Design and Implementation of Access Control Method Based on Correlation Among Files

*Yuki Kodaka(y_kodaka@nii.ac.jp), **Hirokazu Hasegawa, **Hiroki Takakura

*The Graduate University for Advanced Studies (SOKENDAI), Japan

**Center for Strategic Cyber Resilience R&D, National Institute of Informatics, Japan

- Name : Yuki Kodaka
- Affiliation : The graduate university for advanced studies(SOKENDAI), Japan
- Position : Student
- Contact information

    Email address : y_kodaka@nii.ac.jp

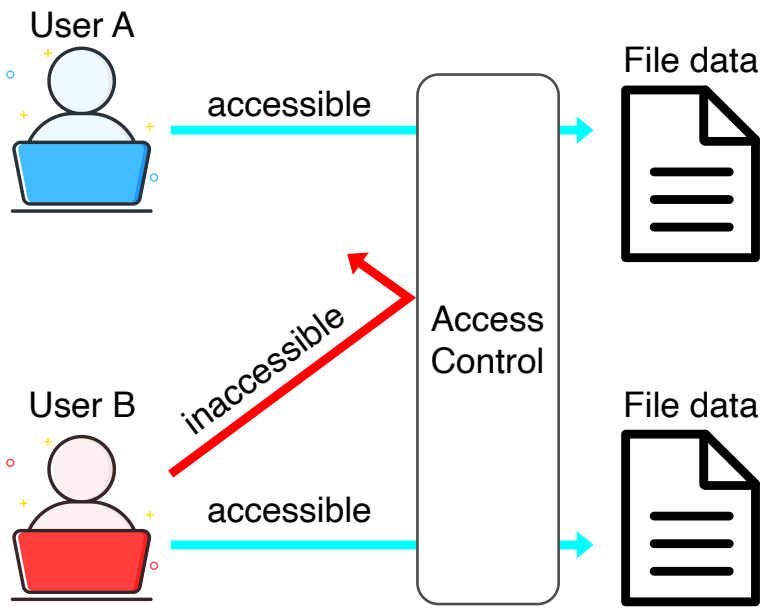
- Research interest :

    Access Pattern Analysis, Inter-file Correlation, Access Control Optimization

File Access Control

- File data = important asset in organization
  - Perspectives of Confidentiality and Business continuity
- Approaches to protect : Encryption・Backup・Access control
- Conventional : Coarse-grained access control due to management cost

-> Fine-grained access control is needed as situation changes



Criteria example for access control decisions

Role : Rank, Affiliation

Attribute : Time, Location

1. Allocation of minimum necessary file access privileges
   - Balance between over and under-access privileges
   - Manual operation = Unrealistic
2. Flexibility to adapt to variable file access demand
   - Demand for file access = Variable as situation change
3. Prevention of confidential information leakage through people

   Ex. Supervisor writes information in a file read by subordinates -> Leakage

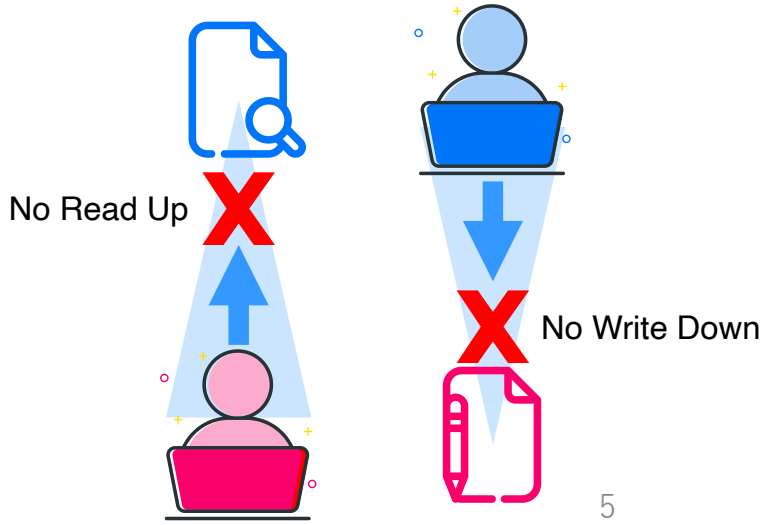
Access Control Method Based on Correlation among Files

Key idea : Criteria for access control decisions = Correlation

- Infer correlation among files from access behavior by users in the same group
- Control file access privileges based on correlation

- Control “Read” and “Write” privileges based on user rank

  cf. Bell-LaPadula model (No read up, No write down)



