Pokročilá kryptologie

Algebraická kryptoanalýza

Mgr. Martin Jureček

České vysoké učení technické v Praze, Fakulta informačních technologií
Katedra informační bezpečnosti

Algebraická kryptoanalýza

„Breaking a good cipher should require as much work as solving a system of simultaneous equations in a large number of unknowns of a complex type.“

C.E. Shannon, 1949

Algebraická kryptoanalýza

Úvod

- Algebraická kryptoanalýza, dále zkráceně AK, je nové odvětví kryptoanalýzy, které si v poslední době získalo velkou pozornost.
- Princip AK spočívá v převodu problému prolomení kryptosystému na problém vyřešení soustavy polynomiálních rovnic nad konečným tělesem.
- Uplatnění AK je hlavně v symetrické kryptografii:
  - příklady blokových šifer: AES, DES
  - príklady proudových šifer: E0 - Bluetooth, Toyocrypt

  ale AK byla aplikovaná i v asymetrické kryptografii.

Algebraická kryptoanalýza

Postup

- Postup algebraické kryptoanalýzy se rozděluje do dvou kroků:

  1: Ze specifických vlastností dané šifry se odvodí soustava polynomiálních rovnic nad konečným tělesem.
  2: Aplikuje se některý postup pro výpočet řešení soustavy, ze kterého se potom odvodí tajný klíč šifry.

- Ale u AK existuje fundamentální problém: vyřešení soustavy polynomiálních rovnic nad konečným tělesem je NP-úplný problém.
- Proč bysme problém prolomení šifry měli převést na problém, pro který neznáme rychlý algoritmus(t.j. s polynomiální složitostí), který ho vyřeší?

Algebraická kryptoanalýza

1.krok

- První krok AK spočívá ve využití struktury šifry a všech jejich aktivit a z nich se sestaví soustava rovnic, která chování šifry pro konkrétní případ popisuje.
- Soustavu rovnic uvažujeme nad konečným tělesem, obvykle nad GF(2).
- Cílem je získat co nejmenší soustavu obsahující polynomy s co nejnižšími stupněmi, pričemž postup, jakým se odvodí rovnice je závislý na konkrétní šifře.
- Když by soustava obsahovala jen lineární rovnice, tak použijeme např. Gaussovu eliminaci, která má kubickou složitost a poměrně rychle se dopracujeme k výsledku.
- Ale dobře navrhnuté šifry poskytují soustavu polynomiálních rovnic, které nedokážeme vyřešit v krátkém čase.

Algebraická kryptoanalýza

2.krok

- Hlavní částí AK je právě 2.krok, ve kterém se vyřeší soustava polynomiálních rovnic nad konečným tělesem.
- Pro lepší představu uveďme, že v případě AES-128 soustava obsahuje přibližně 8000 rovnic s 1600 proměnnými a v případě AES-256 až 22400 rovnic s 4480 proměnnými.
- Existuje několik postupů pro výpočet takových soustav nelineárních rovnic, ale žádný z nich není rychlý.
- Když by pro tento problém existoval efektivní algoritmus(t.j. s polynomiální složitostí), tak by platilo, že P=NP, což by se považovalo za překvapující.

Algebraická kryptoanalýza

Postupy řešení soustavy

- Jednoduchým, ale ne příliš efektivním postupem je postup typu „uhádni a odvoď“(„guess and determine“).
- Spočívá v tom, že „uhádneme“(spočítáme pomocí hrubé síly) hodnoty vhodných proměnných a zbytek soustavy už dopočítáme jednodušeji.
- Dalšími příklady postupů jsou: linearizace, kterou si později uvedeme na příkladě, dále XL algoritmus a v neposledním případě Gröbnerovy báze.
- Gröbnerovy báze se považují za velmi perspektivní metodu a úspěšně byly aplikované např. na AES

Algebraická kryptoanalýza

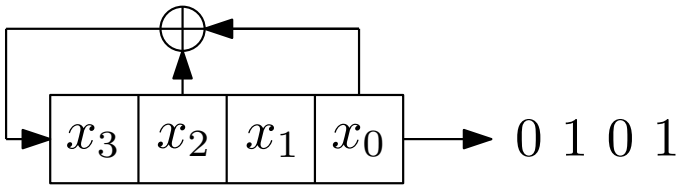
Aplikace AK na některé typy proudových šifer

- Uvedeme si příklady použití AK na tři třídy proudových šifer:
  - LFSR(Linear Feedback Shift Register) - šifra složená jen z jednoho LFSR
  - NLCG(Nonlinear Combination Generator) - šifra složená z vícerých LFSR a nelineární booleovské funkce, podle které se počítá výstup(jejím vstupem jsou jen výstupní bity registrů)
  - NLFG(Nonlinear Filter Generator) - šifra složená jen z jednoho LFSR a nelineární booleovské funkce, podle které se počítá výstup(jejím vstupem jsou všechny bity registru)
- Každý uvedený příklad AK bude typu „known plaintext attack“, což implikuje, že můžeme předpokládat znalost keystreamu.

