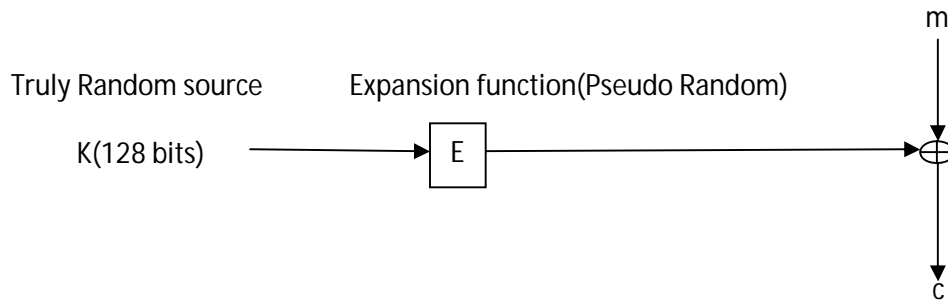


9/4/2009

CSE 508 - Network Security

Stream Cipher



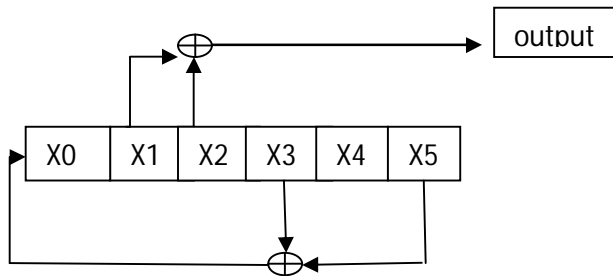
Pseudo random sequence have the following properties

- Look Random
- Can't predict unknown portions of output from some known portions (Next bit tests)
- Given output, adversary should not be able to predict seed
- No repeat
- Efficient
- Long output

Example – Linear Congruential Generator

```
Unsigned int state;  
  
Int E (void)  
{  
State=322349*state + 45656749;  
Return state%2;  
}
```

Linear feedback shift register



- A shift register with all zeros will cause LFSR to output a never ending stream of zeros , which is not useful
- It can in theory generate $2^n - 1$ bit long pseudo random sequence before repeating

Blum Blum Shub

- Pick large primes $p, q, N = pq$
- Pick $X_0 \in \{2, 3, \dots, N-1\}$
- $X_{n+1} = X_n^2 \pmod N$
- $\text{Output}_n = X_n \pmod 2$
- Breaking is equal to factoring N
- Blum Blum Shub is comparatively slow, hence not useful for stream ciphers
- Useful for key generation and high security applications

Probability Distribution and Distinguishability

Definition- A probability distribution function that assigns a probability to each possible value in some set.

$X \leftarrow D$ means Picking an element X from Distribution D

Example Uniform Distribution

$$U_S(s) = 1/|S|$$

Example

$D = U(0,1,2,3,4,5)$

$X \leftarrow D$

X	0	1	2	3	4	5
	1/6	1/6	1/6	1/6	1/6	1/6

X^2	0	1	4	3	4	1
	1/6	1/3	1/3	1/6	1/3	1/3

Statistical Indistinguishability

Two distributions D and D' are ϵ statistically indistinguishable if for all algorithm A

$$\text{Adv } A = |\Pr[A(x)=1 | X \leftarrow D] - \Pr[A(x)=1 | X \leftarrow D']| \leq \epsilon$$

Computational Indistinguishability

Distribution D and D' are t, ϵ computationally indistinguishable if for all algorithms A running in time $\leq t$, $\text{Adv} \leq \epsilon$

Theorem

If $D1 \stackrel{t}{\sim}_{\epsilon_1} D2$ and $D2 \stackrel{t}{\sim}_{\epsilon_2} D3$ then $D1 \stackrel{t}{\sim}_{\epsilon_1 + \epsilon_2} D3$

Proof – Let A be any algorithm running in time t

$$|\Pr[A(x)=1 | X \leftarrow D1] - \Pr[A(x)=1 | X \leftarrow D3]| =$$

$$|\Pr[A(x)=1 | X \leftarrow D1] - \Pr[A(x)=1 | X \leftarrow D2] + \Pr[A(x)=1 | X \leftarrow D2] - \Pr[A(x)=1 | X \leftarrow D3]|$$

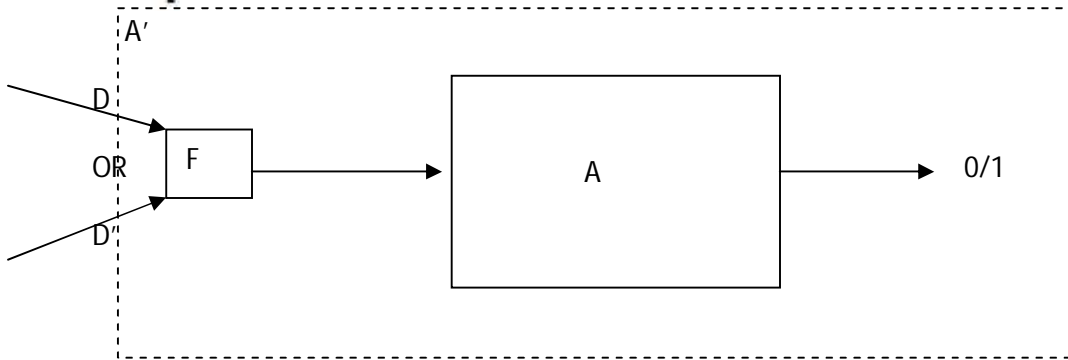
$$\leq |\Pr[A(x)=1 | X \leftarrow D1] - \Pr[A(x)=1 | X \leftarrow D2]| + |\Pr[A(x)=1 | X \leftarrow D2] - \Pr[A(x)=1 | X \leftarrow D3]|$$

$$\leq \epsilon_1 + \epsilon_2$$

Data Processing Inequality

Theorem - If $D \stackrel{t}{\sim} D'$ and f is any function computable in time t' then

$$F(D) \stackrel{t-t'}{\sim}_{\epsilon} F(D')$$



Proof (by contrapositive) - Suppose $\exists A$ running in time $t-t'$ such that $\text{Adv}_{F(D), F(D')} A > \epsilon$

$$\begin{aligned} & \Pr[A'(x)=1 | X \leftarrow D] \\ &= \Pr[A(f(x))=1 | X \leftarrow D] \\ &= \Pr[A(x)=1 | X \leftarrow f(D)] \end{aligned}$$

$$\begin{aligned} & \Pr[A'(x)=1 | X \leftarrow D'] \\ &= \Pr[A(f(x))=1 | X \leftarrow D'] \\ &= \Pr[A(x)=1 | X \leftarrow f(D')] \end{aligned}$$

$$\begin{aligned} \text{So, } \text{Adv}_{D, D'} A' &= | \Pr[A'(x)=1 | X \leftarrow D] - \Pr[A'(x)=1 | X \leftarrow D'] | \\ &= | \Pr[A(x)=1 | X \leftarrow f(D)] - \Pr[A(x)=1 | X \leftarrow f(D')] | \end{aligned}$$

$$\text{ADV} > \epsilon$$