Solving weakened cryptanalysis problems of Bivium cipher in the volunteer project SAT@home

Oleg Zaikin, Alexander Semenov, Mikhail Posypkin

ISDCT SB RAS, Irkutsk, Russia
IITP RAS, Moskow, Russia

SAT approach

There are many practically important problems for which the existence of effective (polynomial) algorithms of their solving has not been proven.

Most of these problems are NP-hard.

However, many of their special cases arise in practical applications:

- planning of production
- designing and verification of hardware
- proving of program correctness

SAT approach

Therefore it is very important to have methods for their solving that don't have polynomial complexity, but are effective in practice.

Such methods can cope with the numerous special cases of NP-hard problems of huge dimensions.

One of the most efficient in terms of software implementations is SAT approach.

SAT approach

SAT approach is based on reducibility of the considered original problem to a Boolean satisfiability problem (SAT), usually in the form of equation CNF=1.

Solution of SAT problem is an Boolean vector – assignments of CNF variables which force CNF to become 1.

From this assignments one can easily get solution of and original problem.

SAT approach

Example of CNF with 3 clauses over 5 variables:

$$C = (x_1 \vee \overline{x_2}) \cdot (x_2 \vee x_3 \vee \overline{x_4}) \cdot (\overline{x_3} \vee x_4 \vee \overline{x_5})$$

This CNF is satisfiable, for example on (11001).

Real hard SAT problem can consist of thousands variables and millions clauses.

Real SAT problem can be unsatisfiable.

Logical cryptanalysis

In logical cryptanalysis an original problem is reduced to a SAT problem, then after solving SAT problem we can easily obtain secret key from satisfying assignment.

Term appeared in 2000 in the paper of Italian researchers Massaci and Mararo.

They reduced DES cryptanalysis problem to a SAT problem and successfully solved some weakened problems of DES with the help of SAT solvers.

Bivium cipher

The Bivium stream cipher uses two shift registers of a special kind. The first register contains 93 cells and the second contains 84 cells.

Cryptanalysis problem:

Based on the known fragment of keystream (200 bits) we search for the values of all 177 bits of registers cells.

Bivium SAT encoding

We used TranAlg tool (developed in ISDCT SB RAS) for creating SAT encoding for Bivium cryptanalysis problem.

Created CNF consists of 777 variables and 12800 clauses. First 177 variables encode register state (we need to find it), last 200 encode known keystream.

First 177 variables are special: other variables depends on them.

Logical cryptanalysis of Bivium cipher

Monte Carlo method: choose a set of K variables, solve a small randomly chosen subset of subproblems from 2^K, then make estimation for 2^K.

For example, randomly choose and solve 1000 subproblems from 2^40, calculated average time (for 1000 subproblems) and multiply it to 2^40.

In a SAT problem encoding cryptanalysis problem for Bivium there is usually 1 satisfying assignment.

So, we have to solve about half subproblems (in average) to find this assignment.

Logical cryptanalysis of Bivium cipher

German researchers (Eibach, Pilz, Vokel) made estimation for Bivium logical cryptanalysis (2008 year)

1.637×10^{13} seconds on one CPU core, set of 45 variables (“manually” found set)

