Strathclyde Water Processing (SWaP) testbed

Matthew Calder, Mujeeb Ahmed, Shishir Nagaraja
{first.last}@strath.ac.uk
Strathclyde Security Group

Introduction

This poster presents the Strathclyde Water Processing (SWaP) testbed, an industrial control system (ICS) for security research. SWaP will be used to (a) assess the calibration security of industrial IoT protocols, (b) design and test the defence technologies, (c) assess the effectiveness of side-channel based defences, and (d) share the datasets with the security community. We designed and built the testbed ourselves using a combination of industrial and off-the-shelf components to provide the greatest flexibility and accessibility.

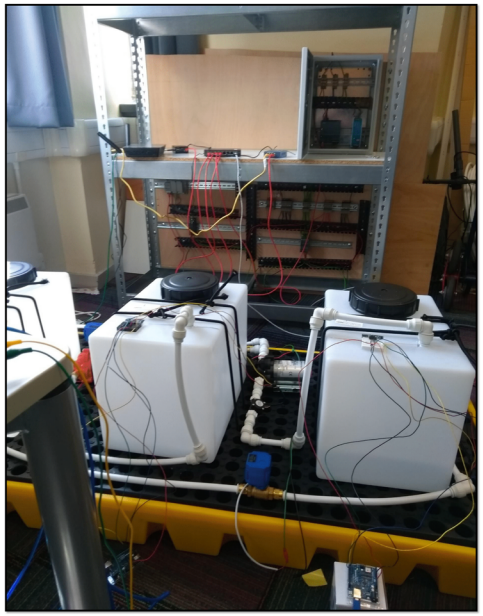
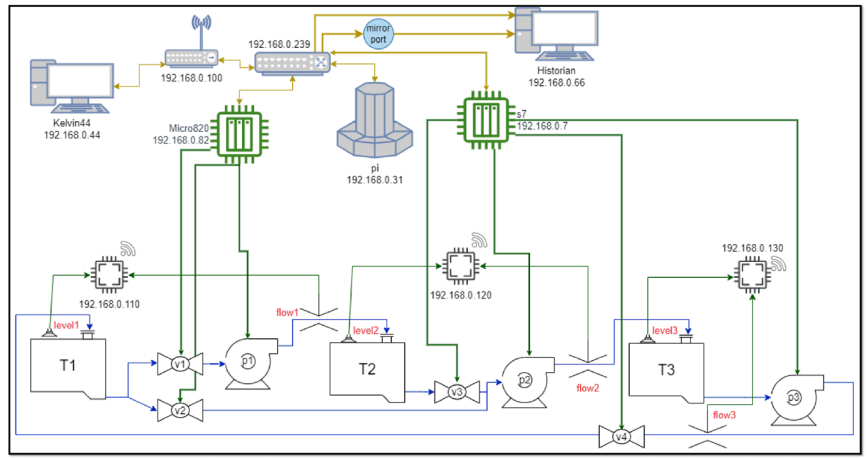


Photo of the testbed showing water tanks and control systems in the background

Testbed Design

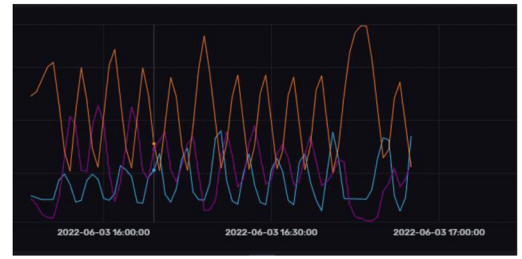
SWaP consists of a 3-stage water process autonomously controlled by 2 PLCs. The local communications between sensors, actuators, and PLCs are realized through wired and wireless channels.
Each tank has a flow and level sensor connected to an Arduino that sends the readings to a Raspberry Pi Remote-IO which collects them and sends them on to each PLC. The PLCs then use these readings to make decisions and control the process.
Each stage has a motorized valve and a pump (stage 1 has an extra valve) which are activated by relays connected to output pins of the PLCs.
Currently, tank 3 acts as the consumer which fills fully before emptying completely while tank 1 and 2 are providers which try to remain full while supplying tank 3.
The testbed is highly modular and is intended to be easy to reconfigure and extend in the future .



