

Audit Trail Pattern Analysis for Detecting Suspicious Process Behavior

Andreas Wespi, Marc Dacier, Hervé Debar, and Mehdi
M. Nassehi

{anw, dac, deb, mmn}@zurich.ibm.com

Global Security Analysis Lab
IBM Zurich Research Laboratory

RAID '98, Louvain-la-Neuve, Belgium, Sep. 14-16, 1998



Agenda

- Context
- Problem
- Approach
- Results
- Conclusions

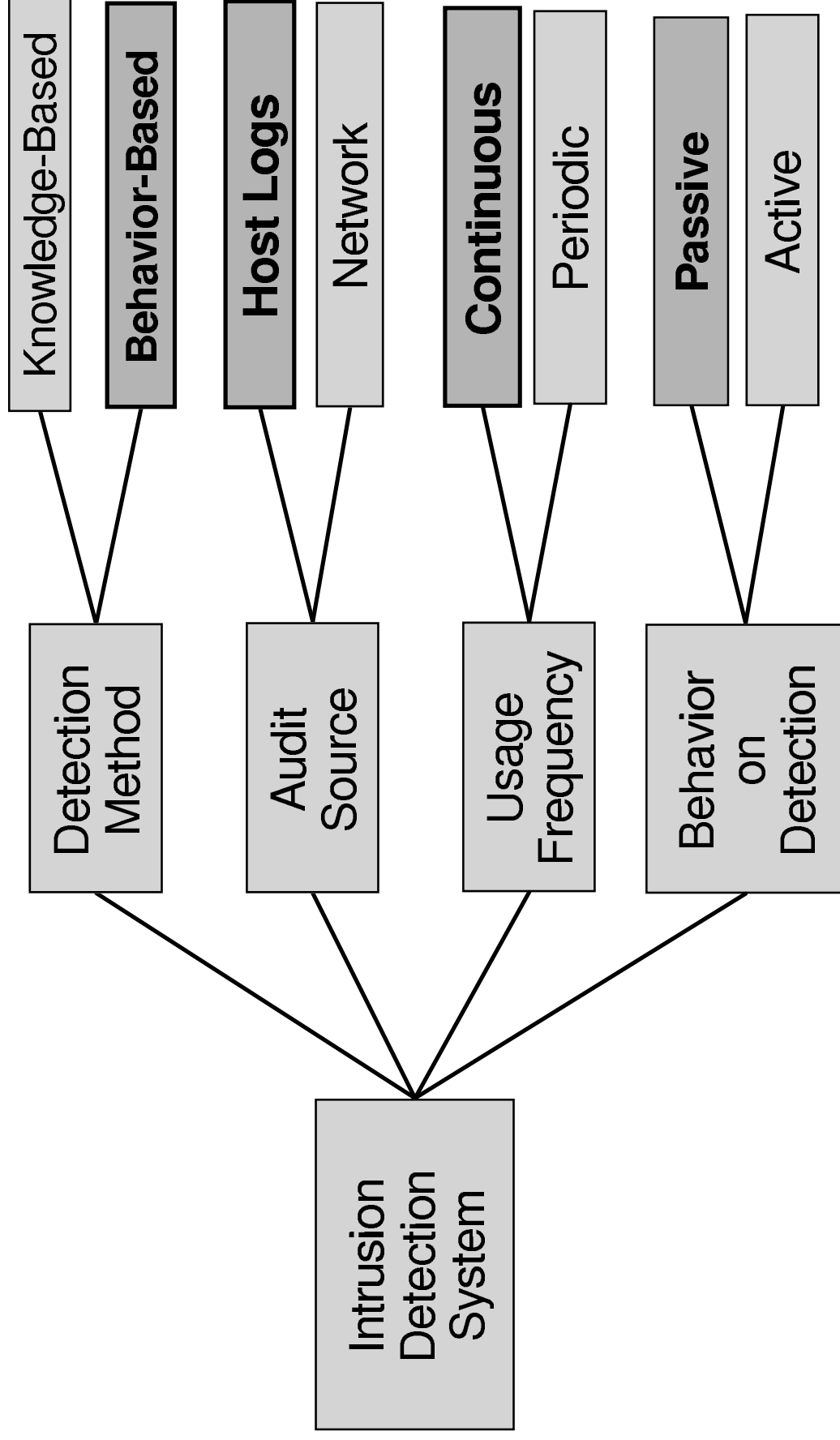


Preliminary Work

- S. Forrest et al., "A Sense of Self for Unix Processes", IEEE Symposium on Security and Privacy, Oakland, 1996.
- H. Debar et al., "Fixed vs. Variable-Length Patterns for Detecting Suspicious Process Behavior", ESORICS '98, Louvain-la-Neuve, 1998.