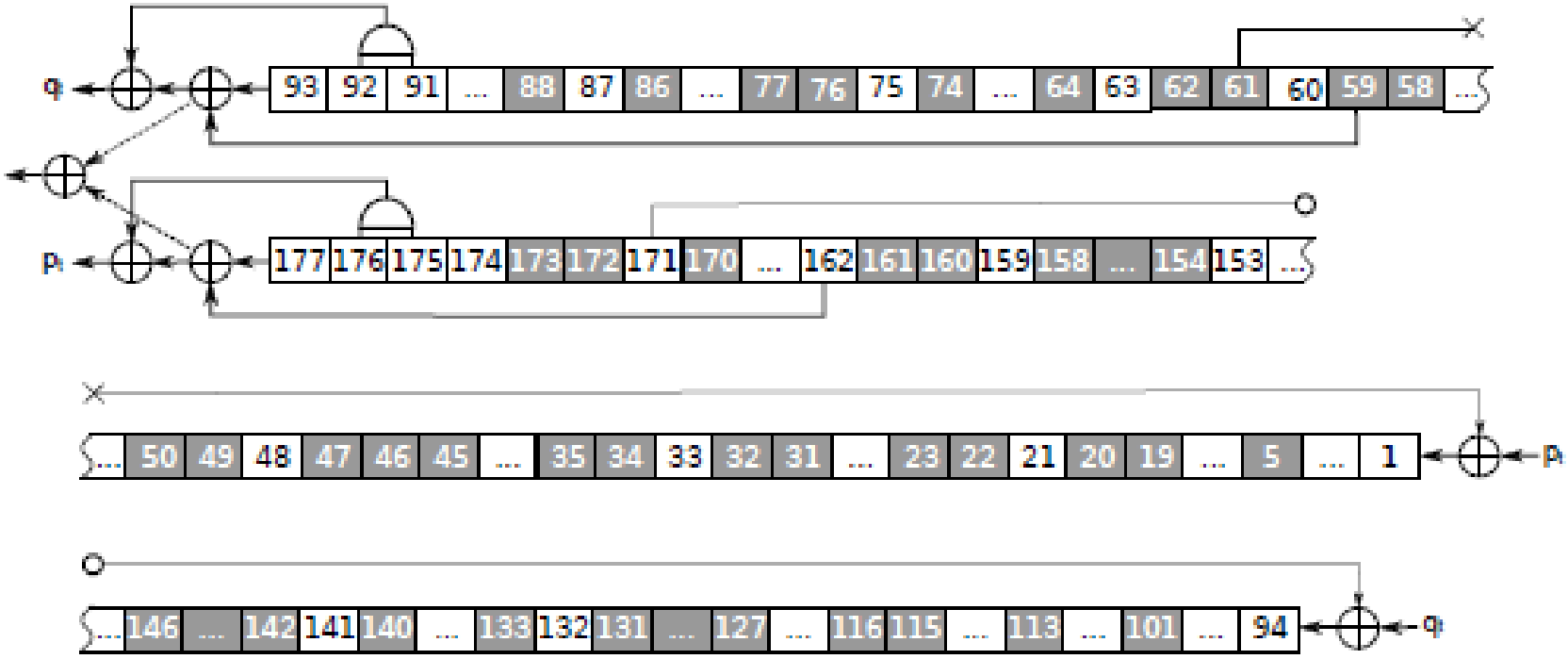
Our estimation (2014)

2.085×10^{11} seconds on one CPU core,

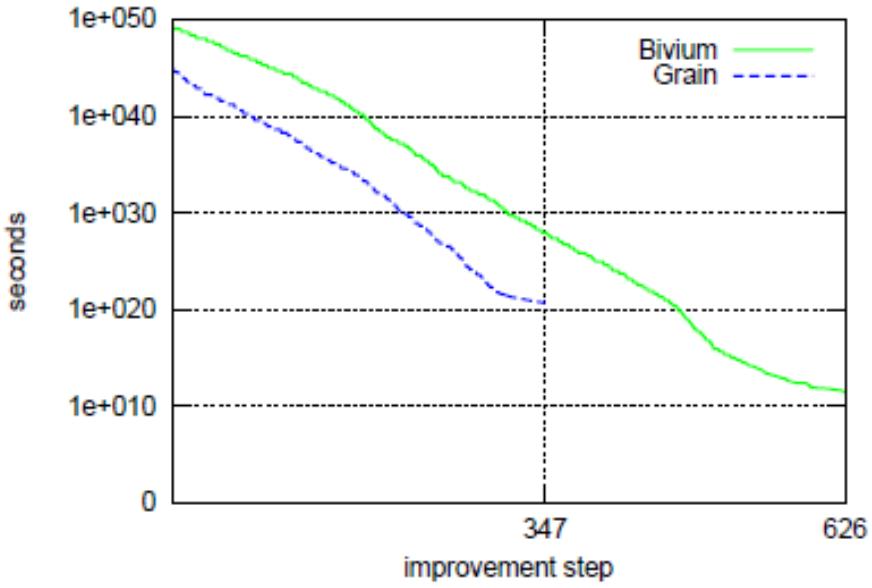
set of 50 variables (automatically found set)

