseteq \cdots XP$
- Ideally, Turbo Charging provides a $\alpha^{f(r,k)}n$ heuristics with an exchange neighborhood of r.

Definition

DOMINATING SET (DS)

Instance: A graph $G = (V, E)$

Parameter: k

Question: Does G have a dominating set $S \subseteq V$, $|S| \leq k$, such that every vertex in $V \setminus S$ is adjacent to a vertex in S ?

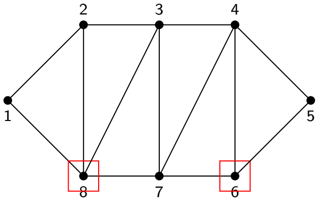


Figure: Example of DOMINATING SET (DS)

Terminology

- *open/closed neighborhood*
  - $N_G(S) = \cup_{v \in S} N(v)$
  - $N_G[S] = N_G(S) \cup S$
- Measure
  - $utility(v)$ = the number of non-dominated neighbors of the vertex v
  - $vote(v) = utility(v) + \frac{1}{\sum_{u \in N(v)} utility(u)}$ [10]
- Edit Distance
  - $d_e(G, G') = \left|\left(E(G) \setminus E(G')\right) \cup \left(E(G') \setminus E(G)\right)\right|$
  - $d_v(D, D') = |(D \setminus D') \cup (D' \setminus D)|$

Greedy Heuristics

- Greedy Chvátal [3]
- Greedy Vote [10]
- Greedy Vote GRASP [10]

Dynamic FPT Heuristics

- FPT TURBO I [4]

Definition

Dynamic Dominating Set (DDS)

Instance: A graph G and a graph G' with $d_e(G, G') \leq k$, a dominating set $S \subseteq V(G)$; a positive integer r.

Parameter: (k, r)

Question: Does there exist a set of vertices $S' \subseteq V(G') = V(G)$ with $d_v(S, S') \leq r$ and with S' being a dominating set of G'?

Definition

Greedy Improvement of Dominating Set (Greedy-DS) [4]

Instance: A graph G; a list L of the vertices of G ordered from highest degree to lowest degree, $L = (v_1, \ldots, v_l, \ldots, v_{l+k} = v, \ldots, v_n)$; a set of vertices $D \subseteq V(G)$ that dominates the set $V' = \{v_1, \ldots, v_{l+k}\}$; a partition $D = D_1 \cup D_2$ where D_1 dominates the set of vertices $\{v_1, \ldots, v_l\}$ and $|D_2| \leq k$

Parameter: k

Question: Is there a set of vertices $D' \subseteq V(G)$ such that D' dominates the vertices of $V' \cup D$, with $D_1 \subseteq D'$ and $|D'| < |D|$?

Greedy Improvement of Dominating Set (Greedy-DS)

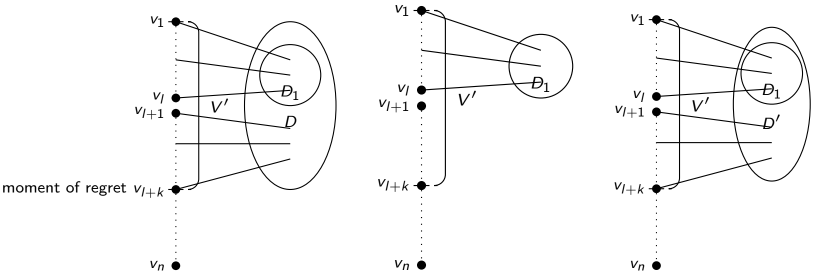


Figure: Illustration of Greedy-DS

Greedy DDS

Require: Two graphs G and G', parameters k and r, and a dominating set solution S for G.

```
1: Divide graph G' into three parts: S, B, and C, where
2: C ← V(G') \ N_{G'}[S]
3: B ← N_{G'}[S] \ S;
4: Apply the reduction rules [4] on G' to get a reduced instance G* with B* and C; ▷ C is a vertex cover for G* and in turn
   B* is an independent set for G*
5: if B* and C are the same as history record, return S;
6: Traverse vertices in B* and C to get a vertex set P which presents different neighbor types of C; ▷ |P| ≤ slant2^k because
   |C| ≤ slantk
7: apply rule (R??) on the neighbor types of C so as to remove impossible vertices in P;
8: findSolution ← false;
9: do
10:     Choose r vertices from P to construct a set D;
11:     if D dominates C then
12:         findSolution ← true;
13:         goto Line 19
14:     end if
15: while (all combinations of r vertices are not tried)
16: if findSolution = false then
17:     D ← C;
18: end if
19: S' ← S ∪ D
20: Return S' as the dominating set solution of of graph G';
```