Our IDS Component



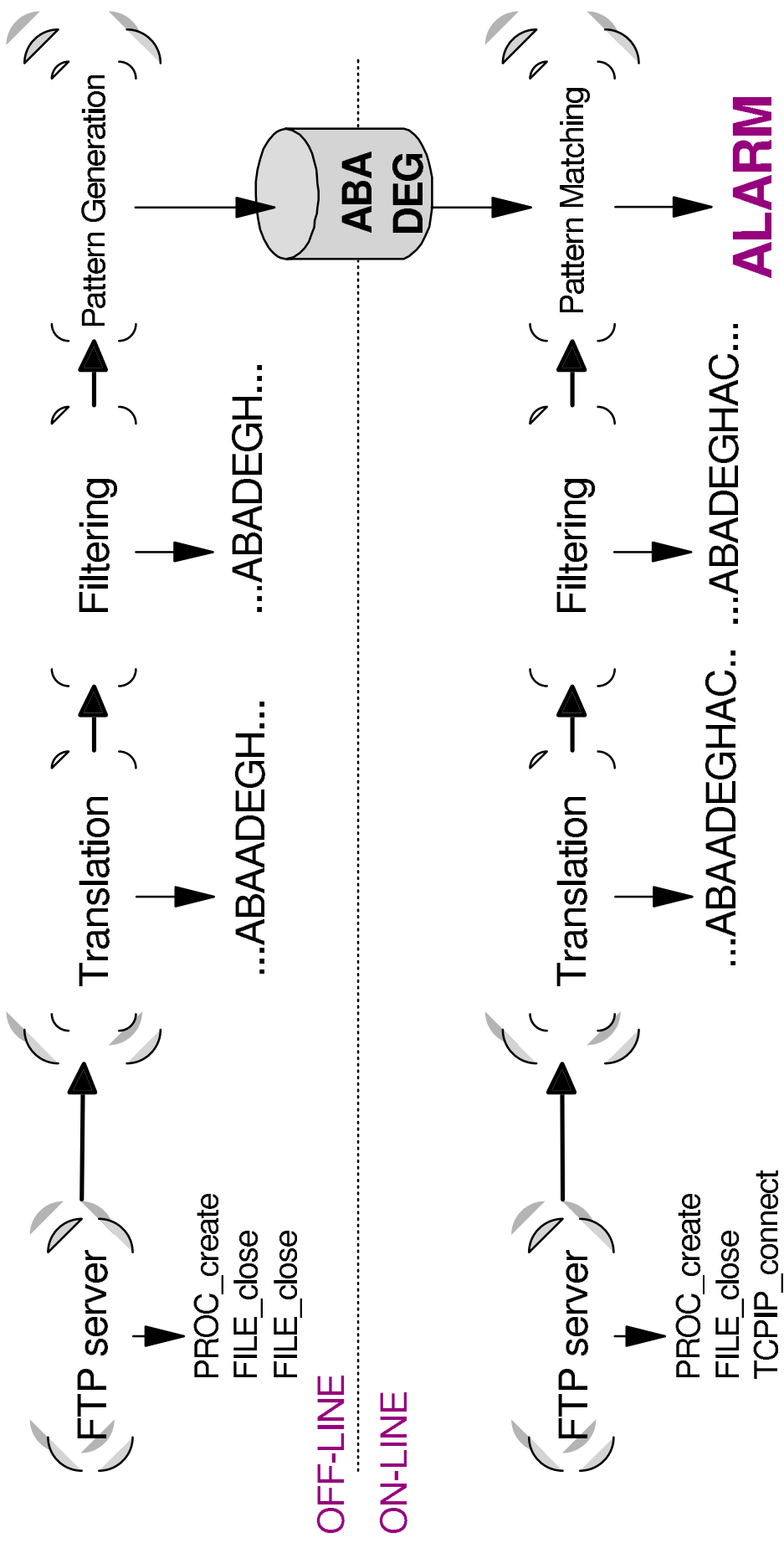


Principle

- A UNIX process is characterized by the audit events / system calls it generates.
- Different invocations of the same process have common subsequences of audit events (patterns).
- The patterns can be used to model the normal behavior of a given process.
- Intrusions are assumed to exercise abnormal paths in the executable code. The corresponding audit events cannot be covered by any patterns.



Architecture





Pattern Matching

- Juxtaposed patterns.
- Criterion to raise an alarm is the number of consecutive uncovered events (e.g. 3).

Pattern table:

F	D	E
B	C	

Match:

B	C	F	D	E	J	K	B	C
---	---	---	---	---	---	---	---	---

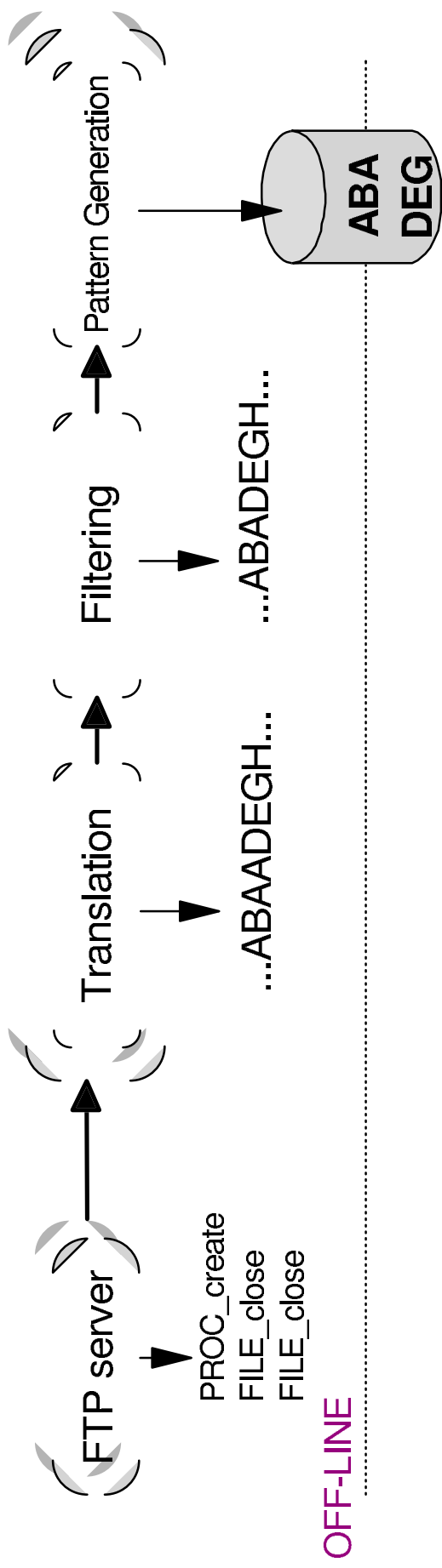
ALARM:

B	C	F	D	E	J	K	L	B	C
---	---	---	---	---	---	---	---	---	---



PROBLEM

How to Build the Pattern Table?





How to Build the Pattern Table?

- Different approaches to train a system:
 - Learn from a running system.
 - Compose manually some test cases.
 - Use Functional Verification Tests (FVT) as provided by software developers.
- Determine the patterns which
 - are typical for a process, and
 - can be used by the pattern matching process to differentiate between normal and abnormal behavior.