Create Correlation graph among files inferred from user access behavior

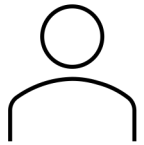
Determine whether a file is accessible or not

① Access history

User A

Timestamp,Filepath
2023-05-10T09:00:00,FileA
2023-05-10T09:15:00,FileB
2023-05-10T09:30:00,FileC
…

User B

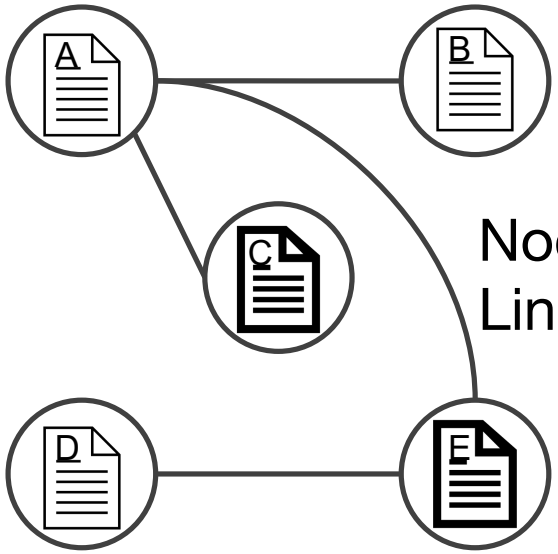


Timestamp,Filepath
2023-05-10T11:00:00,FileD
2023-05-10T11:15:00,FileE
2023-05-10T11:30:00,FileA
…

② File access order

FileA-FileB-FileC
FileD-FileE-FileA

③ Correlation graph among files



Node : File
Link : Correlation

Method

- Element of adjacency matrix exceed threshold

  -> two files are correlated

- Determination formula

(e.g., threshold = 0.8)

$$Matrix(File_{old}, File_{new}) \geq 0.8$$

Correlation between files=$Matrix(Fileold, Filenew)$
$Fileold$= File already accessed by User u
$Filenew$= File newly accessed by User u

Example of list of accessed file

User 1：A, D, E

User 2：C, E

Example of determination

File already accessed by User 2

File accessed newly by User 2

	A	B	C	D	E	F
A		0.3	0.8	0.9	0.9	
B	0.3		0.3	0.6	0.8	1.3
C	0.8	0.3				
D	0.9	0.6				
E	0.9	0.8				
F		1.3				

■ Revocation of access privileges

If no access is recorded for a certain period

-> Determine there is no need for access

User A

No access to file “f” for a certain period

Revoke

ACL for File “f”

