1001 – SCIENTIFIC AND TECHNOLOGICAL RESEARCH PROJECT SUPPORT PROGRAM

PROJECT APPLICATION FORM

Project Title: *Multi-Purpose Encryption Optimization in a Federated Learning Framework for Critical Infrastructures*
Project Leader: Assoc. Prof. Dr. Oğuzhan Ceylan
Institution/Organization Where the Project Will Be Conducted: Kadir Has University

SUMMARY

Project Summary

Critical infrastructure (CI) such as **energy, transportation, communications, and healthcare services** are vital for both **national security** and **social welfare**. **Industrial Control Systems (ICS)** and **Internet of Things (IoT)** devices used in these infrastructures typically operate with constraints such as **low processing power**, **limited energy capacity**, and **restricted network bandwidth**. This situation renders **traditional centralized artificial intelligence solutions** unfeasible. In this context, **Federated Learning (FL)** offers a suitable learning paradigm for critical infrastructure by providing advantages such as **data privacy, low latency,** and **local model training**.

However, during the implementation of **Secure Aggregation** protocols used in FL, issues such as **high energy consumption**, **increased latency**, and **computational load** arise, particularly due to the limited hardware capabilities of edge devices. In this context, the project will model these three fundamental objectives separately to create a multi-objective optimization problem, aiming to develop **Pareto optimal solutions** among these functions.

Pareto optimality defines situations where a solution cannot be improved without worsening the others when improving any objective function. Within the scope of this project, solutions will be generated on **the Pareto front** that provides the best balance between **energy efficiency**, **communication delay**, and **privacy level**; thus, **dynamic encryption protocols** suitable for every infrastructure condition will be proposed.

To this end, **meta-heuristic algorithms** such as **the Whale Optimization Algorithm (WOA)** and **Genetic Algorithm (GA)** will be used to solve **the multi-objective encryption optimization problem**. The results obtained will be tested in both **virtual test environments** (e.g., ICSSIM) and **realistic data sets** (e.g., Edge-IIoT) and subjected to performance analysis.

As a project output, **energy-efficient**, **high-privacy**, and **low-latency encryption solutions** will be developed for edge devices operating in FL processes; this will enable **end-to-end secure artificial intelligence applications in critical infrastructures**. The project will contribute to **national cybersecurity strategies** and offer **a pioneering approach to FL security solutions for critical infrastructures**.

Researchers: The project will be led by Associate Professor Oğuzhan Ceylan and supervised by Professor Hasan Dağ, with a team of seven members consisting of two Ph.D. students, one master's student, and two undergraduate students, comprising a total of seven members (two academics and five scholarship researchers). Undergraduate and graduate researchers will be selected from individuals who are active students in the Management Information Systems, Electrical and Electronics Engineering, or Computer Engineering departments. Work is planned to be carried out over 36 months through 4 work packages during the project period.

Keywords: Federated Learning (FL), Secure Aggregation , Industrial Control Systems (ICS), Masking, Homomorphic Encryption (HE), MPC, Multi-Objective Optimization, Whale Optimization Algorithm, Genetic Algorithm, Cybersecurity, and Energy Efficiency

Title: Multi-objective Encryption Optimization within a Federated Learning Framework for Critical Infrastructures

Summary

Critical infrastructures—including **energy systems**, **transportation**, **communication**, and healthcare—are essential to both **national security** and the **continuity of societal functions**. The **Industrial Control Systems (ICS)** and **Internet of Things (IoT)** devices used in these systems are often **resource-constrained**, with **limited processing capabilities**, **energy**, and **bandwidth**. This makes **centralized AI models** inefficient or infeasible. **Federated Learning (FL)** offers a viable alternative by supporting **on-device training** and ensuring **data privacy** with **reduced communication latency.**

However, in FL, the use of **Secure Aggregation protocols** introduces significant overhead, especially when implemented on **low-power edge devices**. These protocols, relying on **masking**, **homomorphic encryption (HE)**, or **multi-party computation (MPC)**, can cause **high latency**, **excessive energy usage**, and **computational load**. Addressing this challenge requires a **multi-objective optimization** approach that balances these conflicting requirements.

This project aims to solve the multi-objective optimization problem based on three criteria to identify **Pareto optimal** solutions—those where **none of the key objectives (energy, latency, privacy) can be improved without compromising another**. The **Pareto front** will be used to provide **a range of optimized trade-off configurations**, tailored to different infrastructure types and device capabilities.

To this end, **metaheuristic algorithms** such as the **Whale Optimization Algorithm (WOA)** and the **Genetic Algorithm (GA)** will be employed to solve the **multi-objective secure aggregation optimization problem**. These methods will be tested and validated on both **virtual testbeds** (e.g., ICSSIM) and **realistic datasets** (e.g., Edge-IIoT), ensuring robust and comparative evaluation.