PROBLEM

Fixed-Length Patterns

Audit Sequences

A B D E C D D A

C D A B E



C D A B A B

Pattern Table

3	C	D	
3	A	B	
1	D	E	
1	D	A	
1	A	B	E



number of
patterns



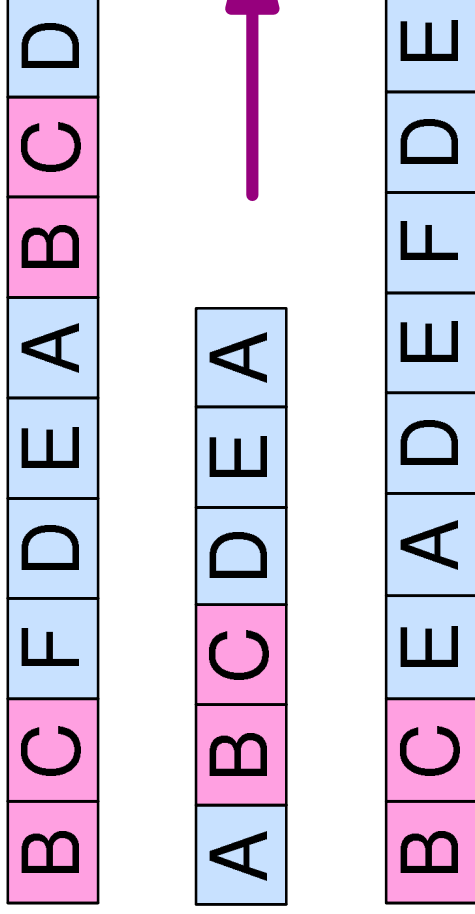
What Is the Ideal Pattern Length?

- A small pattern table helps to speed up the pattern matching process.
- Short patterns result in a small pattern table, long patterns result in a large pattern table.
- Long patterns seem to be process-specific, whereas short patterns may occur in several unrelated processes.
- Preliminary results show that long patterns produce more false positives than short patterns.



Variable-Length Patterns

Audit Sequences



Pattern Table

4	B	C		
4	D	E		
3	E	A		
2	A	B	C	D
2	A	B	C	
2	B	C	D	
2	D	E	A	
2	F	D	E	
2	A	B		
2	F	D		



APPROACH

First Reduction Step

Audit Sequences

B C F D E A B C D

A B C D E A

B C E A D E F D E

All Patterns

4	B	C			
4	D	E			
3	E	A			
2	A	B	C	D	
2	A	B	C		
2	B	C	D		
2	D	E	A		
2	F	D	E		
2	A	B			
2	F	D			

Maximal Patterns

4	B	C			
4	D	E			
3	E	A			
2	A	B	C	D	
2	D	E	A		
2	F	D	E		

Teiresias - a Novel Pattern Discovery Algorithm



APPROACH

Second Reduction Step

Audit Sequences

B C F D E A B C D

A B C D E A

B C E A D E F D E



Maximal Patterns

4	B	C		
4	D	E		
3	E	A		
2	A	B	C	D
2	D	E	A	
2	F	D	E	



Covering Patterns

4	B	C		
4	D	E		
3	E	A		
2	A	B	C	D
2	F	D	E	



Teiresias



Reduction Algorithm

B C F D E A B C D

A B C D E A

B C E A D E F D E

1)

B C F D E

E A

2)