Estimations for Bivium

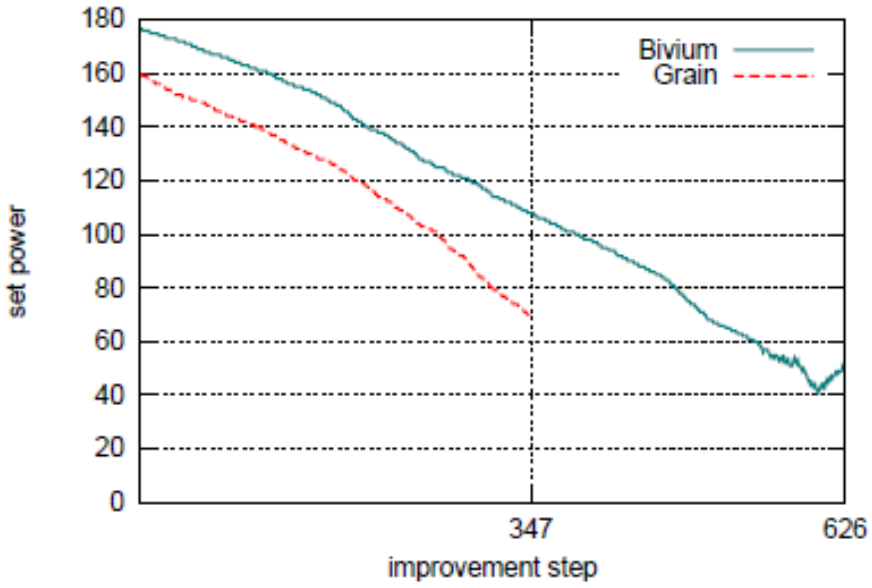
We used MPI program PDSAT with tabu search optimization algorithm. PDSAT started from first 177 variables (they encodes register state) then it can add/remove any of these 177 variables.



Estimations for Bivium



(a) Dynamics of F_{best}



(b) Dynamics of $\left|\tilde{X}_{best}\right|$

Why weakened cryptanalysis problems?

By *BiviumK* we denote a weakened problem for Bivium with known values of K variables (in corresponding SAT problem) encoding last K cells of the second shift register.

Germans solved several Bivium33 problems. It's quite easy problem (about 1 minute on 1 CPU core).

Why weakened cryptanalysis problems?

With the help of sets found by PDSAT we solved several Bivium16, Bivum14 and Bivium12 problems (for every such problems own estimation procedure required).

Bivium10 is quite hard problem.

For Bivium10 we found a set of 40 variables with estimation 6.7 × 10^8 seconds on one CPU core.

Estimation for 1000 cluster cores: 186 hours.

Solving Bivium10

We launched 3 Bivium10 problems on cluster with time limit 0.01 second for every subproblem.

Each problem was running about 2 days on 960 cluster CPU cores.

About 1.2 % of subproblems were interrupted (more then 1 billion subproblems).

These subproblems were solved in SAT@home

Volunteer computing project SAT@home

Started on September 2011 by ISDCT SB RAS and IITP RAS.

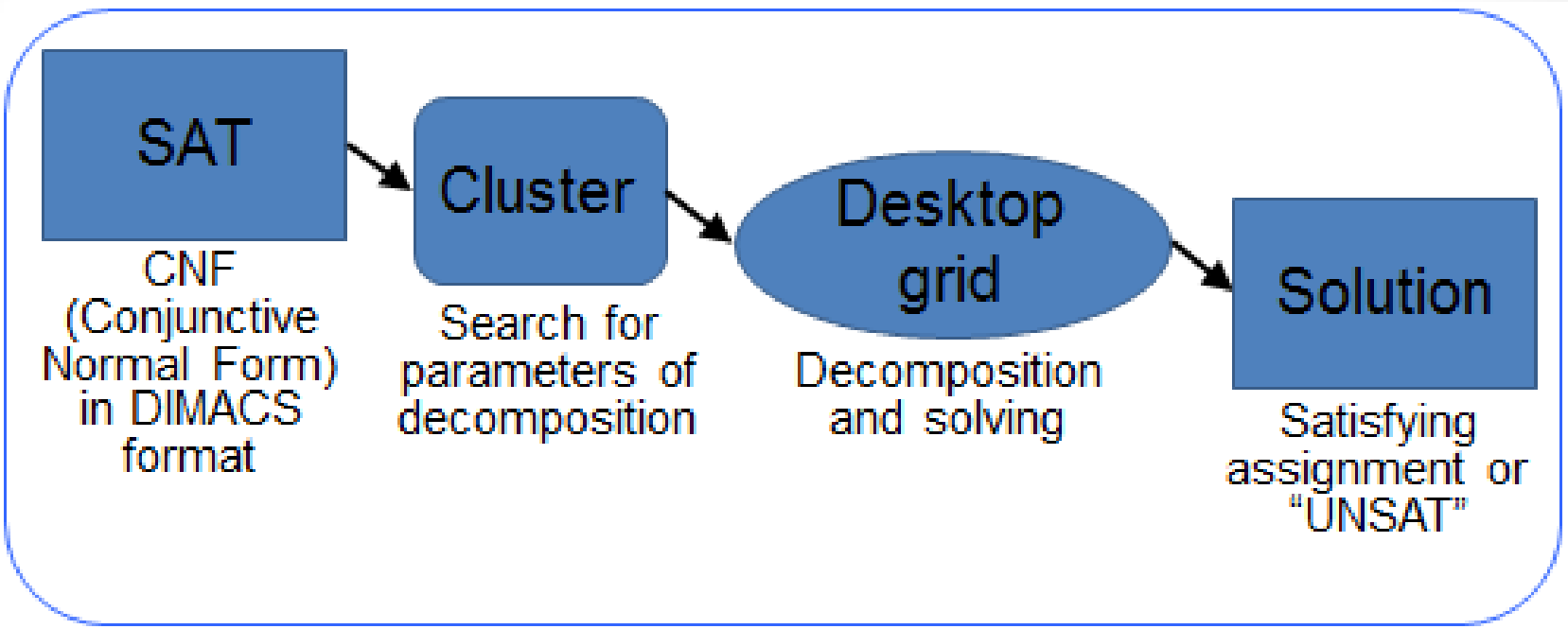
Goal - solve hard large-scale problems that can be effectively reduced to SAT.