Consequently, the project will develop **energy-aware**, **privacy-respecting**, and **low-latency secure aggregation protocols**, enabling the safe integration of **AI-based cyber threat detection systems** into **real-world critical infrastructure environments**. The outcomes will significantly contribute to **national cybersecurity frameworks** and pioneer a new direction in **secure federated learning architectures**.

The project will be carried out by a team of seven members in total—two academics and five scholarship-supported researchers—under the leadership of Assoc. Prof. Dr. Oğuzhan Ceylan (Principal Investigator) and with the guidance of Prof. Dr. Hasan Dağ. The team includes two doctoral students, one master's student, and two undergraduate students actively enrolled in Management Information Systems, Electrical-Electronics Engineering, or Computer Engineering programs. The research will be organized into four work packages and is planned to span 36 months.

Keywords: Federated Learning (FL), Secure Aggregation, Industrial Control Systems (ICS), Masking, Homomorphic Encryption (HE), MPC, Multi-Objective Optimization, Whale Optimization Algorithm, Genetic Algorithm, Cybersecurity, Energy Efficiency

1. UNIQUE VALUE

1.1. Importance of the Topic and Original Value of the Project:

Critical Infrastructure Systems (CIS) encompass information and industrial control systems that could lead to loss of life, economic damage, and national security vulnerabilities in the event of a cybersecurity breach [1]. These systems are designated as "critical" because they provide vital functions or connect critical components [2]. **The federated learning (FL) approach** plays an important role in reducing the latency, bandwidth, and privacy risks created by large data transfers to central servers**,** thanks to the ability to process data at the source [1,2].

Supervisory Control and Data Acquisition (**SCADA) systems**, in particular, stand out as central control and monitoring systems for industrial facilities and critical infrastructure. These systems oversee the operations of vital sectors such as production processes, energy distribution, water and waste management, and transportation. In recent years, the proliferation **of IoT devices** in the first (bottom) layer of the three-layer structure of SCADA systems, as shown in Figure 1, has increased the number of sensors, devices, and control units integrated into SCADA systems. While this integration offers advantages such as **smart monitoring and real-time data collection**, it also introduces new vulnerabilities to cyberattacks.

In this context, the IoT structure in SCADA systems is of critical importance for both **data collection** and **remote monitoring/control**. However, the limited processing power, memory, and energy capacities of IoT devices bring with them the risks and performance bottlenecks of centralized data transmission. This situation makes the FL approach even more attractive, while enabling the reduction of the cyber attack surface through on-site data processing.

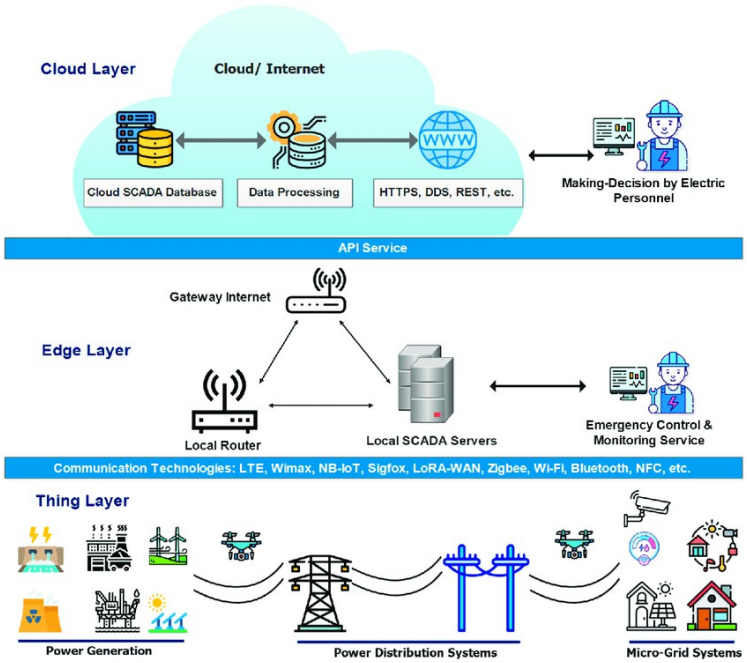

Figure 1 - *An example of a three-tier SCADA architecture [12]*

In a previously conducted and completed TÜBİTAK 1001 project (code: 120E487), studies were carried out on **FL and text analysis (BERT-based)**. In this project, attack/anomaly detection was performed in a distributed environment, particularly based on system logs; however, the detailed optimization of secure aggregation protocols (e.g., cryptographic parameter sizes, latency-energy balance) and the joint evaluation of different protocol types (masking + MPC + HE) remained limited. **This new work** aims to build on our previous experiences by integrating cryptographic methods with multi-objective optimization to create a secure aggregation protocol with higher attack resilience, lower latency, and greater efficiency (lower energy consumption).