B C E A D E F D E

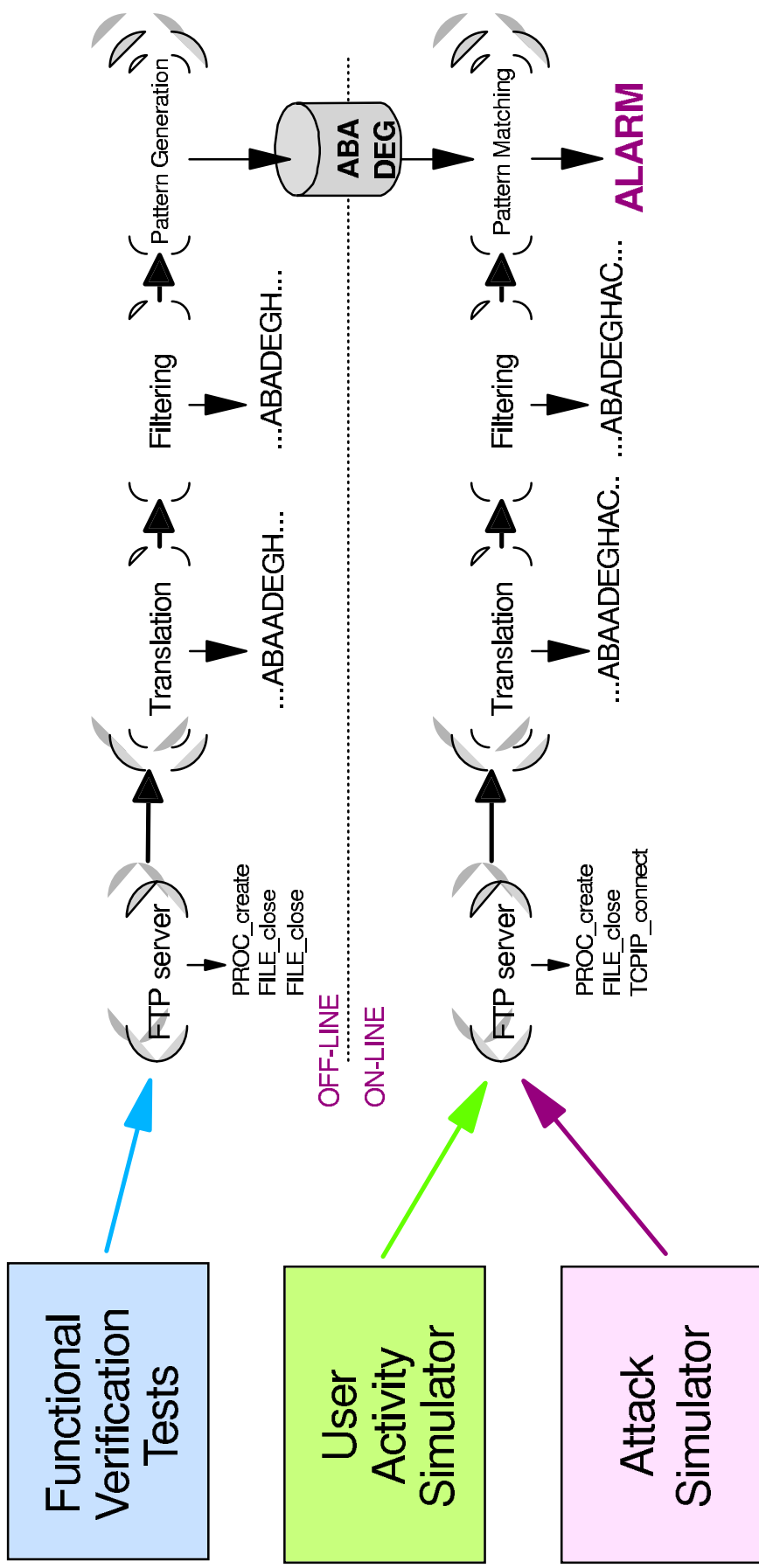
4	B	C				
2	D	E				
2	E	A				
8	A	B	C	D	X	
3	D	E	A			
3	F	D	E			

IrCoverage

4	B	C				
4	D	E				
2	E	A				
	A	B	C	D	X	
0	D	E	A			
6	F	D	E			X



Test Environment





Pattern Statistics

Process examined: ftpd

Test sequences	58
Events	23'302
All patterns	167'187
Maximal patterns	554
Covering patterns	71



Comparison Fixed-/Variable-Length

Process examined: ftpd

Method	Mean pattern length	Number of Patterns	Uncovered events (26'000)	False positives	Attacks detected (10)
Fixed length	3	114	185	0	8
Fixed length	4	152	437	4	9
Variable length	10	71	47	0	8



Conclusions

- We have developed a method to build tables of variable-length patterns. The tables have less but longer entries than tables of fixed-length patterns.
- Tables of variable-length patterns can be constructed that how a low number of uncovered events in normal user sessions but still detect (nearly) all attacks.
- The algorithm to create the pattern table, the pattern matching algorithm, and the alarm criterion are heavily interrelated.