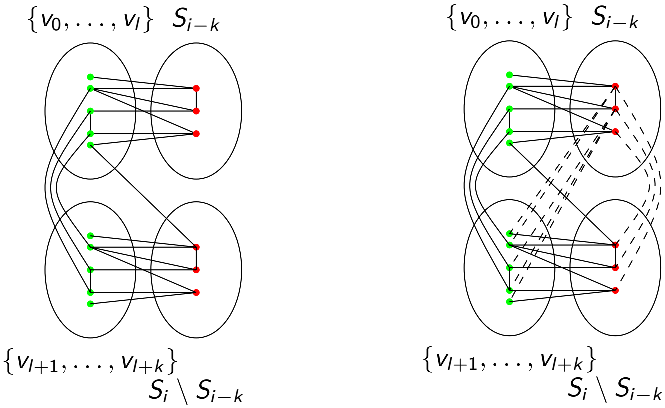
Developing a new hybrid algorithm

A new FPT TURBO HYBRID algorithm.

- Connected Components;
- Applying reduction rules;
- *U*sing alternative measure(s) for selecting solution elements;
- Using dynamic re-ordering of vertex list on the measures;
- Checking whether the solution(s) obtained are minimal;
- *A*pplying an appropriate LS heuristic;
- Adding a *heuristic guarantee*;
- Specifying an appropriate `Moment_of_Regret` function

Hybrid DDS

Input: a graph $G = (V, E)$; parameters k and $rupper$ such that $rupper = k - 1$ (initially)
Rank a list L of vertices in G from lowest to highest utility;
Get the vertex v_0 of the lowest utility;
Get the highest utility vertex $u_0 \in N_G[v_0]$;
$S_0 \leftarrow \{u_0\}$;
Initialize the graph G_0 with $\{u_0, v_0\}$ and the edge between u_0 and v_0;
$i \leftarrow 0$;
do
  $i \leftarrow i + 1$;
  Rank the list L of the vertices(Vote or utility);
  if v_i is dominated by S_{i-1} **then**
    $G_i \leftarrow G_{i-1} \cup \{v_i\}$;
    $S_i \leftarrow S_{i-1}$;
  else
    Get the highest utility vertex $u_i \in N_G[v_i]$;
    $S_i \leftarrow S_{i-1} \cup \{u_i\}$;
    Construct G_i from G_{i-1} with $\{v_i, u_i\}$ and incident edges in G ;
    if is_moment_of_regret(G_i,u_i,S_i,S_{i-k}) **then**
      $r \leftarrow \min(r_{uppper}, |S_i| - |S_{i-k}| - 1)$;
      $\hat{G}_i \leftarrow$ a virtually constructed graph from G_i by adding $\leq 2k$ edges between S_{i-k} and $V(G_i) \setminus N_{G_i}[S_{i-k}]$;
      $S'_i \leftarrow \mathrm{DDS\ FPT}\ (G = \hat{G}_i, G' = G_i, S = S_{i-k},\ k = |V(G_i) \setminus V(G_{i-k})|, r = r)$;
      $S_i \leftarrow min(S_i, S'_i)$ ▷ Get the minimum size set
    end if
  end if
while (Not all the vertices are dominated);
Return the final S_i as the dominating set solution for G;