~~Allow~~	~~User A：Read~~
Allow	User B：Write
Allow	User C：Read

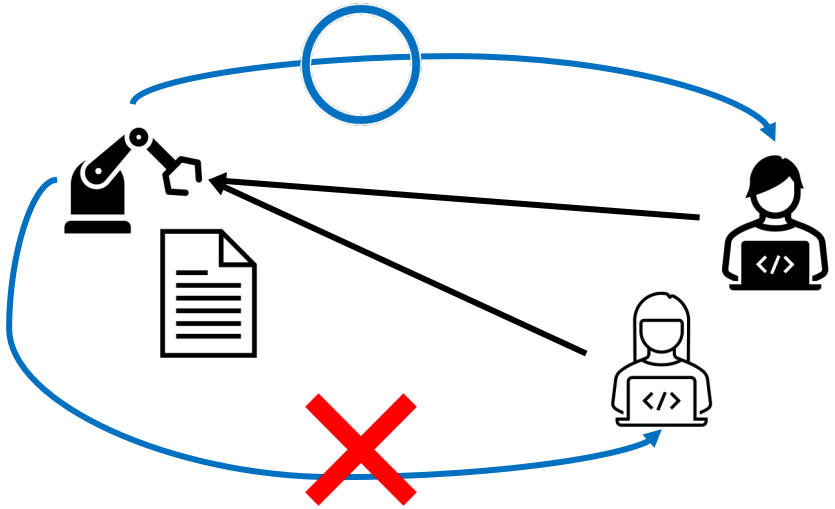
■ Addition of access privileges

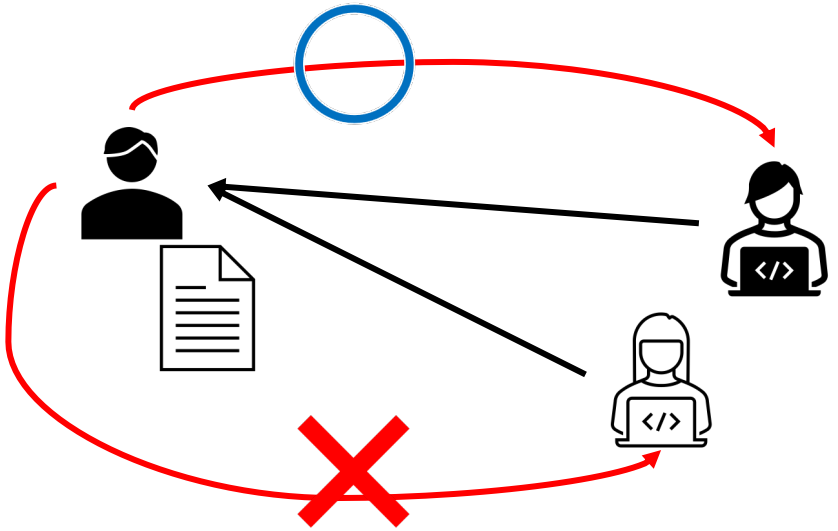
If user accesses unauthorized file

-> Automatic determination of privileges based on file correlation

If automatic determination : Denial

->manual determination by file owner





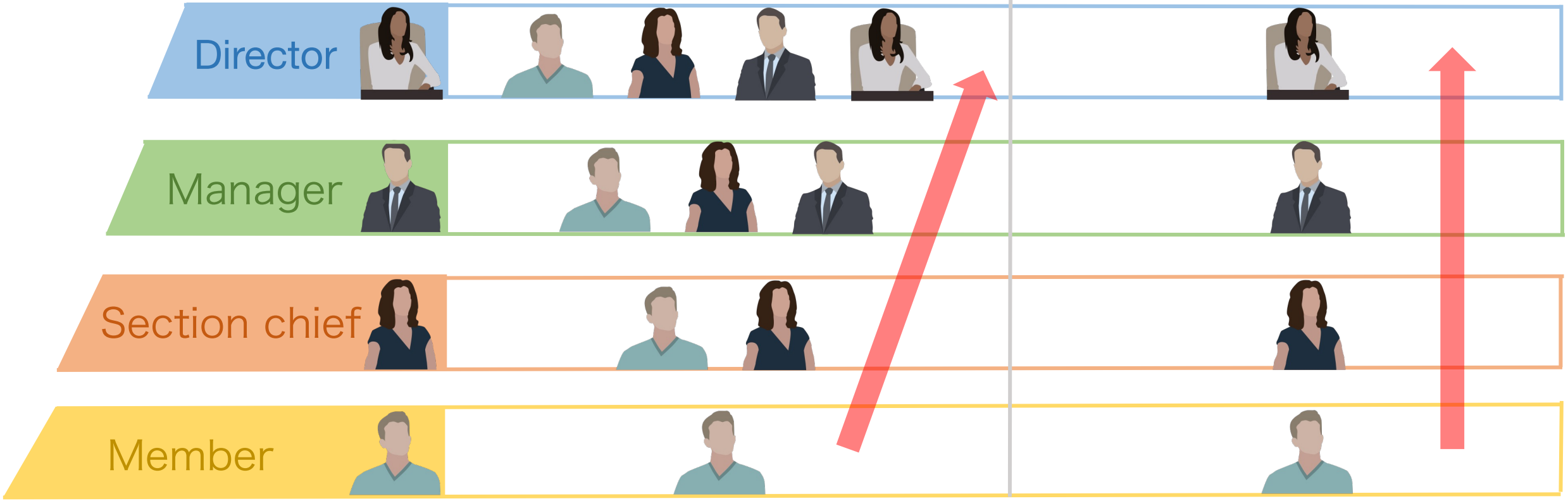
For Bell-LaPadula
-> 2 type graph
for each rank

