Threat Modelling with the GDPR towards a Security and Privacy Metrics Framework for IoT Smart-farm Application

Steph Rudd and Hamish Cunningham
Department of Computer Science, University of Sheffield, Sheffield, U.K.

Keywords: IoT, Smart-farms, GDPR, Privacy, DPIA, Energy Consumption, Risk Mitigation, Threat Modelling.

Abstract: This paper considers a balance between privacy and security provisions for IoT devices constrained by processing ability, energy consumption, and storage. Risk-driven testing is invoked in conjunction with the seven GDPR principles towards a metrics framework suitable for such an energy-conscious network within the domain of IoT-oriented smart-farms. The resulting metrics framework demonstrates how the influence of privacy can minimise processing requirements, whilst threat modeling assures security. The research concludes that several redundant security provisions can be replaced with privacy alternatives that improve energy efficiency.

1 INTRODUCTION

With IoT the aim is always to reduce power - often constrained devices rely on the limited power supply of batteries. Within the remit of IoT agriculture, this aim extends to enable crop growth automatically, without internet dependence, and supported by green energy towards sustainably powering constrained devices. Unfortunately, security implementation is a threat to device lifespan. It is heavy in both power and processing - a threat to the longevity of the environment, yet necessary to protect it from malicious and accidental downfall. This research proposes a solution to the security and threat balance - how to reduce the burden of security whilst protecting against downtime using privacy functions pertinent to the General Data Protection Regulation (GDPR), and Soft Systems Modelling (SSM), for a perspective beyond cryptography metrics.

The outcome is a framework with agnostic application to IoT environments exemplified in the field of aquaponic smart-farming, with the following contributions:

1. Content-driven: Vulnerabilities assessed on value and impact enabled selectivity in protection.
2. Energy reduction: The framework prioritised green-farming towards sustainability.
3. Worldview: SSM and GDPR provided valuable insight into atypical threat factors.
4. Substitution: Privacy functions ranked by energy use demonstrated GDPR application.
5. Additional attributes: Modelling using a world-view enabled domain-specific data attributes.

The study begins with explaining the issues pertaining to de facto standard Transport Layer Security (TLS), and vulnerabilities of IoT. Next, related work considers modelling methods suitable for the environment, risk-driven metrics and privacy substitution. Assessment and scoring is then defined before the framework is invoked for testing. The results demonstrate great benefit of such a framework for modelling security and privacy at the design stage by employing the principles of TLS, SSM, GDPR and threat modeling, but in a way to benefit and sustain IoT operations very different to traditional security application.

2 PROBLEM BACKGROUND

The problem is separated into an overlap between security and privacy, nature and vulnerabilities of IoT, and the paradox of TLS, as the de facto security standard for both traditional and constrained environments.

2.1 Security and Privacy

The misconceptions pertaining to the domain of privacy are likely due to the long-standing definitions of Confidentiality, Integrity, and Availability (CIA), belonging to security (Aminzade, 2018). Whereas security definitions CIA relate to encryption, attestation of

Rudd, S. and Cunningham, H.
Threat Modelling with the GDPR towards a Security and Privacy Metrics Framework for IoT Smart-farm Application.
DOI: 10.5220/0010920800003194
In *Proceedings of the 7th International Conference on Internet of Things, Big Data and Security (IoTBDS 2022)*, pages 91-102
ISBN: 978-989-758-564-7; ISSN: 2184-4976

non-tampering, and proof of origin respectively, privacy definitions refer to the rights of a living subject through anonymity, pseudonymisation, and tokenisation (Varanda et al., 2021). Anonymisation means the inability to identify a living person amongst a data set - a series of values exemplified by health data of a group, for example, where many subjects will belong to a group of age or gender, but none shall stand out as an individual. Anonymisation is often used in statistics for geography, health, demographics, and risk factors like insurance. Pseudonymisation, compared to anonymisation, does not alleviate data of all identifiable information - but reduces the link of a dataset with the original identity of an individual.

Pseudonymisation and anonymisation as privacy functions overlap with security, as they could mean an encryption scheme, or a nickname with meaning to operatives but not outsiders. They could also mean a reference code, an artificial identifier indexed against the real data in a central server. Tokenisation is also similar to encryption, but often operates on a simple substitution to disguise data such as credit card numbers. Tokenisation can therefore be ‘unlocked’ without a key, particularly with well known credit card tokenisation algorithms used as part of the Cardholder Data Environment (CDE) (El Alloussi et al., 2014), of PCI DSS (Razikin and Widodo, 2021).

The approval of PCI DSS for use of cryptographic alternatives such as pseudonymisation for sensitive data, indicates that negation of heavily mathematical functions is entirely safe, practical, and even encouraged under the governing bodies of financial regulation. This work investigates where it is safe to substitute security functions for privacy.