Algebraická kryptoanalýza

Příklad použití AK na LFSR

- Uvažujme následující jednoduchý LFSR délky 4:



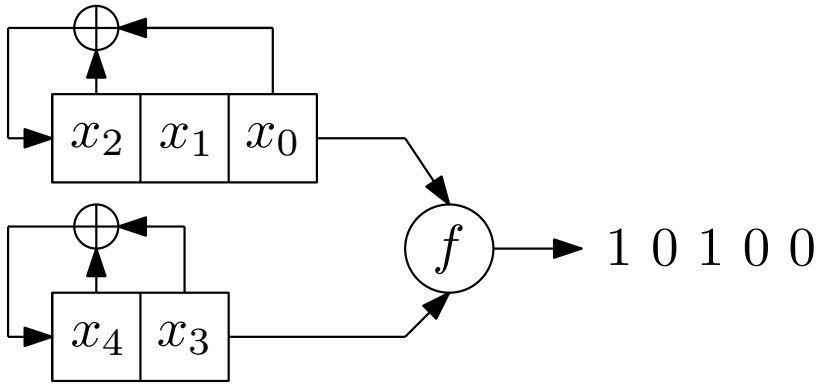
- Naší úlohou je získat hodnoty $x_0, \ldots, x_3 \in \mathrm{GF}(2)$.

posun	stav LFSR	výstup	rovnice
1.	(x_3, x_2, x_1, x_0)	x_0	$x_0 = 0$
2.	$(x_0 \oplus x_2, x_3, x_2, x_1)$	x_1	$x_1 = 1$
3.	$(x_1 \oplus x_3, x_0 \oplus x_2, x_3, x_2)$	x_2	$x_2 = 0$
4.	$(x_0, x_1 \oplus x_3, x_0 \oplus x_2, x_3)$	x_3	$x_3 = 1$

- V tomto případě je už soustava přímo vyřešená.

Příklad použití AK na NLCG

- Uvažujme následující případ složený ze dvou LFSR:



  a nelineární funkce $f(v_1, v_2) = v_1 + v_1 v_2$, kde v_1, resp. v_2 je výstupní bit prvního, resp. druhého registru.

- Cílem je spočítat hodnoty $x_0, \ldots, x_4 \in \mathrm{GF}(2)$.

Odvození soustavy

- První registr si označme R_1 a druhý R_2 a jejich výstupní bity po řade v_1 a v_2

posun	stav R_1	stav R_2	v_1, v_2
1.	(x_2, x_1, x_0)	(x_4, x_3)	x_0, x_3
2.	$(x_0 \oplus x_2, x_2, x_1)$	$(x_3 \oplus x_4, x_4)$	x_1, x_4
3.	$(x_0 \oplus x_1 \oplus x_2, x_0 \oplus x_2, x_2)$	$(x_3, x_3 \oplus x_4)$	$x_2, x_3 \oplus x_4$
4.	$(x_0 \oplus x_1, x_0 \oplus x_1 \oplus x_2, x_0 \oplus x_2)$	(x_4, x_3)	$x_0 \oplus x_2, x_3$
5.	$(x_1 \oplus x_2, x_0 \oplus x_1, x_0 \oplus x_1 \oplus x_2)$	$(x_3 \oplus x_4, x_4)$	$x_0 \oplus x_1 \oplus x_2, x_4$

Algebraická kryptoanalýza

Soustava polynomiálních rovnic

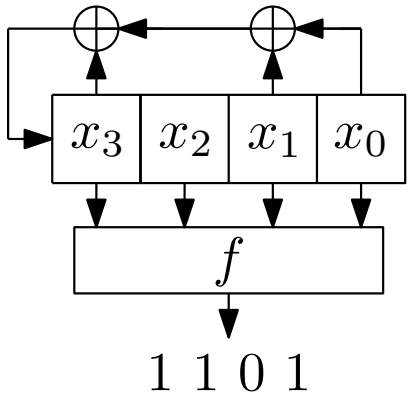
$$
\begin{aligned}
x_0 + x_0x_3 &= 1 \\
x_1 + x_1x_4 &= 0 \\
x_2 + x_2x_3 + x_2x_4 &= 1 \\
x_0 + x_2 + x_0x_3 + x_2x_3 &= 0 \\
x_0 + x_1 + x_2 + x_0x_4 + x_1x_4 + x_2x_4 &= 0
\end{aligned}
$$