“Read” determination graph

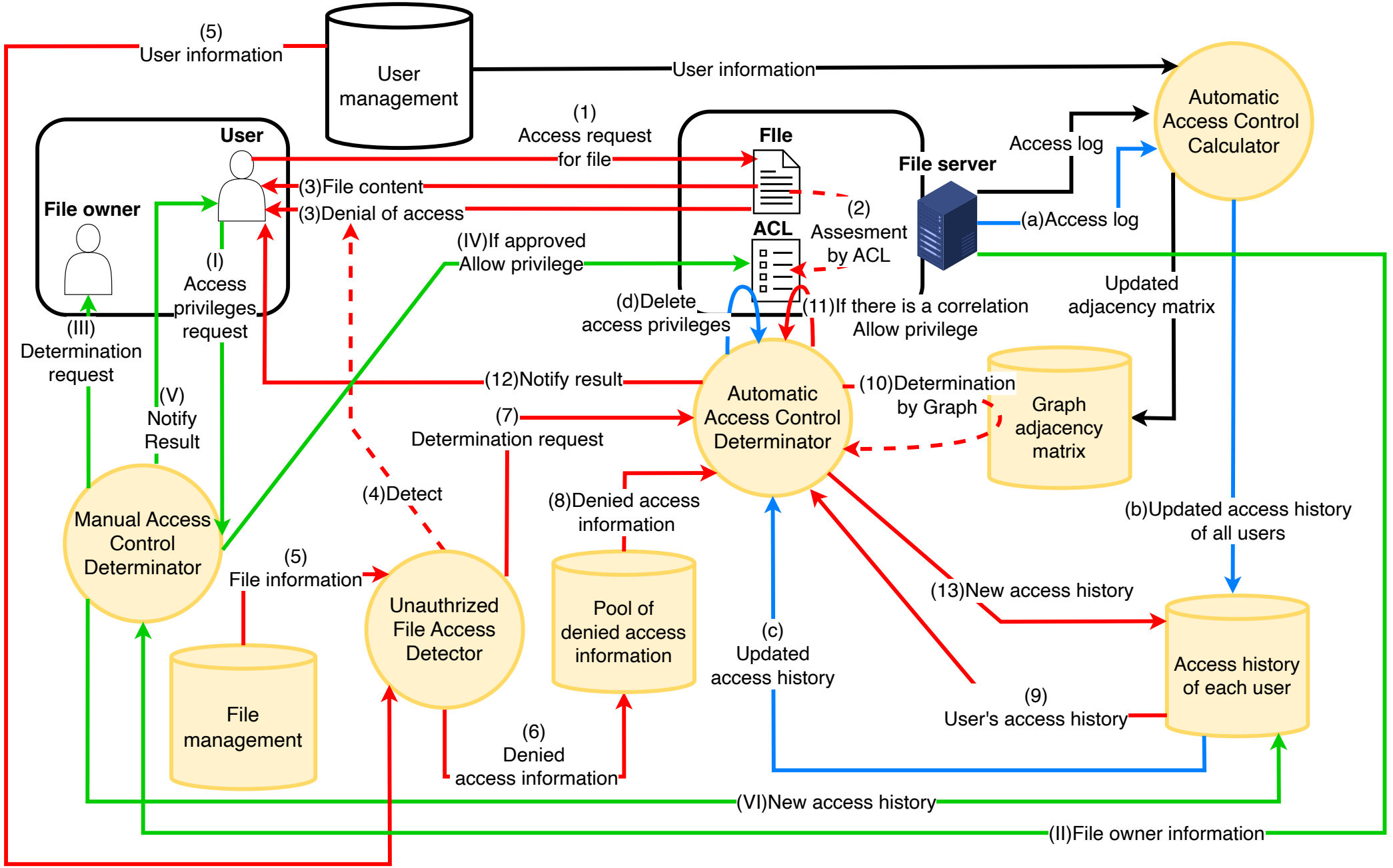
from “Read” access history
of same rank and below