Physical and networking layout of the testbed showing wiring and water piping

Data Collection

The historian regularly queries each PLC for their status, pin and sensor values and stores them in a time series database. The switch also provides a copy of all network traffic during the run.
We have a dataset of normal operation and datasets for a series of different attacks on the process such as sensor manipulation to cause tank overflows.



Example data collection – showing the water depth in each tank over time

Future Work

PLC and scan cycle fingerprinting
Expanded equipment
- Additional protocols
- Additional sensors
- A variety of different PLCs and actuators
- More stages and processes

This work has been produced as part of the Robustness-as-Traceability (RoasT) project

RoasT Project

YASM: Yet Another Surveillance Mechanism

K Ludvigsen*, Shishir Nagaraja (PI), A Daly
Strathclyde Security Group
*: kaspar.rosager-ludvigsen@strath.ac.uk

University of Strathclyde Glasgow

Abstract

Many types of surveillance exist on anything from smartphones to IoT devices, but most of them are not as ubiquitous and intrusive as Client Side Scanning (CSS) for Child Sexual Abuse Material Detection (CSAMD). Apple proposed to scan their software and hardware for such imagery. While CSAMD was since pushed back, the European Union has decided to propose forced CSS to combat and prevent child sexual abuse via a new regulation, and deliberately weaken encryption on all messaging services. CSS represents mass surveillance of personal property, in this case pictures and text, proposed by Apple without proper consideration of privacy, cybersecurity and legal consequences. We first argue why CSS should be limited or not used at all, and briefly discuss some clear issues with the way pictures cryptographically are handled and how the CSAMD claims to preserve privacy. Afterwards, in the second part, we analyse the possible human rights violations which CSS in general can cause within the regime of the European Convention on Human Rights. The focus is the harm which the system may cause to individuals, and we also comment on the proposed European Union Regulation. We find that CSS by itself is problematic because they can rarely fulfil the purposes which they are built for. This comes to down to how even `perfect detection software' is inadequate to achieve automated CSAM detection, as seen with antivirus software. Secondarily, the costs for attempting to solve issues such as CSAM far outweigh the benefits, and this is not likely to change regardless of how the technology develops. We furthermore find the CSAMD as proposed is not likely to preserve the privacy or security in the way of which it is described in Apple's own materials. We also find that the CSAMD system and CSS in general would likely violate the Right to a Fair Trial, Right to Privacy and Freedom of Expression. This is because the pictures could have been obtained in a way that could make any trial against a legitimate perpetrator inadmissible or violate their right for a fair trial, the lack of any safeguards to protect privacy on national legal level, which would violate the Right for Privacy, and it is unclear if the kind of scanning which would be done here could pass the legal test which Freedom of Expression requires, making it likely violate this as well. Finally, we find significant issues with the proposed Child Abuse Regulation. This is because it relies on techno-solutionist arguments without substance, disregards conventional knowledge on cybersecurity and does not justify the independence and power of a ``centre'' to help solve the problem.

Preprint:
https://arxiv.org/abs/2205.14601

Client-side Scanning

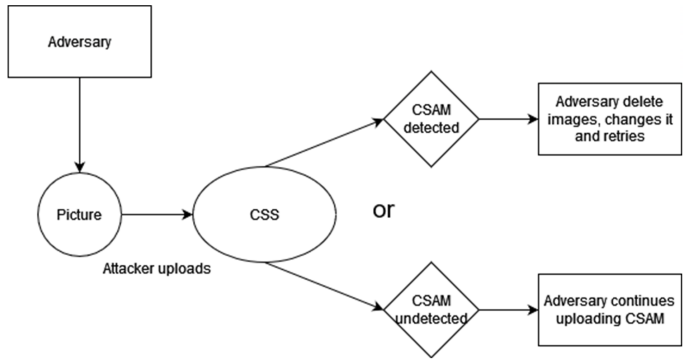
CCS is **scanning on the device**, while server-side scanning is the contrary. We propose a proof as to why perfect CSAM detection does not exist:

Proof

1. Assume that there exists a perfect CSS technique and that this is implemented into a system, and it could, e.g., be used to detect CSAM.
2. Assume there is a machine enforceable definition of CSAM which can be used to automate detection.
3. The CSAM poster can use a ***Sybil device (fake account on a test device)*** to test whether the injection of an image triggers detection. If detection is successful, the CSAM poster does not post it on their genuine device voiding detection. Therefore, perfect CSAM detection software does not guarantee the detection and isolation of CSAM posters.

CSS generally is both a backdoor and at the very least creates cybersecurity issues through its existence and its design. Furthermore, CSS cannot fulfil the goals they pose, because the things they are designed to catch (malware, CSAM) are impossible to perfectly define and therefore always identify. Proof 1 shows how circumvention is possible as well, and all of this combined shows that CSS should not be used over other, more certain or hybrid solutions. Malware is identified through a combination of tools, not just the antivirus, and CSAM must be stopped at its roots (people), not only when it is shared.

Figure 1 below shows the first circumvention from Proof 1.



European Human Rights Convention

Right to a Fair Trial

CSS will likely violate this right, as photos/messages/other data is:

1. Not independent of the individuals they come from (unlike blood, which is).
2. Require authorisation or warrants for every action.

No CSS currently has these rights. Violation of the Right to a Fair Trial may lead to the evidence being inadmissible, which means that the evidence cannot be used in a trial. This could lead to potentially guilty individuals escaping prosecution. Furthermore, it is a waste of police and investigatory resources.
CSS can also be used easily for entrapment or planted evidence, states or adversaries merely have to send or transfer files to devices to trigger them, potentially causing the investigation and prosecution of innocents.

Right to Privacy

CSS may not be used to interfere in the Right to Privacy, unless it is **"necessary in a democratic society"**, which is interpreted as including "pressing social need" in case law. This means that CSS or whoever uses it must justify what they do through this lens – systems such as the proposed CSAMD does not do this.
This is further exasperated since CSS will by its existence violate the Right to Privacy, as it is akin to an agent living in your house and constantly going through your belongings.

Issues with other rights

These systems will likely have the potential to violate both Freedom of Expression and Freedom of Assembly, if the dataset which is used to identify data with is changed to reflect for example **political parties, specific persons, location data** and more. We evaluate and find that CSS would find it very difficult to do this kind of surveillance without violating these rights in the paper.

Concluding Remarks

In the paper:

- We argued why CSS should not be used,
- Discussed the vulnerabilities of CSAMD specifically (not on this poster)
- Showed that they will like likely violate the Right to a Fair Trial, Right to Privacy and Freedom of Expression and perhaps Right to Assembly in the European Human Rights Convention.
- That the CSAM proposal by the European Commission has the risk of breaking all encryption, and like CSAMD, not actually lead to less CSAM (not on this poster).

From this, we recommend mainly that CSS should not be used in the first instance, and if you absolutely have to, you need to have legal (in statutes or equivalent) mandate to even run the system, and have enough safeguards to satisfy the European Court of Human Rights.

Legal Grounds for ROAST systems in the EU

Kaspar Rosager Ludvigsen
Kaspar.rosager-ludvigsen@strath.ac.uk

Abstract

The calibration of IoT and all other types of devices and cyberphysical systems is increasingly becoming more important, because of the role which these systems play in both our everyday lives and in industrial settings. Everything from surgical, industrial and military robots, critical infrastructure and industrial control systems - all rely on accuracy and the best calibration possible. Another layer to this is adversaries, who are interested in manipulating or otherwise abusing issues which wrongful calibration can cause. To accommodate these potentially increased dangers, we suggest that IoT systems (and others) going forward, should be Robust as Traceability (ROAST). ROAST entails an increased focus on auditing and always knowing, when, where, who, and what was calibrated. To accommodate this, we here show some examples of existing legal rules which can be used to mandate increased security and focus on calibration, for the safety and security of everyone going forward. In the final paper and here, we use two examples from EU law - GDPR and the Medical Device Regulation. For the purposes of practicality, we focus on calibration, traceability and verification. As this is part of the ROAST PETRAS project, further elaboration on how these ideas can be feasibly realised in the systems themselves will follow, but this part of the research show that it is possible to do, even within the current legal rules.