2.2 Nature and Vulnerabilities of IoT

IoT networks by nature are chaotic, constrained, and heterogeneous - opposite of the capabilities and resources of typical computers such as desktop machines and smartphones. However, particularly in the case of microcontrollers (as opposed to microcomputers), their constraint and simplicity provides natural protection. In the way that a calculator cannot be exploited through lacking architecture, a microcontroller is naturally without Operating System (OS) or the problems that go with it. Developments of microcontroller-oriented developments such as Serial Peripheral Interface Flash File System (SPIFFS) (Espressif, 2021a), Mongoose OS (Espressif, 2021b), and the Arduino webclient (Microcontrollerslab, 2019), increase the threat landscape by removing this natural robustness of IoT simplicity. Attempting to transform machines comparable with those of the 1990’s towards modern expectations, weighs heavily on the storage and energy restrictions. TLS is a reflection of such unreasonable expectations; it is centralised, rigid, and contradicts notions of heterogeneous freedom and spontaneity. However, TLS is strong, widespread, and available for most, if not all, communications protocols.

2.3 Rigidity of TLS

As the long-standing de facto standard in online security, TLS has carried ubiquity onto IoT, with libraries for both software and embedded devices, and dedicated hardware acceleration in some microcontrollers such as the ESP32 (ESP32, 2021). However, the benefits of TLS end at the convenience of ubiquity. Designing an agnostic security application to satisfy the same standards as TLS would be quite a challenge, particularly since the CIA functions are readily available. Therefore the challenge is not to reinvent TLS security, but disrupt everything about it apart from the ubiquitous functions - as a challenge centric to carbon-neutrality and device longevity, this is motivation enough to accept.

Now, TLS relies on invoking a channel to agree on the authenticity of the server which a client wishes to connect with. The client interrogates the server, to ascertain it is who it claims to be, before agreeing to send over any sensitive information such as username and password credentials. Subsequently the server’s identity is confirmed through the medium of an X.509 certificate, or more specifically, the Digital Signature Algorithm (DSA) on that certificate, and then the secure session begins. This is the way online transactions have taken place for decades.

Now, if the network entities do not intend to use web browsers, Certificate Authorities (CA), an OS, SPIFFS, or any other complications pertaining to their detriment, they do not need the certificate either. The reason for this is, without a browser, the certificate need not be kept, since it is for browser display behind the TLS handshake. Without the browser, devices can authenticate perfectly well using regular TLS functions and without all the unnecessary infrastructure. This deals with the centralised aspect that would otherwise stifle equality and distribution throughout the network.

To address energy consumption this is where risk assessment becomes so valuable. Since the TLS channel must no longer be invoked due to removing the centralised infrastructure, security functions can be employed with disregard to the rules of that channel. This allows freedom to employ CIA or privacy functions on a very flexible basis. It is important to embrace the standards and reasoning behind TLS, and so related works on risk-driven assessment for addressing the energy-security-privacy balance shall now be investigated.

3 RELATED WORK

Related work considers how a requirements analysis of the smart-farm domain would be fulfilled, threat modelling strategies, relevance of privacy under the GDPR, and security under the CIA.

3.1 Domain Requirements

From a business perspective, production input compared with output justifies domain viability. In the context of a smart-farm, this balance is delicately subject to the power vs processing tradeoff (Pavlović et al., 2021); a paradigm in which microcontrollers provide system monitoring and control ideally under the jurisdiction of green energy. If that green energy can supply all the energy required to fulfill the paradigm of carbon-neutral or negative crop production, the application is feasible, and if it can't, it isn't. Since security application threatens resources of constrained networks with such energy burdens, Soft Systems Modelling (SSM) (Sarrala, 2021), is a good place to scope basic requirements and eliminate frivolous procedures.

The 'seven stages' is an activity model based on domain context, achievable changes and actionable points, whereas 'two stream' model refers to logic-based analysis with cultural and political influence, similar to the 'four main activities' model. The CATWOE model identifies people, processes and environments which contribute to a situation, and is the most appropriate approach given that security should be tailored to an environment and its vulnerabilities. On this basis of environmental influences, other notable business models include SWOT and TOWS; Strengths Weaknesses Opportunities and Threats from internal and external direction.

3.2 Threat Modelling Strategies

There are several risk-driven methodologies for security at the systems design phase. STRIDE is the most mature (Mauri and Damiani, 2021), and provides threat-specific variants to arrange property violations identifiable in Data Flow Diagrams (DFDs) of the system design. An immediate problem with STRIDE in the smart-farm domain is the presumption of specific vulnerabilities according to the top ten, which in a heterogeneous environment, may not be applicable.

PASTA (Shin et al., 2021), describes a seven-stage model of activities to combine business objectives and technical requirements. It is attacker-centric, considers governance, architecture and operations, and may reflect the concerns of the domain better than STRIDE.

LINDDUN (Robles-Gonzalez et al., 2020), largely concerns privacy, with a systematic approach influenced by DFD-illustrated entities, data stores, processes and data flows. With reference to the aforementioned overlap between privacy and security, LINDDUN will be a valuable approach.

CVSS (Walkowski et al., 2021), provides a numerical score of a vulnerability based on its characteristics. The CVSS consists of three metrics groups (base, temporal and environmental), and is often used in conjunction with other threat-modelling methods. CVSS is not useful for the smart-farm security application design, but as part of a penetration test following implementation, towards remediation. It will not be used.

