NETWORK LOG ANALYSIS TOOL - USER MANUAL

SAÉ 1.05 - Data Processing Project
BUT Networks & Telecommunications - Year 1
IUT de Roanne - 2025/2026

TABLE OF CONTENTS

WHAT'S THIS ABOUT?

The Problem

We had a major network issue at the India production site - super slow connection, packets dropping everywhere, and the network administrator spotted **2 suspicious activities** but couldn't figure out what was causing them.

Traditional tools (WireShark, manual config checks) gave us nothing. Zero results.

The Solution

I built this toolkit to analyze tcpdump captures automatically. It has two versions:

- **Python scripts** - Powerful, automated, professional-grade
- **Excel VBA macro** - For users who prefer working in spreadsheets

Both detect the same 3 types of network attacks:

1. **SYN Flood** - DDoS attempts
2. **Port Scanning** - Network reconnaissance
3. **Traffic anomalies** - Suspicious packet patterns

This tool was specifically designed to identify the two suspicious activities reported on the India site.

SETUP (DO THIS FIRST)

Python Version Requirements

Operating System: Windows 10+, Linux, macOS
Python: Version 3.8 or higher
Required Library: pandas

Installation Steps

1. **Install Python** (if you don't have it):
   - Go to python.org and download
   - **IMPORTANT:** Check "Add Python to PATH" during installation
2. **Install pandas** (open cmd/terminal):
3. pip install pandas

Takes about 30 seconds.

4. **Verify installation:**
5. python --version
6. # Should show: Python 3.x.x
7.
8. pip list | grep pandas
9. # Should show: pandas x.x.x

Excel VBA Requirements

Microsoft Excel: 2016 or newer
Macros: Must be enabled (see setup below)
Operating System: Windows only

PYTHON VERSION - QUICK START

What You're Doing

The Python script reads your tcpdump .txt file, parses it, detects anomalies, and generates 3 report files automatically.

Steps

1. Download the Project

Clone from GitHub or download ZIP:

```
git clone https://github.com/Mehdi-pxl/sae105.git

cd sae105-network-analysis
```

2. Navigate to Scripts Folder

```
cd scripts
```

3. Run the Analysis

Using default file location:

```
python analyse_reseau.py
```

Specifying your own file:

```
python analyse_reseau.py --fichier ../data/DumpFile.txt
```

Or shorter version:

```
python analyse_reseau.py -f /path/to/your/file.txt
```

4. Watch the Output

The console will show real-time progress:

```
============================================================
  SAÉ 1.05 - NETWORK LOG ANALYSIS (TCPDUMP)
============================================================


[INFO] Reading file: ../data/DumpFile.txt

[INFO] 1523 valid lines parsed

[INFO] 245 hexadecimal lines ignored


[OK] DataFrame created with 1523 rows


Preview of first 5 rows:

     Heure      IP_Source     IP_Dest  Port_Dest Flags Protocole

0  15:34:04   192.168.1.100  192.168.190.130   50019  [P.]     TCP

1  15:34:05   192.168.1.100  192.168.190.130   50020  [S]      TCP

...


[INFO] Anomaly analysis in progress...

[INFO] 2 anomaly(ies) detected
```

```
[OK] CSV report generated: ../rapports/DumpFile/rapport_20260115_143000.csv

[OK] JSON report generated: ../rapports/DumpFile/rapport_20260115_143000.json

[OK] Markdown report generated: ../rapports/DumpFile/rapport_20260115_143000.md


============================================================

            ANALYSIS COMPLETE

============================================================


1523 packets analyzed

2 anomaly(ies) detected

Reports saved in: ../rapports/DumpFile/
```

5. Check Your Results

Three files are created in rapports/[filename]/:

File	Purpose	Open With
rapport_YYYYMMDD_HHMMSS.csv	Excel spreadsheet	Excel, LibreOffice
rapport_YYYYMMDD_HHMMSS.json	Web interface data	Any text editor, web apps
rapport_YYYYMMDD_HHMMSS.md	Human-readable report	Markdown viewer, Notepad

EXCEL VBA VERSION - QUICK START

What You're Doing

The Excel macro imports your tcpdump file directly into Excel, creates formatted tables, detects anomalies, and generates a color-coded "Anomalies" worksheet.