- Dokázali byste ji vyřešit?
- (např. z první rovnice je jasné, že $x_0 = 1$)

Algebraická kryptoanalýza

Příklad použití AK na NLFG

- Uvažujme následující případ jednoho LFSR:



  a nelineární funkce $f(x_0, x_1, x_2, x_3) = x_0 + x_0x_1 + x_1x_3$.
- Cílem bude opět spočítat hodnoty $x_0, \ldots, x_3 \in \mathrm{GF}(2)$.

Algebraická kryptoanalýza

Odvození soustavy

posun	stav registru
1.	(x_3, x_2, x_1, x_0)
2.	$(x_0 + x_1 + x_3, x_3, x_2, x_1)$
3.	$(x_0 + x_2 + x_3, x_0 + x_1 + x_3, x_3, x_2)$
4.	$(x_0, x_0 + x_2 + x_3, x_0 + x_1 + x_3, x_3)$

Soustava polynomiálních rovnic

$$\begin{aligned} x_0 + x_0x_1 + x_1x_3 &= 1 \\ x_1 + x_0x_2 + x_2x_3 &= 1 \\ x_2 + x_0x_3 + x_3^2 &= 0 \\ x_3 + x_1x_3 + x_3^2 + x_0^2 + x_0x_1 &= 1 \end{aligned}$$

- Dokázali byste uvedenou soustavu vyřešit?

Algebraická kryptoanalýza

Poznámky

- V případě verze NLFG platí, že stupeň generovaných rovnic je shora omezený stupněm nelineární funkce f.
- Odhad na maximální stupeň rovnic má pro AK velký význam, protože AK je efektivnější u soustav s nízkým maximálním stupněm.
- Důležitým faktem generování soustavy rovnic je ten, že tento krok nezávisí na konkrétních hodnotách(keystreamu).
- Proto sestavení rovnic můžeme vykonat ještě před samotným útokem a proto hlavním a výpočtově nejsložitějším krokem bude právě 2.krok - výpočet soustavy.
- *Challenge: Zkuste sestavit soustavu rovnic pro šifru A5/1, která je složená z 3 LFSR a nelineárního prvku, který určuje, které registry se v aktuálním čase posunou.*

Algebraická kryptoanalýza

Linearizace

- Nakonec si uvedeme jeden základní postup na řešení soustavy polynomiálních rovnic nazvaný linearizace.
- Podstatu algoritmu můžeme shrnout do následujících třech kroků:

1: V každé rovnici každý výraz, který je ve tvaru součinu, nahradíme novou proměnnou.
2: Vyřešíme soustavu lineárních rovnic(např. pomocí Gaussovy eliminace).
3: Řešení dosadíme do původní rovnice a ověříme jeho správnost.

- Protože linearizací soustavy můžou vznikat lineárně závislé rovnice, běžně se používají různé vylepšení.

Algebraická kryptoanalýza

Linearizace - příklad 1

- Uvažujme následující soustavu nad GF(2):

$$\begin{aligned} x + xy &= 1 \\ x + y &= 1 \\ x + y + xy &= 1 \end{aligned}$$

- V 1.kroku algoritmu nahradíme výraz xy proměnnou z:

$$\begin{aligned} x + z &= 1 \\ x + y &= 1 \\ x + y + z &= 1 \end{aligned}$$

- V 2.kroku lehce spočítáme řešení: $x = 1$, $y = 0$, $z = 0$
- V 3.kroku ověříme, že řešení je skutečně správné.

Linearizácia - príklad 2

- Nutnost ověřování výsledku ilustruje následující příklad:

$$\begin{aligned} x + xy &= 1 \\ x + y &= 0 \\ x + y + xy &= 0 \end{aligned}$$

- V prvním kroku nahradíme výraz xy proměnnou z:

$$\begin{aligned} x + z &= 1 \\ x + y &= 0 \\ x + y + z &= 0 \end{aligned}$$

- Vyřešením soustavy dostaneme: $x = 1$, $y = 1$, $z = 0$.
- Ale po dosazení do 1.rovnice dostaneme: $1 + 1 * 1 \neq 1$, takže ověřování výsledku je nevyhnutelné.