Persona non Grata (PnG) (Mantha et al., 2021), considers the motivations and skills of human attackers by characterising attackers as archetypes for system misuse towards specific goals. Use case scenarios are generally helpful in system design, and so this will be a beneficial security application of such scenarios.

TRIKE, was created as a security audit framework from a defensive perspective, understanding actors, assets, intended actions and rules based on a DFD, and then dividing each one into one of two categories - elevation of privilege or denial of service. Since this framework assumes that users are very active in the system, it is irrelevant - the aquaponics system aims for autonomy. The same disregard applies for the VAST threat model.

3.3 GDPR and the CIA

Privacy by Design (PbD) has been proposed as a Quantitative Threat Modeling Methodology (QTMM) (Luna et al., 2012), by risk assessment based on privacy-related attacks. The difference between privacy-oriented assessment is that other forms of threat modelling are largely agnostic of involved data subjects (Sion et al., 2019). The GDPR sets 99 articles and 173 recitals towards protecting the privacy of personal and sensitive data for subjects - and processes to attain that protection, with and without security functions.

Although the aquaponics system is largely devoid of personal and sensitive data belong to human subjects (we shall assume the fish don't mind), then the usefulness of the GDPR pertains to assessing a 'data-oriented' or 'content-based' attack risk, and ideally will result in substitution of security functionality with the less mathematical privacy functionality. Following the GDPR principle of 'necessary processing of data and nothing more than absolutely necessary', the challenge is to reduce protected content, and processing wherever possible.

GDPR privacy functions can be realised in several ways:

- Data masking (Larrucea et al., 2021): refers to disclosure of data with modified values. This could be character shuffling, tokenization, encryption, or character substitution - one of the oldest forms of cryptography. Substitution is where a value is exchanged for another value, to make identification or reverse engineering difficult. This character set will be the same throughout the policy, but may turn into a centralised process.
- Pseudonymisation (Kangwa et al., 2021): data de-identification that substitutes private identifiers with false identifiers or pseudonyms. "John Smith" could be changed to "Mark McConnell", but this would require some form of centralised database.
- Generalisation or minimisation (Goldsteen et al., 2021): the removal of data to prevent identification, such as a house number of an address, so accuracy of data remains whilst not being too specific.
- Perturbation (Broen et al., 2021): the use of round-numbering and adding random noise to modify a dataset. The set of values must be proportionate to noise to avoid disturbance, making this method quite complex.
- Synthetic data (Slokom and Larson, 2021): is algorithmically-generated information with no relation to the real case. These datasets represent patterns from real data by invoking standard deviation, linear regression, medians and other statistical methods. It is of no use to our comparatively simple model.
- Permutation (Barati et al., 2021): data swapping or shuffling rearranges dataset attributes so that they do not fit original information, often by switching columns in databases.

4 ASSESSMENT METHODS

This section considers the purpose of the security application with appropriate risk-based methodologies, how objective scoring can be calculated given the application parameters, and the influence of GDPR compliance.

4.1 Addressing Purpose

The purpose of the security application does not require typical considerations (Ankele et al., 2019), and the outcome will not be maximum security, granularity of roles or rules enforcement. IoT demands an unusual assessment method; to reduce security as much as safely possible for the longevity and autonomy of the domain.

Privacy Impact Assessment (PIA), and Privacy Threat Analysis (PTA), have been derived from security assessment models but do not consider privacy systematically, and so LINDDUN was invoked to satisfy this concern towards attributes Identification and Authentication (IA) (Robles-Gonzalez et al., 2020). Since IA are ubiquitous security procedures, the PIA and PTA approaches exemplify objective measurement, and how two entities can operate as a unit towards data reduction.

Although LINDDUN describes the use of DFDs to define a problem space, the smart-farm domain should be considered by content in preference to process flow. The reasons behind invoking metrics for content are the same as GDPR but with the focus on reduction of means rather than addition.

Data Protection Impact Assessments (DPIA), form metrics based on business impact which will then provide a risk score often requiring additional protection to mitigate that risk. Generally this is a positive approach - data can freely move around email, desktop folders, and portal hierarchies tiered by roles. The challenge with typical application is therefore to compose an exhaustive overview of human activity and put in place as many preventative measures as needed. In a smart-farm domain largely operating by autonomy of sensors and without an OS for human-oriented navigation, data cannot as easily be accessed, manipulated, lost or duplicated. As a result, it would be more fruitful to list the content by style of data dictionary and align privacy and security laws to each attribute, than it would to compile additional processes.

Data dictionaries are useful to exhaust all the possible system attributes proposed for security applications, without limiting the potential for data flow later on, this is because even as systems expand by network entities or sensors, ideally the security principles will not change. There is no need to define a number of maximum sensors as long as we know what type of topology is appropriate, and we do not limit attributes by a finite set. Security must be designed on the basis of infinite network expansion, and this way of designing is useful to ensure that scalability threats such as the data cumulation of blockchain will never occur.