Steps

1. Enable Developer Tab

1. Open Excel
2. **File** → **Options**
3. **Customize Ribbon**
4. Check the box **Developer**
5. Click **OK**

2. Import the VBA Code

1. Press **Alt + F11** (opens VBA editor)
2. **Insert → Module**
3. Copy ALL the VBA code from the artifact
4. Paste it into the module window
5. Press **Ctrl + S** to save
6. Press **Alt + Q** to close VBA editor

3. Enable Macros

When you open the file, you'll see a yellow security warning:

- Click **Enable Content**

4. Run the Analysis

Method 1: Developer Tab

1. Click **Developer** tab
2. Click **Macros**
3. Select AnalyserLogsReseau
4. Click **Run**

Method 2: Keyboard Shortcut

1. Press **Alt + F8**
2. Select AnalyserLogsReseau
3. Click **Run**

5. Select Your File

A window pops up:

1. Navigate to your tcpdump .txt file
2. Select it
3. Click **Open**

6. Wait for Processing

You'll see status messages:

- "Reading file in progress..."
- "Lines processed: 100, 200, 300..."
- "Analysis: SYN Flood Attack..."
- "Analysis: Port Scan..."

For large files (10,000+ lines), this takes 2-3 minutes. **Be patient!**

7. View Results

After processing, you'll see:

Main Worksheet:

- Complete log data in a formatted table
- Columns: Heure, Source, Destination, Port, Flags
- Automatic borders and colors

Anomalies Worksheet:

- Color-coded security alerts
- Red = CRITICAL (immediate action)
- Orange = HIGH (urgent review)
- Yellow = MEDIUM (monitor)
- Summary box showing total packets and anomalies detected

8. Export to CSV (Optional)

If you need a CSV file:

1. Press **Alt + F8**
2. Select ExporterCSV
3. Click **Run**
4. Choose save location
5. Click **Save**

📖 UNDERSTANDING THE OUTPUT

CSV File Structure

The generated CSV has these columns:

Column	Description	Example
Heure	Timestamp (HH:MM:SS)	15:34:04
Source	Full source address	BP-Linux8.ssh
IP_Source	Source IP or hostname	BP-Linux8
Port_Source	Source port	ssh (or 22)
Destination	Full destination address	192.168.190.130.50019
IP_Dest	Destination IP	192.168.190.130

Column	Description	Example
Port_Dest	Destination port	50019
Flags	TCP flags	[P.], [S], [F]
Protocole	Network protocol	TCP, UDP, DNS

Common TCP Flags

Flag	Meaning	What It Does
S	SYN	Initiates connection
.	ACK	Acknowledges data
P	PSH	Push data immediately
F	FIN	Closes connection
R	RST	Resets connection
U	URG	Urgent data

Common Ports (Quick Reference)

Port	Service	Description
22	SSH	Secure shell (remote login)
80	HTTP	Web traffic (unencrypted)
443	HTTPS	Web traffic (encrypted)
53	DNS	Domain name lookups
25	SMTP	Email sending
3389	RDP	Windows remote desktop

Markdown Report Structure

The .md file includes:

1. **Analysis Summary**
   - Total packets analyzed
   - Number of anomalies detected
   - Time range
2. **Security Alerts Table**
   - Severity level (with emoji)

  - Anomaly type
  - Source IP
  - Detailed statistics

3. **Detailed Descriptions**
   - Full explanation of each detected anomaly
4. **Recommendations**
   - Actionable steps based on severity
   - Critical: Immediate blocking required
   - High: Urgent investigation
   - Medium: Monitor closely

WHAT TO LOOK FOR (ANOMALIES)

1. SYN Flood Attack (SERIOUS)

What It Is: An attacker sends massive amounts of SYN packets to overwhelm the target server. It's a type of DDoS attack.