Solved problems:

- A5/1 logical cryptanalysis problem;
- New pairs of orthogonal diagonal Latin squares of order 10.

App versions: windows, linux

Solving scheme



- While finding parameters of decomposition we take into account peculiarities of the original problems
- CNF encoding of the original problem is decomposed into a set of independent CNFs

Solving Bivium10 problems in SAT@home

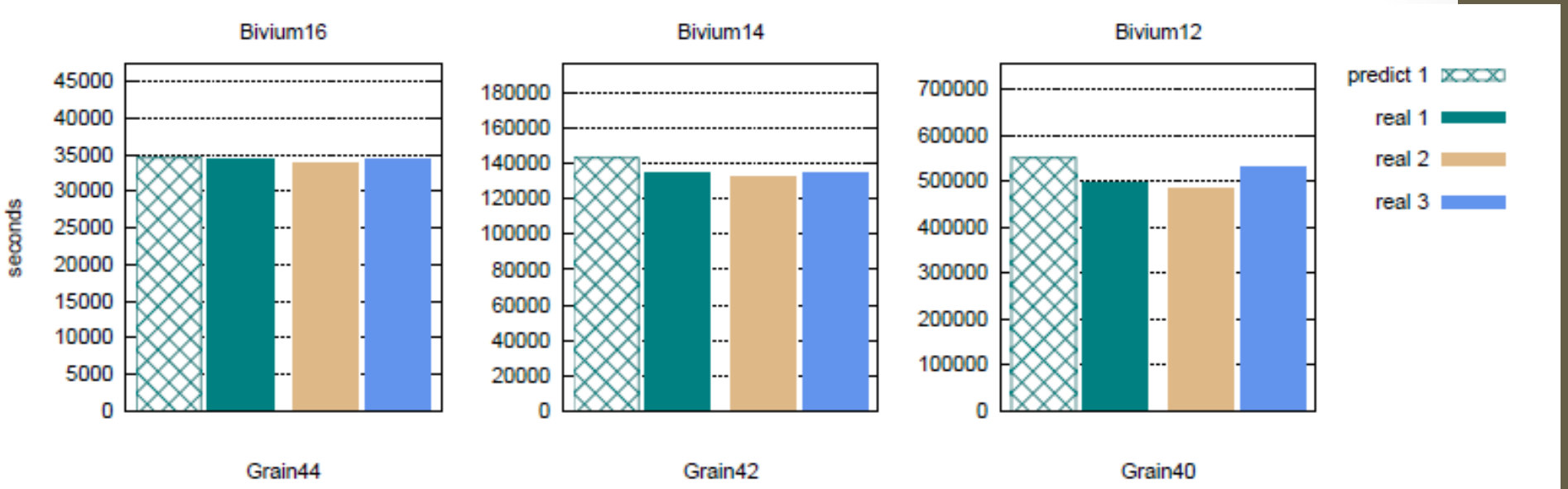
3 Bivium10 problems were solved, each for about 2 weeks.