4.2 Objective Scoring

The Open Web Application Security Project (OWASP), provides a risk rating template (OWASP, 2021), to objectively evaluate risk severity by multiplying likelihood of attacks from threat agents and vulnerabilities, by impact on technical systems and business continuity. Factor scores are ranked 1-9, based on a 'worst case scenario', and are fulfilled by the following criteria:

Estimating Likelihood: Threat Agent Factors.

- Skill level: How technically skilled the threat agents are.
- Motive: How motivated the group is likely to be, on desirability of reward.
- Opportunity: Resources and opportunities required.
- Size: The origin and position of threat agents against the system.

Estimating Likelihood: Vulnerability Factors.

- Ease of discovery: How easy it was to discover the vulnerability.
- Ease of exploit: How easy it is to perform the exploit.
- Awareness: The assumed knowledge of the exploit to the threat agents.
- Intrusion detection: How likely it is that the exploit will be detected.

Estimating Impact: Technical Factors.

- Loss of confidentiality: How much data could be disclosed and how sensitive it is.
- Loss of integrity: How much the data could become corrupted and how damaged it would be.
- Loss of availability: How much of the service could be lost, and the vitality of that service.
- Loss of accountability: The traceability of the threat to their origin.

Estimating Impact: Business Factors.

- Financial damage: Resulting financial loss from an exploit.
- Reputation damage: The extent to which reputation damage would cause the business.
- Non-compliance: How much exposure non-compliance would introduce.
- Privacy violation: The amount of resulting personal and sensitive information disclosure.

Resulting scores in low risk severity are not a concern, and medium scores are considered tolerable with monitoring, but high and critical results require immediate attention. To expand this template for IoT smart-farm application, additional factors and PbD solutions should be considered based on the results of the SSM testing; CATWOE, BATWOVE, SWOT and TOWS.

4.3 GDPR Compliance

The GDPR (Chikukwa, 2021), is a set of requirements composed of 173 recitals within 99 articles, to address seven principles. Big data systems have addressed GDPR requirements by mapping them to IT design for compliance of collection, storage, and analytics (Rhahla et al., 2021). The seven principles are outlined below with implications of smart-farm compliance.

Lawfulness, Fairness and Transparency. Whenever processing personal data, there must be a good reason for doing so, and the user must have given consent, be necessary to fulfil a legal obligation, for the protection of vital interests of a natural person or a public task for the interest of the public, and finally, with proof of legitimate interest. Reasons for processing must be transparent and reasonable.

Purpose Limitation. There must be a specific, explicit and legitimate reason for collecting this data. This reason must be justified in a sense that there is no alternative, and that data must be collected for this purpose.

Data Minimisation. Only collect the smallest amount of data possible for the purpose, additional data is not permitted. For example, email subscriptions should be limited to email addresses and not names, phone numbers or addresses.

Accuracy. Accuracy should be ensured on collection and storage, which is the responsibility of the collector.

Storage Limitation. The length of time data is stored must be justified. In the UK, the standard for data retention is seven years. Integrity and confidentiality Although integrity is always necessary and consistent by definition between privacy and security, confidentiality in a privacy sense does not necessarily mean encryption of data, but to preserve the privacy of an identifiable subject, whilst keeping it secure from internal and external threat.

Accountability. This is the proof that an organisation is complying with the GDPR and actually doing it - appropriate measures must be in place which supervising authorities can request at any time. This can be achieved by privacy by design and default.

5 METRICS AND RESULTS

Results of SSM are presented to outline the requirements of the application. From the CATWOVE and TOWS, additional attributes were deduced, and the data attributes listed by sensitivity. Security and privacy functionality was then ordered by energy consumption, and finally the full metrics framework demonstrated benefits of privacy and SSM design.

5.1 CATWOVE

CATWOE (Moumivand et al., 2021), and BATWOVE (Ireland et al., 2012), were combined to produce Customers, Actors, Transformation, World-view, Owner, Victims and Environment (CATWOVE), as a framework for defining business stakeholder perspectives. From the resulting concerns of this analysis, attributes can be considered for threat modelling and subsequently aligned with quantitative risk severity scores and solutions.
Customers: Consumers, green interests, investors, farmers, food retailers.
Actors: Researchers, engineers, plant and computer scientists, green energy interests.
Transformation: Security energy consumption to extend smart-farm battery life.
Worldview: Reducing security strength wherever safely possible, based on data content.
Owner: The model ignores TLS recommendations and instead uses ‘threshold’ security.
Victims: Each type of content must be considered, and the TLS channel will not be used.
Environment: All content must be considered, and TLS recommendations are ignored.

The security application intends to reduce energy waste towards environmentalism for the benefit of any system operating constrained devices. However, there is a threshold approach to the security levels of this solution, and that threshold is considerably lower than TLS recommendations stipulate. The metrics framework is designed to quantify the risks of such threshold-based practice to justify safety.