(a) G_{i-k} (b) $\hat{G}_i$

Figure: $G_i \longrightarrow \hat{G}_i$

- any dashed-edges are removed by reduction rule so don't effect k

A new Moment-of-regret function

- In the new `is_moment_of_regret` we allow $(|S_i| - |S_{i-k}|) \geqslant MOR_{threshold}$.
- Each time the moment-of-regret happens the parameter value of r can vary
- So r for each time should be the minimum among
  1. $(|S_i| - |S_{i-k}| - 1)$;
  2. at least one less than $(|S^t{}_i| - |S_{i-k}| - 1)$, $S^t{}_i$ is the solution size of heuristic guarantee;
  3. the input argument r_{upper}. (New input value)

Performance on the KONECT data set

All algorithms presented in this paper are implemented in the Java programming language. The experiments are run on a computer of the OSX Yosemite operating system with a CPU of 3.5 GHz, 6-Core Intel Xeon E5, and 32 GB memory.

Exact solutions shown using Fomin et al. [5], implemented using the hybrid method of Abu-Khzam et al. [2] given

- KONECTS [1, 6]
- KONECTS (Massive graphs $|V| > 4900$)
- DIMACS (Complement) [8]
- BSHOLIB [**?**]

Performance on the KONECT data set

Data Set			Opt	Greedy Chvátal		Greedy Vote GRASP		FPT Turbo I		FPT **Turbo Hybrid**	
Data set name	$\lvert V \rvert$	$\lvert E \rvert$	**Size**	**Size**	**Time**	**Size**	**Time**	**Size**	**Time**	**Size**	**Time**
Zebra	27	111	4	4	<0.01	4	<0.01	4	0.02	**4**	12.1
Zachary karate club	34	78	4	4	<0.01	4	<0.01	4	<0.01	**4**	8.03
Dolphins	62	159	14	17	<0.01	15	0.03	$\boxed{\mathbf{14}}$	0.01	$\boxed{\mathbf{14}}$	4.1
David Copperfield	112	425	18	19	0.01	18	0.11	19	0.03	**18**	4.32
Jazz musicians	198	2742	13	14	0.02	13	0.22	13	0.05	**13**	5.02
PDZBase	212	244	NA	54	0.01	52	0.58	52	0.31	52	41.3
U.Rovira i Virgili	1133	5451	210	229	0.77	213	246	217	1.44	$\boxed{211}$	770
Euroroad	1174	1417	384	471	0.57	400	271	$\boxed{392}$	1.23	397	1590
Hamsterster	1858	12534	241	253	1.9	242	667	245	6.37	242	1630
Hamsterster Full	2426	16631	416	437	4.59	416	1740	422	12.2	**416**	4680
Facebook (NIPS)	2888	2981	10	10	0.05	10	9.92	10	0.59	**10**	29.3

Table: Algorithm performance on the KONECT data sets (Time in sec).

- FPT Turbo Hybrid out-perform others
- FPT Turbo I better size/time ratio

Performance on the large KONECT Data Sets

Data Set			Greedy Chvátal		FPT **Turbo Hybrid***	
Name	$\|V\|$	$\|E\|$	**Size**	**Time**	**Size**	**Time**
Power Grid	4941	6594	1588	3.7	$\boxed{1499}$	37.6
Pretty Good Privacy	10680	24316	2862	129	$\boxed{2732}$	287
arXiv astro-ph	18771	198050	2509	628	$\boxed{2456}$	555
CAIDA	26475	53381	2422	447	$\boxed{2406}$	2411

Table: A comparison of the algorithm performance on the large KONECT data sets (Time in sec).

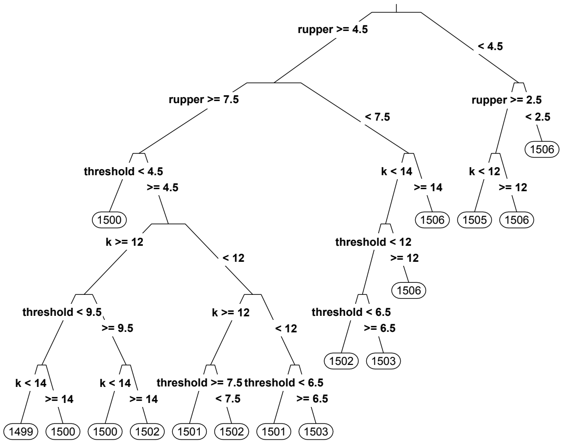
- Due to the size of these graphs, the exact algorithm and GRASP Local Search heuristic were not able to process the graphs.
- NOTE FPT Turbo runs faster than Greedy for some sets.