GDPR

Accuracy is a core concept the processing of personal data.

We see this in Article 5(1)(d):

> *1. Personal data shall be:*
> ...
> *(d) accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay ('accuracy')}\end{quote}*

GDPR only applies to personal data, but we know from practice and the economics surrounding data that this can be understood in an expanded manner, as it is very difficult to anonymise or pseudonomise personal data.

We interpret Article 5(1)(d) as including:

- **Calibration.**
- **Traceability.**
- **Verification.**

Accuracy cannot be obtained without these steps, meaning that the personal data which is gathered or which the system uses accuracy to gather will be inaccurate or outright wrong, leading to a violation of the GDPR, which can lead to fines. This incentivises calibration and so on, but does not include a definition of how these things could be done in practice, which means that standards and certifications play a role for the practical application.

Medical Device Regulation

This regulation does not have a central article we can cite, but we can provide arguments from several spots, including Article 5(1) and (2):

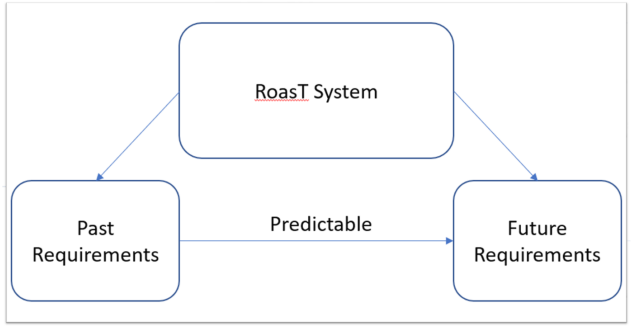
> *1. A device may be placed on the market or put into service only if it complies with this Regulation when duly supplied and properly installed, maintained and used in accordance with its intended purpose.*
>
> *2. A device shall meet the general safety and performance requirements set out in Annex I which apply to it, taking into account its intended purpose.*

Article 5(1) entails ``properly installed'' and ``maintained'' which we interpret as including calibration, traceability and verification.

This is further confirmed in Annex I, chapter 1, both in regards to the mandatory and enforceable risk management systems, where it is clearly stated that medical devices must function as intended and not harm anyone, users or patients alike. This cannot be reached with calibration, traceability and verification, and will be much easier to realise if the systems are ROAST as well.

Unlike the GDPR, product legislation in the EU like the Medical Device Regulation can lead to forceful withdrawals or bans of products, so the consequences for not calibrating, tracing and verifiying properly can be much more severe, such as product bans or forced withdrawals.

Figure 1.



Legal Based Evolution

ROAST systems must not only easily trace and reveal everything for auditing, inspections or lawsuits, but they should also be robust. Robustness, or synonyms or equivalents in other sciences such as dependable, entail both increased safety but also long term viability.

The economics of the IoT environment is directly against the environment of the planet - products are created and made obsolete (and not serviced) at increasingly ridiculous rates, which is why we suggest a new type of mechanism which can let these systems exist for longer, without being impacted by changing legal requirements**: Legal-based evolution**.

The idea is that ROAST IoT devices, or other systems, must be complaint with both current and future predictive legal requirements. The best example for this is privacy and the GDPR. There is no doubt that future legal requirements will entail more privacy and better protection of the individual, which is why better PETS or similar mechanisms can both fulfil current and future compliance. We can extrapolate this to both cybersecurity in general, as well as calibration, traceability and verification in general. We illustrate the idea in Figure 1.

Future Work

- Explore economic Costs of requiring ROAST systems over existing regulation frameworks
- Further explore how legal-based evolution can be implemented into cyberphysical systems in general.