5.2 TOWS

Strengths, Weaknesses, Opportunities and Threats (SWOT), is a classic business strategy tool and TOWS is a variant of the same acronym to extend SWOT from an internal environment to an external one (Javaid, 2021). The aim of this model is to draw domain-specific attributes such as communication ranges and architectural details, which will impact the attributes of the metrics framework later on.
Internal Strengths: No browser, restricted internet access, hardware-accelerated crypto functions.
External Strengths: Range of communications protocols, one-to-one device relationships in BLE.
Internal Weaknesses: Battery life, overheads of cryptography, frequency of data transmission.
External Weaknesses: Lack of solar and wind in the Northern hemisphere and at night.
Internal Opportunities: Selective use of cryptography, data volume, privacy functions, protocols.
External Opportunities: Green energy - solar, wind, piezoelectric etc for charging device batteries.
Internal Threats: Sabotage by authorised access, downtime. External threats: Physical tampering, malicious attempts at network downtime by protocols.

SO: Strengths to Maximize Opportunities. Protocols belonging to IoT such as BLE, MQTT and LoRaWAN, are considerably lighter than the internet, and drain battery life at a much slower rate. As a result, a combination of reducing computation by selective cryptography, utilising green energy can sustain the network without mains electricity. In conjunction with the carbon dioxide absorbed by crop production, this can become a carbon-negative environment.

ST: Strengths to Minimise Threats. The use of short-range communications protocols reduces the threat landscape to that limited range. Use of the internet opens the network to threats of a global scale, whereas the use of BLE for example, restricts the threat to anyone outside of a 100m range at most.

WO: Minimise Weaknesses by Taking Advantage of Opportunities. With the exclusive one-to-one device relationships supported by BLE, the threat landscape becomes even smaller. Each device owns its own public-private key pair and long-term relationships promote the notion of infrequent authentication, as relationships are static until the system fails. As a result, the relationships stay the same, recognise each other, and use little power in re-authenticating.

WT: Minimize Weaknesses and Avoid Threats. This is about defensive strategies to minimise loss rather than promote success. The notion of ‘adequate’ or ‘threshold’ security is emphasised here to maintain a good balance between power and computation, where security is lowered as much as safely possible, whilst battery life benefits as much as possible by suffering from the least feasible computation.

5.3 Additional Attributes

IoT-specific attributes derived from the SSM study were: Time to exploit: The longevity of the data in the system. Physical range: The distance between two devices of each protocol. Key length by time: Key strength of relevant authentication keys.

5.4 Data Attributes

The ‘data dictionary’ was ordered by rank of sensitivity, where 1 was most sensitive, 10 the least:

1. Private keys: Private keys of the Public-Private key pair of device authentication should be a long-standing, impenetrable set for delivering data with guaranteed confidentiality.
2. AES keys: These keys are not as long-standing but should be capable of full confidentiality. HMAC-SHA keys: This key ensures integrity and authentication of each message.
3. Initialisation Vector (IV): This key is for hardening session ciphers.
4. Instruction code: This does not require confidentiality for the most part, but small amounts do.
5. Readings: These require integrity and authentication, but not confidentiality.
6. MAC address: This can be spoofed and should be connected with other security.
7. IP address: This can be spoofed and should be connected with other security.
8. Digital signatures: Proof of sender origin, and can be transferred without any CIA.
9. Public keys: Intended to be public identifiers of each device to decrypt transmissions.

5.5 Functionality Reduction

Functions of TLS can be applied individually and at the discretion of an application that does not require abiding by heavyweight regulations to function. This is because a TLS channel will not be employed, but the TLS library will. The challenge is first to determine the least energy-consumptive TLS application, and then if privacy functions can be used for substitution. Due to the vast number of TLS functions, applications, protocols and attack scenarios, it is most straightforward to rank functions by strength and energy consumption, where the strongest is highest consumption, but least risk severity. Using this method, a similar outcome to the OWASP DPIA was produced: Threat agent factors have been removed as it is good practice to assume agents are motivated and skilled.

5.5.1 Security Application

This set considers what kind of basic protection the data attributes require compared with what they are worth as a reward from an exploit. It is a good way of comparing typical TLS and internet applications against IoT by applying TLS-standards to deduce a result and comparing it against the IoT environment to determine which properties could be lightened or removed given the flexibility.

Table 1 describes protection requirements where the difference between privacy and confidentiality represent GDPR and encryption techniques respectively. From an IoT perspective, readings need not be fully protected, because it makes no difference if threat agents can see them or not. It does however matter that they are integrity-assured and with proof of origin. Digital signatures can be delivered in plaintext with zero protection, as they represent proof of origin. In these examples, the two most common data transmissions require no encryption - a huge energy saving to the IoT system that is not available when employing TLS.

Table 1: Protection: How much protection is required for this attribute.

Severity score	**Attribute**
1	Plaintext, no CIA or GDPR required.
2	Integrity and availability required.
3	Integrity, availability, and some privacy required.
4	Integrity, availability, and guaranteed confidentiality required.

Table 2 represents the value of the exploit reward and correlates with how much security it should have by that value. Low value material need not employ heavyweight protection, and ideally all data that can be transparent should be to reduce costs. This table differs from the first in a sense of motivation - whereas any network implementer will be inclined to protect data rigorously, if there is no motive to steal it, they should be discouraged from doing so.

Table 2: Value: Value of exploited data on success.

Severity score	**Attribute**
1	No value at all.
2	Value in conjunction with other data exploits.
3	Reputation and business disruption.
4	Full downtime, financial reward or kudos.