Power Grid: Effect of parameters on Time

- Solution size were measured on the Power Grid instance. chosen as it provided a large enough solution size to give a range of results.
- 823 results were recorded
  $1 \geq r_{upper} \geq (k-1), 3 \geq k \geq 15$ and $1 \geq threshold \geq k$
- These results have been classified using a Recursive partitioning [11]
- Best solution size were obtained with $r_{upper} \geq 5$ threshold t between $5 \geq t \geq 9$ and $k = 12$.

Power Grid Decision Tree - Effect of parameters on solution

Effect of r upper, k, and
Moment of Regret Trigger (threshold) on solution size



Effect of parameters on Time

- Of the 823 tests, only 43 obtained the minimum solution of 1499.
- this result was notably less than the greedy Chvátal solution of 1588.
- It seem interesting to then consider the execution times of these optimal solutions.

Descision Tree - Effect of parameters on Time

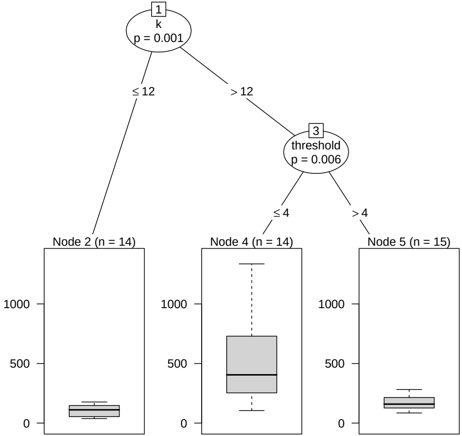


Figure: Execution times for moment of regret (threshold) r_{upper}, and k resulting in optimum solutions (KONECT–Power Grid Network). (Generated using the R ctree package [7])

Conclusion and Future Work

- Happy that the FPT procedure can have useful results
- For best results set $r > 5$ and use a none-trivial moment of regret function
- Obviously ordering effects the result
- New LS heuristics should be available soon to compare the results
- seems to give best results on scale-free graphs rather than graphs where greedy heuristics gave an optimum solution.

References

U. rovira i virgili network dataset – KONECT, April 2015.

Faisal N. Abu-Khzam, Michael A. Langston, Amer E. Mouawad, and Clinton P. Nolan, *A hybrid graph representation for recursive backtracking algorithms*, Frontiers in Algorithmics: 4th International Workshop, FAW 2010, Wuhan, China, August 11-13, 2010. Proceedings (Berlin, Heidelberg), Springer Berlin Heidelberg, 2010, pp. 136–147.

Vasek Chvátal, *A greedy heuristic for the set-covering problem*, Math. Oper. Res. **4** (1979), no. 3, 233–235.

R. G. Downey, J. Egan, M. R. Fellows, F. A. Rosamond, and P. Shaw, *Dynamic dominating set and turbo-charging greedy heuristics*, Tsinghua Science and Technology **19** (2014), no. 4, 329–337.

Fedor V. Fomin, Fabrizio Grandoni, and Dieter Kratsch, *Measure and conquer: Domination - A case study*, Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings (Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, eds.), Lecture Notes in Computer Science, vol. 3580, Springer, 2005, pp. 191–203.

Roger GuimerÃ, Leon Danon, Albert DÃŋaz-Guilera, Francesc Giralt, and Alex Arenas, *Self-similar community structure in a network of human interactions*, Phys. Rev. E **68** (2003), no. 6, 065103.

Torsten Hothorn, Kurt Hornik, and Achim Zeileis, *Unbiased recursive partitioning: A conditional inference framework*, Journal of Computational and Graphical Statistics **15** (2006), no. 3, 651–674.

David S. Johnson and Michael A. Trick, *The Second DIMACS Implementation Challenge*, `fpt://dimacs.rutgers.edu/pub/challenge`, 1993.

Stephen Milborrow, *rpart.plot: Plot 'rpart' models: An enhanced version of 'plot.rpart'*, 2018, R package version 2.2.0.