“Write” determination graph

from “Write” access history
of same rank only



Architecture of System

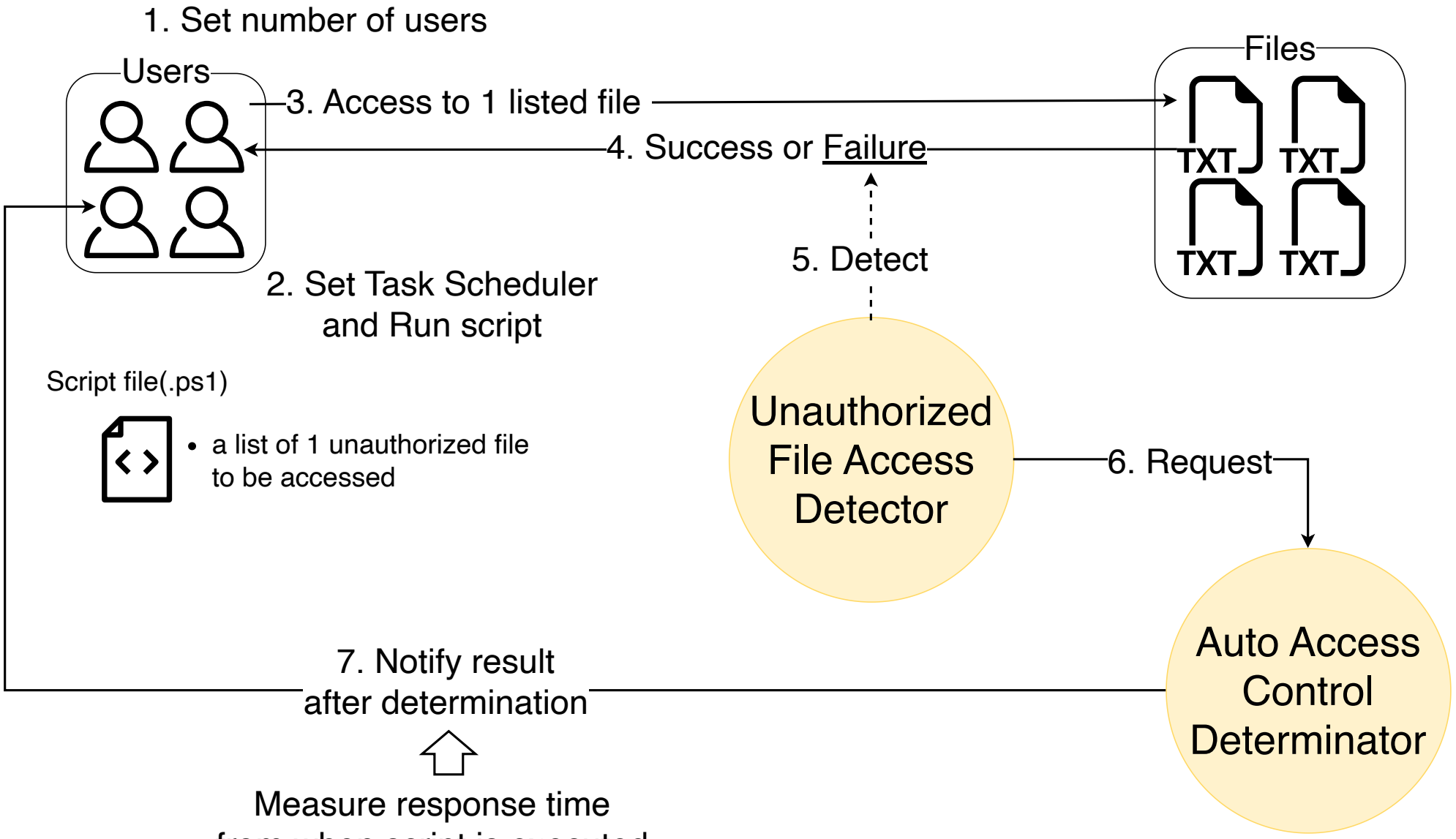


- Brief Implementation
  - Hardware : Intel® NUC 8 Pro Kit (NUC8v7PNH)
  - OS : Windows 11 pro edition
  - User : 6 general users and 3 users for proposed system
  - File : 10 files

- Verification of determination time by proposed system
  - Measure response time

    from when users accesses one unauthorized file

    to when determination results are notified to user


1. Set number of users
Users
3. Access to 1 listed file
4. Success or Failure
Files
TXT
TXT
TXT
TXT
5. Detect
2. Set Task Scheduler
and Run script
Script file(.ps1)
• a list of 1 unauthorized file
to be accessed
Unauthorized
File Access
Detector
6. Request
Auto Access
Control
Determinator
7. Notify result
after determination
Measure response time
from when script is executed
to when determination results are notified to user

Number of Users	Response time (seconds)						
	1st trial	2nd trial	3rd trial	4th trial	5th trial	5-trial Average	Average per User
1 (A)	6.85	6.22	6.62	5.96	6.69	6.47	**6.47**
2 (A/B)	7.96	7.92	6.51	6.71	7.82	7.38	**3.69**
3 (A/B/C)	7.91	9.56	7.34	9.49	8.58	8.58	**2.86**
4 (A/B/C/D)	8.61	8.26	8.35	8.79	8.16	8.43	**2.11**
5 (A/B/C/D/E)	8.79	8.95	9.56	9.10	9.46	9.17	**1.83**
6 (A/B/C/D/E/F)	11.20	12.77	9.91	12.53	11.84	11.65	**1.94**

- Average response time per user decrease

  as number of users increase

-> Efficiency of system improves after second case

<-> First determination = Bottleneck of efficiency

Limitation

Implementation Environment : Simplified

Verification experiment : Not high load

-> Although implementation of designed system was possible

Need to verify system scalability in more practical environment

- Propose Access Control Method based Correlation among Files
- In brief environment,

  Verify determination time by proposed system
- Results show

  First determination is bottleneck of efficiency of system

-> Future work should address above issues

  and verify system scalability in more practical environment