Table 3 is unique to IoT and was deduced from the SSM earlier. Since most of the communications will take place within a small and restricted environment, the threat landscape is not representative of global applications, and this further reduces the need for heavy security.

Table 3: Physical range: Accessibility of communications protocol.

Severity score	**Attribute**
1	1:1 relationships, up to 100m range.
2	1:1 relationships, up to 10km range.
3	M:N relationships, up to 1000 nodes.
4	M:N relationships, global range and unlimited nodes.

5.5.2 Energy Consumption

This set considers how long data attributes are in use compared against the security applied to them. It is useful to compare TLS default features in energy consumption against ideal IoT consumption. As a rule of thumb, the more ephemeral the data, the less heavy the protection need be - but TLS contradicts this by applying top security for all transactions.

Table 4 describes the frequency of reprocessing the security application of data.

Table 5 describes the symmetric, asymmetric, and Elliptic Curve Cryptography (ECC), security tiers. 80-bit security is very easy to break and is not included. 112-bit security has been deprecated by TLS, but could be used for short-lived attributes such as session key exchange prior to key derivation, or One-Time Pads (OTP).

Table 4: Frequency: Of security processing of data.

Severity score	**Attribute**
1	Weeks and months.
2	Days and weeks.
3	Hours and days.
4	Seconds and hours.

Table 5: Security strength: The level of security according to TLS.

Severity score	**Attribute**
1	None.
2	AES-112, RSA-2048, ECC-224, less strong and less costly.
3	AES-128, RSA-3072, ECC-256, quite strong and less costly.
4	AES-256, RSA-15,360, ECC-521, very strong but very costly.

Table 6 describes the sensitivity of data against business impact, and should be assessed on the assumption that data has been attained for malicious intent.

Table 6: Sensitivity: The intended exposure to the public.

Severity score	**Attribute**
1	Designed to be public.
2	Not designed to be public but of no consequence.
3	Designed to be private and could be critical with other data.
4	Designed to be secret and will be critical.

5.5.3 Privacy Substitution

This set is intended to provide potential differences in energy consumption by changing TLS settings. Low outcomes demonstrate little difference in energy use, but attention should be paid to medium, high and critical results, where a large amount of energy can be saved.

Table 7 represents energy savings based on reducing encryption key strengths. At the time of writing, an AES-112 key is safe for around 15 years, and AES-128 is quantum-proof for considerably longer (Barker and Roginsky, 2018). Coupled with the fact that a network operating on such small range communications such as BLE or LoRaWAN, smaller keys are safe for application. In an environment where security is so strong and energy so critical, all efforts towards reduction can safely be made.

Table 7: Encryption: The amount of reduction available by TLS standard.

Severity score	**Attribute**
1	No reduction in encryption strength.
2	Reduce to bit-strength AES-192, RSA-7680 or ECC-384.
3	Reduce to bit-strength AES-128, RSA-3072 or ECC-256.
4	Reduce to bit-strength AES-112, RSA-2048 or ECC-224.

Table 8 represents changes applied to data using privacy functions, ranked by lowest energy (1), to highest energy use (4).

Table 8: Privacy: Function substitution graded by strength.

Severity score	**Attribute**
1	Pseudonymisation or tokenisation.
2	Data masking, substitution, or permutation.
3	Plain text without privacy.
4	Data anonymised by removing attributes.

Table 9 represents changes in data frequency, where attributes such as device authentication can be extended given the permanent nature of device relationships. This is a different scoring system to protection of data, where rather than short-lived data being safer and cheaper, the reduction is placed on using it for longer periods of time in order to process it less.

Table 9: Computation time: Frequency of security or privacy to data.

Severity score	**Attribute**
1	Seconds and hours.
2	Hours and days.
3	Days and weeks.
4	Weeks and months.

5.6 Comparison Results

Generally, results of low are ideal, medium is acceptable, high requires attention and critical equates to running the system towards downtime given poor energy management.

Table 10 shows energy consumption using TLS:

Table 10: TLS results.

Data attribute	**Ref**	**Security application**	**Energy consumption**	**Privacy substitution**
Private keys	1	4	4	4
AES keys	2	4	4	4
HMAC-SHA keys	3	4	4	4
IV	4	3.33	3.50	3.50
Instruction code	5	3.33	3.50	3.50
Readings	6	3	3.25	3.25
MAC address	7	3	2.5	2.5
IP address	8	3	2.5	2.5
Digital signatures	9	4	3.25	3.25
Public keys	10	4	3.25	3.25
Average Score	-	3.567	3.375	3.375
Severity	-	Critical	High	High

5.6.1 Security Application

With TLS, security application is what it is, and there is no flexibility - the data is either sent using TLS and is under full protection, or no security is applied at all. As a result, every result is under full protection. The values vary, as some data ranges from public to secret, and the range of technologies can change such as MAC used for BLE. Range does not make a difference when using TLS, because the assumption is a global scale network of unlimited nodes. All scores return as high or critical.