L. A. Sanchis, *Experimental analysis of heuristic algorithms for the dominating set problem*, Algorithmica **33** (2002), no. 1, 3–18.

Terry Therneau and Beth Atkinson, *rpart: Recursive partitioning and regression trees*, 2018, R package version 4.1-13.

THE
END

Performance on the DIMACS data set

DataSet			Opt	Greedy Chvátal		Greedy Vote GRASP		FPT Turbo I		FPT Turbo Hybrid	
Name	$\|V\|$	$\|E\|$	Size	Size	Time	Size	Time	Size	Time	Size	Time
C1000.9	1000	450079	2	3	0.83	3	2.88	3	3.4	3	2550
C125.9	125	6963	2	2	0	2	0.01	2	0.03	2	5.26
C2000.5	2000	999836	NA	7	5.56	7	42.5	7	8.38	7	150
C2000.9	2000	1799532	NA	3	3.83	3	13.4	3	17.7	3	88.9
C250.9	250	27984	2	2	0.02	2	0.06	2	0.23	2	14.6
C4000.5	4000	4000268	NA	7	25.9	8	234	8	49.8	7	717
C500.9	500	112332	2	2	0.11	2	0.31	2	0.81	2	154
DSJC1000.5	1000	249826	NA	6	1.02	6	6.95	6	2.49	6	1340
DSJC500.5	500	62624	NA	5	0.18	5	1.01	5	0.47	5	115
MANN_a27	378	70551	2	2	0.06	2	0.16	2	0.52	2	47.3
MANN_a81	3321	5506380	NA	2	11.5	2	25.4	2	74.4	2	256
brock200_2	200	9876	NA	5	0.02	4	0.09	4	0.06	4	9.95
brock200_4	200	13089	3	3	0.02	4	0.1	3	0.41	3	12.3
brock400_2	400	59786	3	3	0.09	3	0.31	3	0.41	3	90.1
brock400_4	400	59765	3	3	0.09	3	0.36	3	0.35	3	91.4
brock800_2	800	208166	4	4	0.53	4	2.38	4	1.79	4	850
brock800_4	800	207643	4	4	0.51	4	2.46	4	1.72	4	860
gen200_p0.9_44	200	17910	2	2	0.01	2	0.04	2	0.08	2	9.89
gen200_p0.9_55	200	17910	2	2	0.01	2	0.04	2	0.11	2	10.2
gen400_p0.9_55	400	71820	2	2	0.07	2	0.19	2	0.35	2	61.3
gen400_p0.9_65	400	71820	2	2	0.07	2	0.19	2	0.44	2	55.5
gen400_p0.9_75	400	71820	2	2	0.07	2	0.19	2	0.54	2	54.1
hamming10-4	1024	434176	2	2	0.53	2	1.47	2	2.76	2	2530
hamming8-4	256	20864	2	2	0.02	2	0.05	2	0.08	2	21.3
keller4	171	9435	2	2	0.01	2	0.02	2	0.04	2	9.21
keller5	776	225990	2	2	0.25	2	0.7	2	1.34	2	888
keller6	3361	4619898	2	2	8.47	2	26.7	2	46.7	2	234