Detection Criteria:

- More than **100 SYN packets** from a single IP
- Severity: **CRITICAL** if > 1,000 packets

Example Output:

Type: SYN Flood Attack

Source IP: 192.168.1.100

Packet Count: 1,245

Severity: CRITICAL

Description: IP 192.168.1.100 sent 1,245 SYN packets (flood attack)

What to Do:

1. **Block the source IP immediately** in your firewall
2. Enable SYN cookies on affected servers
3. Contact network security team
4. Document the timestamp and IP for incident report

---

2. Port Scanning (MEDIUM TO SERIOUS)

What It Is: An attacker probes multiple ports to identify vulnerable services. This is reconnaissance before an actual attack.

Detection Criteria:

- More than **10 different ports** targeted by a single IP
- Severity: **CRITICAL** if > 50 ports

Example Output:

Type: Port Scan

Source IP: 10.0.0.25

Ports Scanned: 67

Severity: CRITICAL

Description: IP 10.0.0.25 scanned 67 different ports (network recon)

What to Do:

1. Investigate the source IP (internal or external?)
2. Review firewall logs for the same IP
3. Enable port scan detection on IDS/IPS
4. If internal: Check if machine is compromised
5. If external: Block and report to ISP

---

3. No Anomalies Detected (GREEN)

What It Means: The traffic patterns look normal based on the thresholds.

But Still Check:

- Review the Top 10 IPs manually in the CSV
- Sometimes issues are subtle (not caught by thresholds)
- Verify the time range covers peak usage hours

🔧 TROUBLESHOOTING

Python Issues

Problem: "File not found" Error

Symptoms:

[ERROR] The file '../data/DumpFile.txt' does not exist!

Solutions:

1. Make sure you're running the script from the scripts/ folder
2. Verify the file exists: ls ../data/ (Linux/Mac) or dir ..\data\ (Windows)
3. Try absolute path: python analyse_reseau.py -f C:\full\path\to\file.txt

---

Problem: "No module named 'pandas'"

Symptoms:

ModuleNotFoundError: No module named 'pandas'

Solution:

1. Install pandas: pip install pandas
2. If using virtual environment, activate it first:
   - Windows: venv\Scripts\activate
   - Linux/Mac: source venv/bin/activate
3. Verify: pip list | grep pandas

Problem: No Valid Data Found

Symptoms:

[ERROR] No valid data found in the file!

Possible Causes:

- The file is empty
- Wrong file format (not tcpdump)
- All lines are hexadecimal dumps

Solution:

1. Open the file in a text editor (Notepad, Sublime, VSCode)
2. Check if it contains lines like this:
3. 15:34:04.766656 IP 192.168.1.100 > 192.168.190.130.50019: Flags [P.]
4. If it looks completely different, you might have the wrong format
5. Make sure you're using tcpdump output, not WireShark

Problem: Script Takes Forever

Symptoms:

- Script runs for 10+ minutes

- No progress shown

Solution:

1. Your file is probably huge (> 100,000 lines)
2. Try analyzing just a smaller time window first
3. On Linux/Mac, extract 1000 lines: head -1000 DumpFile.txt > test.txt
4. Analyze the smaller file to test

Excel VBA Issues

Problem: Macros Are Disabled

Symptoms:

- Yellow security warning appears
- Macro buttons don't work

Solution:

1. Click **Enable Content** in the yellow bar at the top
2. If that doesn't work:
    - File → Options → Trust Center
    - Trust Center Settings
    - Macro Settings
    - Select "Enable all macros" (temporarily)

Problem: "Compile Error: Can't Find Project or Library"

Symptoms:

Compile error: Can't find project or library

Solution:

1. Press **Alt + F11** (VBA editor)
2. Tools → References
3. Look for items marked **MISSING**
4. Uncheck any missing references
5. Make sure **Microsoft Scripting Runtime** is checked
6. Click OK and try again

Problem: Excel Freezes During Processing

Symptoms:

- Excel shows "Not Responding"
- Task Manager shows high CPU usage

Solution:

1. **Wait patiently** - Processing 10,000+ lines takes 2-3 minutes
2. Don't click anything or Excel might crash
3. For very large files (> 50,000 lines), use the Python version instead
4. Close other applications to free up RAM