Concluding Remarks

If we combine the idea of accuracy requiring really good calibration with specific product rules, for example the medical device regulation, which mandate the same but in a much more serious manner (to prevent injury), we get the best grounds for why ROAST should be the standard.

If done now, it will allow existing IoT systems to be compliant both in the present and in the future, via legal-based evolution, but also be beneficial for its users and those who are affected by them, while likely not requiring massive costs for the manufacturers.

Selected References

Ross Anderson. *Security engineering: a guide to building dependable distributed systems*. John Wiley & Sons, 2020.
Elisabetta Biasin and Erik Kamenjasevic. *"Cybersecurity of Medical Devices:Regulatory Challenges in the EU"*. In: *The Future of Medical Device Regulation: Innovation and Protection*. 2020.
Tanja Dorst et al. *"Metrology for the factory of the future: Towards a case studying condition monitoring"*. In: *I2MTC 2019 - 2019 IEEE International Instrumentation and Measurement Technology Conference, Proceedings 2019-May* (2019).
Sascha Eichstädt et al. *"Toward smart traceability for digital sensors and the industrial internet of things"*. In: *Sensors 21.6* (2021), pp. 1–15.

Traceability as Verifiability in Robotic Workflows: A Traffic Analysis Approach

Ryan Shah, Mujeeb Ahmed, Shishir Nagaraja
{first.last}@strath.ac.uk
Strathclyde Security Group

Abstract

Connected robots play a key role in automating industrial workflows, making use of sensors and other IoT devices to bring higher levels of efficiency and accuracy compared to humans. Due to the safety-critical nature of operation and the environment in which they operate, the potential for error and resulting liabilities is high. Thus, mitigating these errors and maintaining the highest possible level of operational safety is a key priority. To achieve this, an approach to verify the intent sent by a robot operator matches the corresponding robot actions – encapsulated within a continuous monitoring approach – is a potential solution. In this work, we investigate whether one can accurately predict (fingerprint) robot movements and reconstruct operational workflows by simply monitoring network traffic between a robot and its controller. Using a machine learning traffic analysis approach, we found that we can predict TLS-encrypted robot movements with around 60% accuracy, increasing to near perfect accuracy in realistic settings. This showcases as a potential solution for verifying intent and outcomes in robotics systems (traceability as verifiability).

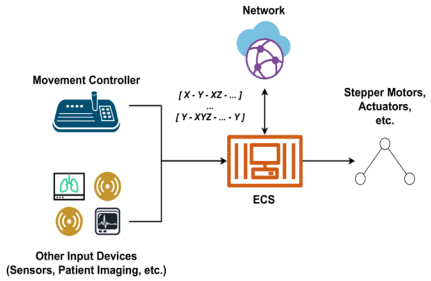


Fig. 1 – Teleoperated Robot Architecture

Introduction

Industrial robotics systems are prominent in various industrial applications and provide higher levels of accuracy and precision, compared to solely relying on human operators.

When Internet-connected, industrial robots are prone to a wide array of attacks to which several countermeasures have been deployed. As a standard, stricter security requirements are to be conformed to, such as the use of appropriate channel security technology (i.e. TLS) to protect confidentiality and integrity of transmitted data.

In this work, we aim to address the key question of whether, even under these strict measures, it is still possible to infer information from the traffic side channel about robot movements as an approach to achieving traceability as verifiability. Further, we question whether this information be used to recover high-level workflows in warehousing settings.

- This is important as collecting this information can be used to verify the intended actions (commands) issued by the controller/operator match the corresponding output from the robot.

System Design

The focus of the study is on teleoperated warehouse robots. To mimic this, we use a uARM Swift Pro operated via an Arduino Mega 2560. The controller runs on a Windows 10 laptop using the uARM Python SDK. The communication protocol follows a master-slave topology using an asynchronous TCP/IP socket. TLS traffic and realistic network simulations are propagated through a Software-Defined Network (SDN) (Fig. 1).

We programmed robot movements along X, Y, Z axis and combinations of these movements. We also programmed these movements with varying speeds and distances, and network parameters (link delay/jitter and packet loss). Finally, we programmed four warehousing workflows: push, pull, pick-and-place and packing operations. Our dataset contains around 150k samples.

