Average versus Worst in Solving Sparse Algebraic Equations

Igor Semaev

Department of Informatics, University of Bergen, Norway

Rump session, Crypto2007

Notation

- F_q finite field with q elements;
- $X = \{x_1, x_2, \ldots, x_n\}$ variables from F_q;
- X_i subsets of X of size l;
- f_i polynomials over F_q in variables X_i.

Problem

- Look for all solutions in F_q to the nonlinear equations

$$f_1(X_1) = 0, \ldots, f_m(X_m) = 0,$$

- Equations are called l-sparse.
- Motivation: cryptanalysis. E.g. DES $\quad m = 512$ Boolean equations, $n = 504$ variables, at most $l = 14$ variables in each equation

Worst case

$q = 2$

- l-**sparse equations** is polynomially equivalent to l-**SAT** with the same set of variables.
- The worst case is the same and complexity bounds are the same.
- The average cases are different.

Our Results

1. Deterministic Agreeing-Gluing1 algorithm to solve l-sparse equations.
2. Simple and practical.
3. Almost no additional memory is required. Keep only initial equations.
4. We estimate the expected complexity.

Probabilistic Model

- The equations $f_i(X_i)$ are chosen:
  1. randomly,
  2. independently of each other,
  3. X_i and f_i have uniform distribution.

- The Algorithm complexity is a random variable.

- Its expectation is rigorously estimated.

- The estimates are being compared with the worst case.

Average versus worst

- Let $q = 2$ and $m = n$.

$l =$	3	4	5	6
the worst case	1.324^n	1.474^n	1.569^n	1.637^n
Agreeing-Gluing1, expectation	1.113^n	1.205^n	1.276^n	1.334^n

.

- Significant difference in the worst and average cases.
- E.g. for $l = 3$ the bounds are

$n =$	100	300	500	1000
the worst case	$1.5\,10^{12}$	$3.6\,10^{36}$	$8.7\,10^{60}$	$7.7\,10^{121}$
Agreeing-Gluing1, expectation	$4.4\,10^{4}$	$8.8\,10^{13}$	$1.7\,10^{23}$	$3.1\,10^{46}$

.

Conclusion

Average systems of sparse algebraic equations

are not so difficult as one may expect.