5.6.2 Energy Consumption

Critical scoring has been saved by the nature of the data, and not the actions of TLS. Readings and MAC addresses owning public status lower the score of energy consumption, but it is still a high score, and that means a short lifespan for battery or green-powered networks.

5.6.3 Privacy Substitution

The keys could all be reduced to recomputing every few months to reflect the strength of AES. Reducing the cipher to 112-bit would be beneficial; despite the warnings from TLS that 112-bit is inadequate, it is still safe for several years, and the threat landscape is so small without a browser or global access, and atypical communications protocols such as BLE.

In summary, TLS is very high in security and energy consumption, but also very high in mitigation options given the ability to make them.

Table 11 shows energy consumption using a design framework to model an IoT security application:

Table 11: Privacy by design results.

Data attribute	**Ref**	**Security application**	**Energy consumption**	**Privacy substitution**
Private keys	1	3	2.75	2.75
AES keys	2	3	2.75	2.75
HMAC-SHA keys	3	2.33	2.25	2.25
IV	4	2.33	2.25	2.25
Instruction code	5	2	2	2
Readings	6	1.33	1.25	1.25
MAC address	7	1	1	1
IP address	8	1	1	1
Digital signatures	9	1	1	1
Public keys	10	1	1	1
Average Score	-	1.8	1.725	1.725
Severity	-	Low	Low	Low

5.6.4 Security Application

The beauty of energy consumption, security and privacy by design is that the consumption can be assessed before implementation. Here security application is the highest consumption of energy because the keys require full CIA, and without it there would be full downtime. However, the nature of other data has changed from full protection to optional privacy (instructional code), and no confidentiality or privacy (readings). The range of device communications has also changed from a presumption of global threat to a small network less than 100m between pairs of devices.

5.6.5 Energy Consumption

Severity scores requiring encryption are higher than others because encryption is computationally expensive. The score is still an overall low, and leaves little room for improvement.

5.6.6 Privacy Substitution

The keys could all be substituted with forms of pseudonymisation or tokenisation, and with an extended lifespan given the strength of AES. Reduction was not available for any attribute since the design indicated a value and sensitivity assessment pointing to the many attributes not requiring full protection. This consistency is a good indication that security and privacy by threat-base modelling offers benefits to the IoT environment.

In summary, the scores were low throughout, and there was little energy reduction ability given the benefits offered by the framework earlier on, whereas TLS returned critical and high, the IoT network returned low.

6 CONCLUSION AND FURTHER WORK

The metrics framework has provided an effective way to identify areas of security available for energy reduction using privacy functions. The following contributions were made:

6.1 Content-driven

Vulnerabilities were based on assessing the content of data attributes and transmissions. Selective security was applied to the majority of data, saving large amounts of energy where TLS would automatically apply high-level security. A lot of data was discovered to be of little value in plaintext, and so making it transparent but retaining integrity assurance and message authentication meant there is less attack motivation.

6.2 Energy Reduction

Whereas risk-driven methodologies aim towards defense, this framework offers the novelty of security reduction towards energy conservation. The biggest threat to an IoT environment is the constraint of power and processing which reduce device lifespan. As long as security is 'adequate' then extending the longevity of the network is a priority. The metrics framework demonstrated how energy reduction can be identified by using risk modelling so that non-threatening security is possible.

6.3 Worldview

The use of SSM allowed a wider analysis not typically employed in security assessments. This was useful to establish the benefits and concerns of an IoT domain away from protection needs, which in turn was very useful for providing adequate protection without superfluous energy use. Since the communications protocols were so short in range compared to internet use, and many of the transmissions were deemed low value, the majority of computing could be alleviated.

6.4 Substitution

The use of GDPR allowed a list of available functions not typically employed in security design. Since most of the data could be generalised or pseudonymised in whole or part, selectivity of standard TLS functions was applied. Substituting security functions for privacy functions reduced the energy consumption of the domain enormously, and the metrics framework demonstrated the areas in which substitution could make high and critical differences to energy use.

6.5 Additional Attributes

Modelling using a worldview enabled domain-specific data attributes. Attributes not normally considered in risk assessment for security applications made a big difference in energy use. By considering the domain from a business perspective, attributes such as range and the value of content during data transmissions gave insight into how security functions could either be reduced to privacy or removed completely. Although TLS does not provide the flexibility to do this, the functions belonging to TLS can be applied to retain strength and ubiquity of security whilst avoiding the high costs.

The framework was designed to have agnostic application - future work could include health care, blockchain and industry towards sustainability in reduced energy use.

REFERENCES

Aminzade, M. (2018). Confidentiality, integrity and availability – finding a balanced IT framework. *Netw. secur.*, 2018(5):9–11.

Ankele, R., Marksteiner, S., Nahrgang, K., and Vallant, H. (2019). Requirements and recommendations for IoT/IIoT models to automate security assurance through threat modelling, security analysis and penetration testing. In *Proceedings of the 14th International Conference on Availability, Reliability and Security*, number Article 102 in ARES '19, pages 1–8, New York, NY, USA. Association for Computing Machinery.