The features we collected for which variations are observed include: *packet time*, *frame length*, *IP length*, *TCP length*, *bytes-in-flight*, *ack round-trip-time* and *TLS record length*.

From a workflow reconstruction approach, we identify 5 key research questions:
- Can we detect individual robot movements on each axis?
- Can we detect combinations of robot movements?
- Is it possible to reconstruct operational workflows that correspond to patterns of movements?
- How is movement fingerprinting affected by distance and speed of movement?
- How do realistic network conditions impact movement fingerprinting?

Initial observations from basic frequency analysis and eye-balling variations in packet features show that these simple traffic analysis approaches are not enough to answer our challenges. This led us to take a neural network approach to workflow reconstruction.

Our neural network consists of one input layer of 16 neurons for each of our features, one Dense hidden layer of 108 neurons using ReLU activation and Adam for optimisation, and an output layer using softmax activation with categorical cross-entropy loss (Fig. 2).

Movement	Precision	Recall
X	70%	85%
Y	69%	54%
Z	80%	63%
XY	21%	60%
XZ	68%	92%
YZ	81%	31%
XYZ	72%	97%

Table 1: Baseline Classification Results

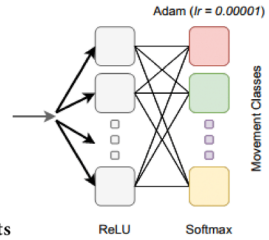


Fig. 2 – Neural Network Architecture

Operation	Recovery Rate	Pos Changes
Push	97%	2–3
Pull	97%	2–3
Pick-and-Place	84%	7–9
Packing	88%	6–9

Table 2: Workflow Reconstruction Results

Operation	Recovery Rate	Pos Changes
Push	45%	2–3
Pull	47%	2–3
Pick-and-Place	51%	7–9
Packing	48%	6–9

Table 3: Tor Procedure Reconstruction Results

Evaluation

Baseline

In accordance with our challenges, the first two aim to determine whether one can verify individual movements and combinations of movements. As a baseline (lowest speed and distance), we find that accuracy averages around 60% (Table 1). We found Y-based movements had lowest precision due to similarities in traffic features between them.

Workflow Reconstruction

For the third challenge, we examine whether workflows can be reconstructed from patterns of movements. Specifically, we look at push, pull, pick-and-place, and packing operations from warehousing workflows. Results are found in Table 2.
- Quantifying the accuracy of reconstruction demonstrates the ability to verify that entire robot workflows are conducted properly (i.e. within logistics supply chains).

We observe workflows can be recovered more accurately that individual movements, averaging around 90%.

Experimental Parameters

For the remaining challenges, we examine the impacts of movement speed and distance, packet loss and link delay. For both movement distance and speed, we found improvements compared to baseline (up to a certain point), with speed showing better accuracy than distance. As for network parameters, in both cases we observe near perfect recall under all tested conditions (delay of 10/50/100/1000ms or loss of 10/25/50%).

Countermeasures (Tor)

We also examine Tor as a potential hindrance to our verifiability approach. Given its success as a defence in other applications such as website fingerprinting, it is a reasonable consideration that Tor could be used to further strengthen the security of robotics deployments and thus, should also be evaluated to investigate its impact on verification accuracy.
- We use a Tor hidden service and monitor inbound traffic on the service host to capture robot movement traffic.
- We captured samples from multiple autonomous systems and over 20 random circuits.

Overall, we found that Tor averages around a 20% decrease in accuracy compared to the baseline. For workflow reconstruction, we find that Tor does also reduce recovery rate by at least a factor of 2, but accuracy is comparable with baseline results.
- Latency does not (overall) present with much overhead (<1s), but in safety-critical contexts (i.e. surgical robotics) this is less than desirable.

Future Work

- Exploration of other side-channels (i.e. acoustic, RF)
- Investigating impact of other security measures (i.e. padding/mixing) on accuracy

This work has been produced as part of the **Robustness-as-Traceability** (RoasT) project