Start of the experiment: 2014-02-07

№–	Date	Users	Problem	Value
1	2014-02-13 20:49:40 UTC	Retupmoc from Retupmoc / dimawalker	Bivium16_2	8CECD32330575840
2	2014-03-16 17:14:30 UTC	Crystal Spirit from Crystal Dream / Shaman from Czech National Team	Bivium14_1	81BB1DBCA1C33564
3	2014-03-20 19:06:14 UTC	madx from BOINC RUSSIA / kunzea	Bivium14_0	C78CD4937363FF09
4	2014-04-10 07:43:21 UTC	mi5rys from Crystal Dream / Alexandr Burnashev from Russia Team	Bivium10_0	C78CD4937363FF09
5	2014-05-07 11:10:08 UTC	Pavel_Kirpichenko from Astronomy.Ru Forum / Crystal Spirit from Crystal Dream	Bivium10_1	81BB1DBCA1C33564
6	2014-05-26 12:02:09 UTC	http://vk.com/boinc from Colombia / alnb from Russia Team	Bivium10_2	8CECD32330575840

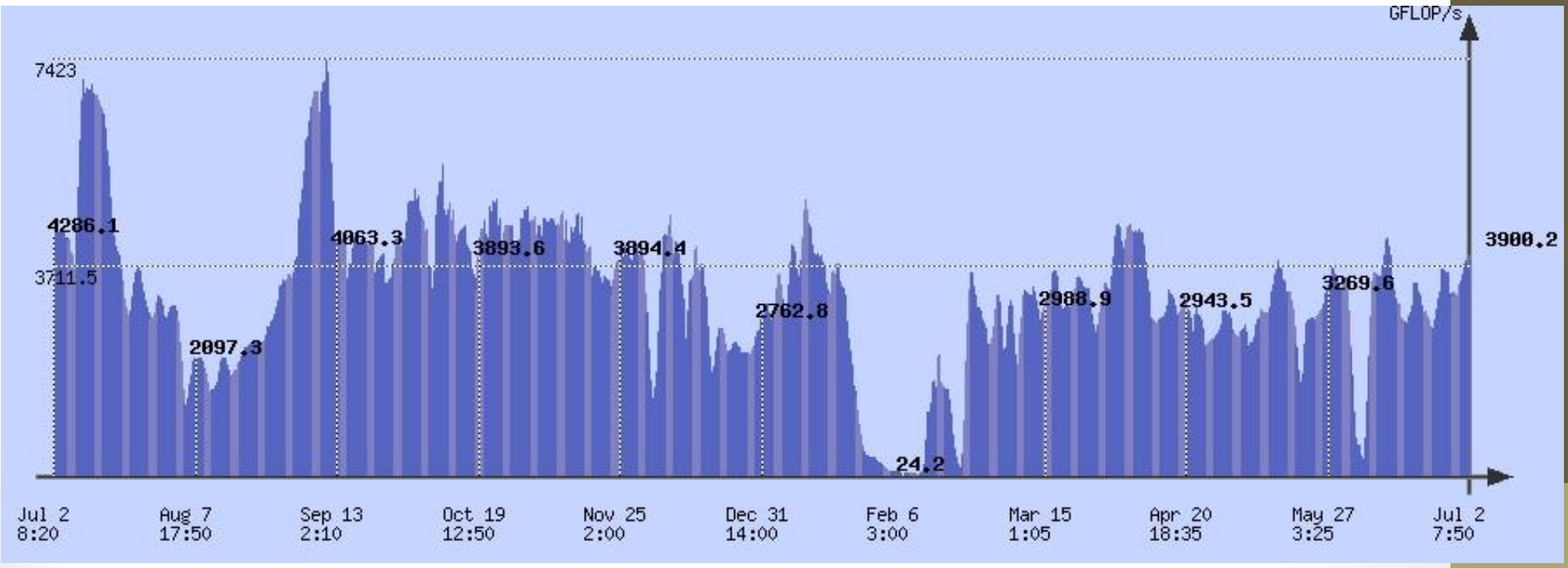
Estimation VS real time

For Bivium10 we have similar comparisons.



SAT@home performance

Current: 3.9 TFLOPs
Peak: 7.4 TFLOPs



Last year dynamics

Conclusions

- Checking estimations obtained for slightly weakened problem cryptanalysis problems is very important
- We used cluster for searching decompositions for Bivium10 and for solving small independent subrpoblems
- We used SAT@home for solving “hard” independent subproblems for Bivium10
- Results for Bivium10 shows good precise of estimations, so we can hope than estimation for Bivium is precise too
- Volunteer computing – is a good tool for solving such weakened problems

Thank you for your attention!