Performance on the DIMACS (Compliment) data set

DataSet			Opt	Greedy Chvátal		Greedy Vote GRASP		FPT Turbo I		FPT Turbo Hybrid II	
Name	$\lvert V \rvert$	$\lvert E \rvert$	**Size**	**Size**	**Time**	**Size**	**Time**	**Size**	**Time**	**Size**	**Time**
brock200_2	200	10024		4	0.007	4	0.044	4	0.011	4	0.042
brock200_4	200	6811		7	0.003	6	0.025	6	0.003	6	0.019
brock400_2	400	20014		10	0.015	9	0.067	10	0.012	10	0.093
brock400_4	400	20035		10	0.005	10	0.032	9	0.006	9	0.217
brock800_2	800	111434		8	0.007	8	0.095	8	0.010	8	0.042
brock800_4	800	111957		8	0.006	8	0.105	8	0.030	8	0.102
C1000.9	1000	49421		25	0.021	24	0.351	31	0.037	24	0.157
C125.9	125	787		14	0.001	14	0.007	14	0.001	14	0.006
C2000.5	2000	999164		7	0.026	7	0.497	7	0.215	7	0.328
C250.9	250	3141		18	0.001	17	0.014	19	0.002	17	0.004
C4000.5	4000	3997732		8	0.089	8	3.073	8	2.604	8	3.730
C500.9	500	12418		21	0.003	20	0.137	29	0.007	21	0.030
DSJC1000.5	1000	249674		6	0.005	6	0.115	6	0.026	6	0.035
DSJC500.5	500	62126		6	0.002	5	0.026	6	0.006	5	0.009
gen200_p0.9_55	200	1990		17	0.001	15	0.011	17	0.002	16	0.004
gen400_p0.9_55	400	7980		19	0.002	19	0.042	20	0.006	19	0.016
gen400_p0.9_65	400	7980		20	0.002	19	0.043	24	0.006	19	0.008
gen400_p0.9_75	400	7980		21	0.003	19	0.038	27	0.007	19	0.008
hamming10-4	1024	89600		15	0.008	14	0.289	15	0.040	15	0.071
hamming8-4	256	11776		4	0.000	4	0.007	4	0.002	4	0.002
keller4	171	5100		6	0.000	6	0.002	6	0.001	6	0.001
keller5	776	74710		11	0.004	11	0.119	11	0.018	11	0.025
keller6	3361	1026582		18	0.081	18	3.749	18	0.759	18	1.051
MANN_a27	378	702		27	0.002	27	0.073	27	0.002	27	0.008
MANN_a45	1035	1980		45	0.017	45	0.770	45	0.016	45	0.087
MANN_a81	3321	6480		81	0.190	81	8.920	81	0.354	81	1.154
p_hat1500-1	1500	839327		3	0.007	3	0.279	3	0.009	3	0.017
p_hat1500-2	1500	555290		5	0.009	5	0.172	5	0.012	5	0.024
p_hat1500-3	1500	277006		11	0.013	11	0.288	11	0.082	11	0.115

Performance on the BSHOLIB data set

DataSet			Greedy Chvátal		Greedy Vote GRASP		FPT Turbo I		FPT Turbo Hybrid	
Name	$\|V\|$	$\|E\|$	**Size**	**Time**	**Size**	**Time**	**Size**	**Time**	**Size**	**Time**
frb30-15-1	450	17827	12	0.16	12	1.62	12	0.3	12	19.3
frb30-15-2	450	17874	12	0.13	12	1.56	13	0.17	12	18.9
frb30-15-3	450	17809	12	0.12	12	1.65	13	0.16	12	18.4
frb30-15-4	450	17831	13	0.14	12	1.62	13	0.13	12	19.6
frb30-15-5	450	17794	13	0.13	12	1.56	13	0.13	12	19.5
frb35-17-1	595	27856	14	0.25	14	3.52	15	0.21	14	33.7
frb35-17-2	595	27847	14	0.23	14	3.61	15	0.2	14	40.7
frb35-17-3	595	27931	15	0.27	14	3.92	16	0.23	14	35.9
frb35-17-4	595	27842	17	0.28	15	4.29	16	0.2	15	33.5
frb35-17-5	595	28143	14	0.25	15	4.24	15	0.2	13	45.5
frb40-19-1	760	41314	16	0.42	16	7.95	17	0.33	16	62.6
frb40-19-2	760	41263	17	0.46	16	7.43	17	0.33	16	62.7
frb40-19-3	760	41095	17	0.46	16	7.47	18	0.33	16	67.5
frb40-19-4	760	41605	17	0.46	16	7.64	17	0.32	16	65
frb40-19-5	760	41619	16	0.45	16	7.64	17	0.3	16	64.8
frb45-21-1	945	59186	18	0.72	18	13.7	19	0.44	18	105
frb45-21-2	945	58624	19	0.78	19	15.2	20	0.43	19	118
frb45-21-3	945	58245	18	0.72	18	13.8	19	0.48	18	116
frb45-21-4	945	58549	18	0.69	18	13.8	19	0.47	18	105