Barati, M., Aujla, G. S., Llanos, J. T., Duodu, K. A., Rana, O. F., Carr, M., and Rajan, R. (2021). Privacy-Aware cloud auditing for GDPR compliance verification in online healthcare. *IEEE Trans. Ind. Inf.*, pages 1–1.

Barker, E. and Roginsky, A. (2018). Transitioning the use of cryptographic algorithms and key lengths. Technical report, National Institute of Standards and Technology.

Broen, K., Trangucci, R., and Zelner, J. (2021). Measuring the impact of spatial perturbations on the relationship between data privacy and validity of descriptive statistics. *Int. J. Health Geogr.*, 20(1):3.

Chikukwa, G. (2021). *A Consent Framework for the Internet of Things in the GDPR Era*. PhD thesis, Dakota State University.

El Alloussi, H., Fetjah, L., and Chaichaa, A. (2014). Securing the payment card data on cloud environment: Issues & perspectives. *International Journal of Computer Science and Network Security (IJCSNS)*, 14(11):14.

ESP32 (2021). https://www.esp32.net. Accessed: 2021.

Espressif (2021a). https://www.espressif.com. Accessed: 2021.

Espressif (2021b). https://www.espressif.com. Accessed: 2021.

Goldsteen, A., Ezov, G., Shmelkin, R., Moffie, M., and Farkash, A. (2021). Data minimization for GDPR compliance in machine learning models.

Ireland, V., Cavallo, A., Omarova, A., Ooi-Sanches, Y., and Rapaport, B. (2012). Approaches in addressing system of systems. In *2012 7th International Conference on System of Systems Engineering (SoSE)*, pages 380–385.

Javaid, N. (2021). Integration of context awareness in internet of agricultural things. *ICT Express*.

Kangwa, M., Lubobya, C. S., and Phiri, J. (2021). Prevention of personally identifiable information leakage in e-commerce via offline data minimisation and pseudonymisation. *Int. J. Innov. Sci. Res. Technol*, 6(1):209–212.

Larrucea, X., Moffie, M., Mor, D., and Others (2021). Enhancing GDPR compliance through data sensitivity and data hiding tools. *JUCS-Journal of Universal Computer Science*.

Luna, J., Suri, N., and Krontiris, I. (2012). Privacy-by-design based on quantitative threat modeling. In *2012*

7th International Conference on Risks and Security of Internet and Systems (CRiSIS), pages 1–8.

Mantha, B., García de Soto, B., and Karri, R. (2021). Cyber security threat modeling in the AEC industry: An example for the commissioning of the built environment. *Sustainable Cities and Society*, 66:102682.

Mauri, L. and Damiani, E. (2021). Stride-ai: An approach to identifying vulnerabilities of machine learning assets. In *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*, pages 147–154.

Microcontrollerslab (2019). Esp32 web server in arduino ide: Control leds.

Moumivand, A., Azar, A., and Toloie Eshlaghy, A. (2021). Combined soft system methodology and agent-based simulation for multi-methodological modelling. *Syst. Res. Behav. Sci.*, (sres.2802).

OWASP (2021). https://www.owasp.org. Accessed: 2021.

Pavlović, N., Šarac, M., Adamović, S., Saračević, M., Ahmad, K., Maček, N., and Sharma, D. K. (2021). An approach to adding simple interface as security gateway architecture for IoT device. *Multimed. Tools Appl.*

Razikin, K. and Widodo, A. (2021). General cybersecurity maturity assessment model: Best practice to achieve payment card Industry-Data security standard (PCI-DSS) compliance. *CommIT (Communication and Information Technology) Journal*, 15(2):91–104.

Rhahla, M., Allegue, S., and Abdellatif, T. (2021). Guidelines for GDPR compliance in big data systems. *Journal of Information Security and Applications*, 61:102896.

Robles-Gonzalez, A., Parra-Arnau, J., and Forne, J. (2020). A linddun-based framework for privacy threat analysis on identification and authentication processes. *Comput. Secur.*, 94:101755.

Sarrala, T. (2021). *Uncovering privacy threats with Soft Systems Methodology: Development of a privacy threat modelling method for today's needs*. PhD thesis.

Shin, Y.-J., Cho, E., and Bae, D.-H. (2021). PASTA: An efficient proactive adaptation approach based on statistical model checking for Self-Adaptive systems. *Fundamental Approaches to Software Engineering*, 12649:292.

Sion, L., Van Landuyt, D., Wuyts, K., and Joosen, W. (2019). Privacy risk assessment for data Subject-Aware threat modeling. In *2019 IEEE Security and Privacy Workshops (SPW)*, pages 64–71.

Slokom, M. and Larson, M. (2021). Doing data right: How lessons learned working with conventional data should inform the future of synthetic data for recommender systems.

Varanda, A., Santos, L., Costa, R. L. d. C., Oliveira, A., and Rabadão, C. (2021). Log pseudonymization: Privacy maintenance in practice. *Journal of Information Security and Applications*, 63:103021.

Walkowski, M., Oko, J., and Sujecki, S. (2021). Vulnerability management models using a common vulnerability scoring system. *NATO Adv. Sci. Inst. Ser. E Appl. Sci.*, 11(18):8735.