Problem: CSV Shows Everything in Column A

Symptoms:

- All data appears in a single column
- Can't read the data properly

Solution:

1. Select column A
2. **Data** tab → **Text to Columns**
3. Select **Delimited**
4. Check **Semicolon** (NOT comma)
5. Click **Finish**

FILE STRUCTURE

Here's how your project should be organized:

```
sae105-network-analysis/
├── data/
│   └── DumpFile.txt          # Your tcpdump file
├── scripts/
│   └── analyse_reseau.py       # Python script
├── rapports/                 # Output folder (auto-created)
│   └── DumpFile/             # Subfolder per file
│       ├── rapport_20260115_143000.csv
│      ├── rapport_20260115_143000.json
│     └── rapport_20260115_143000.md
├── vba/
│  └── AnalyseReseauVBA.bas     # VBA code
└── README.md                   # Project documentation
```

Keep this organized! You might need to go back to old captures later.

MODIFYING DETECTION THRESHOLDS

If you want to adjust when anomalies trigger (too sensitive or not sensitive enough):

In Python Script

1. Open analyse_reseau.py in any text editor
2. Find the function detecter_anomalies() (around line 140)
3. Change these values:

```
# SYN Flood threshold
if nb_paquets > 100:  # Default: 100
   severite = "CRITIQUE" if nb_paquets > 1000 else "ÉLEVÉE"


# Port Scan threshold
if nb_ports > 10:  # Default: 10
   severite = "CRITIQUE" if nb_ports > 50 else "ÉLEVÉE"
```

In Excel VBA

1. Press **Alt + F11** (VBA editor)
2. Find the DetecterAnomalies subroutine
3. Change these lines:

```
' SYN Flood threshold
If dicSYN(cle) > 100 Then  ' Default: 100
   If dicSYN(cle) > 1000 Then  ' Default: 1000


' Port Scan threshold
If dicPorts(cle).Count > 10 Then  ' Default: 10
   If dicPorts(cle).Count > 50 Then  ' Default: 50
```

Recommendations

Lower values = More sensitive

- Catches more issues
- More false positives
- Good for high-security environments

Higher values = Less sensitive

- Only catches serious issues
- Fewer false positives
- Good for busy networks with legitimate high traffic

Test your changes on a known-good capture first!

DEPLOYMENT NOTES (FOR INDIA TEAM)

System Requirements

Hardware:

- CPU: Dual-core 2.0 GHz minimum
- RAM: 4 GB minimum (8 GB recommended for large files)
- Disk: 500 MB free space
- Network: Not required (analysis is offline)

Software:

- OS: Windows 10+, Ubuntu 20.04+, or macOS 10.14+
- Python 3.8 or higher
- pandas library
- Admin/sudo access for installation

Installation Steps (Linux/Ubuntu)

```
# Update package manager
sudo apt update

# Install Python 3 and pip
sudo apt install python3 python3-pip -y

# Install pandas
```

```
pip3 install pandas


# Verify installation

python3 --version

pip3 list | grep pandas
```

Transfer Files

Option 1: Git Clone

```
git clone https://github.com/Mehdi-pxl/sae105.git

cd sae105-network-analysis
```

Option 2: Direct Download

1. Download ZIP from GitHub
2. Extract to /opt/network-analysis/
3. Set permissions: chmod +x scripts/analyse_reseau.py

Testing

1. **Test with small capture** (1,000 lines):
2. cd scripts
3. python3 analyse_reseau.py -f ../data/test_small.txt
4. **Verify output files** are created in rapports/
5. **Check for errors** in console output

Automation (Optional)

Create a cron job to analyze captures automatically:

```
# Edit crontab

crontab -e


# Add this line (runs daily at 2 AM)

0 2 * * * /usr/bin/python3 /opt/network-analysis/scripts/analyse_reseau.py -f /path/to/daily-capture.txt
```

Adjust Thresholds

Indian network profile might differ from French site. Monitor for **1 week** and adjust thresholds if you see:

- Too many false positives → Increase thresholds
- Missing real attacks → Decrease thresholds

Support

For deployment issues, contact:

- **Technical Lead:** [Mehdi Moumite]
- **Email:** [mehdi.moumite@etu.univ-st-etienne.fr]
- **GitHub Issues:** https://github.com/Mehdi-pxl/sae105.git

TECHNICAL SUPPORT

Contact Information

For SAÉ 1.05 Academic Questions:

- Instructor: [Mehdi Moumite]
- Email: [mehdi.moumite@etu.univ-st-etienne.fr]

For Tool Issues:

- GitHub Repository: https://github.com/Mehdi-pxl/sae105.git
- Create an issue with:
  - Clear problem description
  - Error messages (screenshot or copy-paste)
  - Your OS and Python/Excel version
  - Steps to reproduce

Before Contacting Support

1. Check this manual's Troubleshooting section
2. Search GitHub Issues for similar problems
3. Verify you followed installation steps correctly
4. Try with a small test file first

Include This Info in Support Requests

- Operating System and version
- Python version (python --version)
- pandas version (pip list | grep pandas)
- Error message (full text)
- File size and number of lines
- What you were trying to do
- What you expected vs. what happened

ADDITIONAL RESOURCES

Documentation

- **Pandas Official Docs:** https://pandas.pydata.org/docs/
- **Python Regex Guide:** https://docs.python.org/3/library/re.html
- **Tcpdump Manual:** https://www.tcpdump.org/manpages/tcpdump.1.html

Network Security

- **TCP Flags Explained:** https://www.keycdn.com/support/tcp-flags
- **SYN Flood Mitigation:** https://www.cloudflare.com/learning/ddos/syn-flood-ddos-attack/
- **Port Scanning Detection:** https://nmap.org/book/port-scanning.html

Excel VBA

- **Microsoft VBA Reference:** https://learn.microsoft.com/en-us/office/vba/api/overview/excel
- **Regular Expressions in VBA:** https://www.regular-expressions.info/vba.html

BEST PRACTICES

When Capturing Traffic

Do:

- Capture during normal business hours (get baseline)
- Keep original .txt files (in case you need to reprocess)
- Document what time range you captured
- Use filters to avoid capturing sensitive data

Don't:

- Don't capture passwords or authentication tokens
- Don't analyze from an untrusted source
- Don't run captures 24/7 (generates huge files)

When Analyzing

Do:

- Start with small time windows (1 hour) for testing
- Compare against baseline traffic patterns
- Cross-reference with firewall/IDS logs

- Document findings with timestamps and IPs

Don't:

- Don't panic if you see unknown IPs (investigate first)
- Don't edit CSV manually (breaks the analysis)
- Don't block IPs without verification
- Don't ignore "green" results (still review manually)

When Responding to Alerts

Critical (Red) Anomalies:

1. Document the IP address and timestamp
2. Check if it's internal or external
3. Review recent firewall logs for that IP
4. Block immediately if confirmed malicious
5. Email security team with evidence
6. File incident report

High (Orange) Anomalies:

1. Note the IP address
2. Check if it's a known scanner (security audit?)
3. Monitor for 24 hours
4. Escalate if behavior continues

Medium (Yellow) Anomalies:

1. Add to watch list
2. Check daily reports
3. Investigate if pattern repeats

FINAL NOTES

Important Disclaimers

- This tool provides a **starting point** for investigation, not definitive proof
- Always cross-reference with other security tools (IDS, firewall logs)
- Thresholds are tuned for our network - adjust for yours
- False positives can happen - verify before taking action

What This Tool Can't Do

- It can't tell you if blocked traffic is legitimate or malicious
- It can't prevent attacks (it's a detection tool)
- It can't analyze encrypted traffic (HTTPS contents)
- It can't work on live traffic (use tcpdump first)

Data Privacy

- Be careful with sensitive data in captures
- Don't share raw tcpdump files externally
- Follow your company's data retention policies
- Anonymize IPs in reports if needed for external sharing

VERSION HISTORY

Version	Date	Changes
1.0.0	January 2026	Initial release with Python and VBA support

---

End of User Manual

January 2026
Created for SAÉ 1.05 project - BUT R&T