IT-OT Convergence refers to the integration of information technology (IT) with operational technology (OT). While IT is data-centric, OT manages physical processes; this integration creates more efficient but also more vulnerable systems [3,4]. With Industry 4.0 and IoT, systems have become increasingly distributed and interconnected, expanding the cyber attack surface and making it more difficult to protect critical infrastructure [5]. **Particularly in CASs,** the need to implement different security protocols according to the Purdue Layered Security Architecture (see Figure 2)

highlights the importance of a holistic approach [6,7].

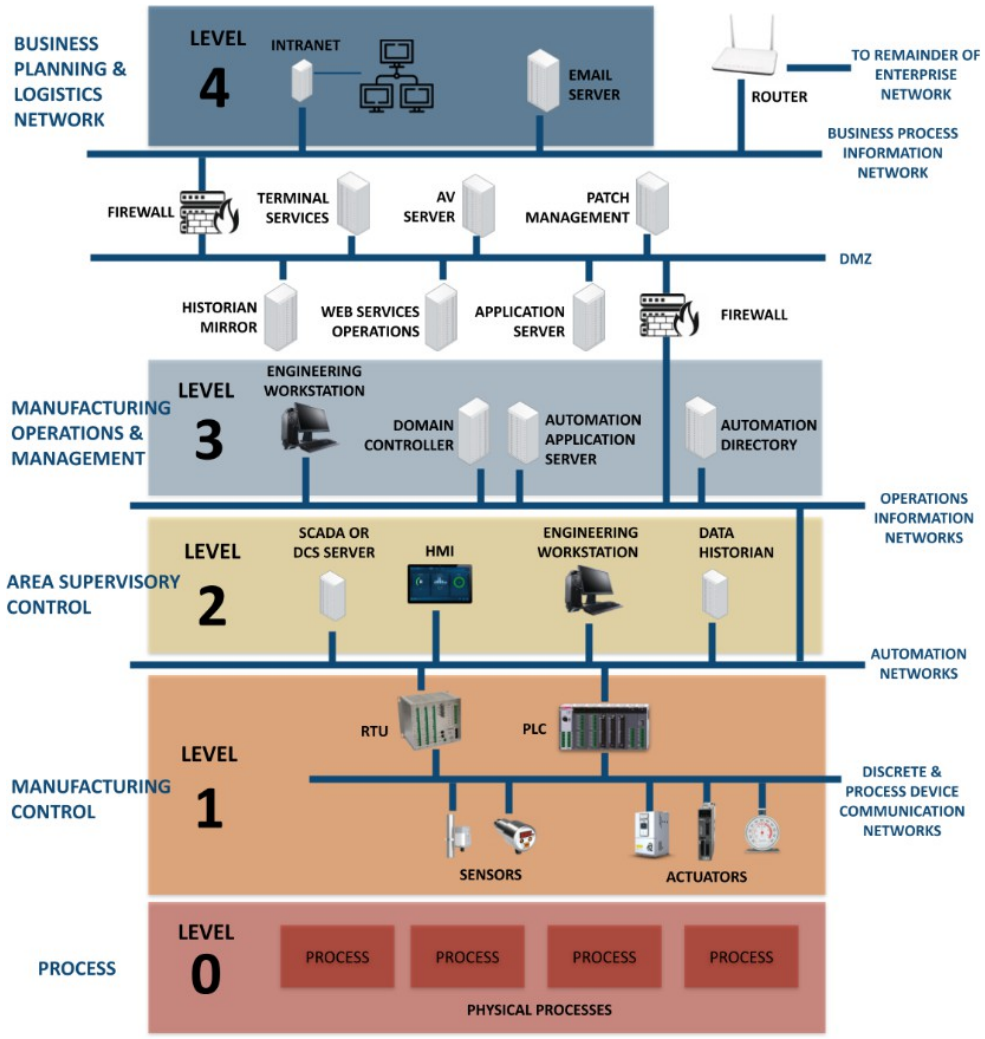

Figure 2 - *Purdue Model [13]*

In this context, the Stuxnet attack (2010, nuclear power plant in Iran, Siemens control systems) symbolizes the beginning of the era of cyber warfare [8]. Examples such as the Siberian pipeline (1982), Chevron refinery (1992), Worcester airport (2007), Maroochy water-waste system (2000), and CSX train line (2003) clearly demonstrate the consequences that vulnerabilities in KASs can have [9–11]. These incidents highlight **the inadequacy of traditional security measures** and the need for comprehensive defense systems that can process IT and OT logs, supported by natural language processing and federated machine learning.

Industrial Control Systems (ICS) and IoT devices have become integral parts of these infrastructures. However, the data generated by these devices is highly sensitive and directly impacts operational processes. **The FL model** enables data to be trained locally without being extracted from the source, based on the principle that participating devices only transmit model updates (parameters or derivatives) without sharing their data [14,15]. It is imperative that the confidentiality of these updates is preserved; otherwise, malicious attackers could infer the original data from the parameter values [15,21,29]. This is where "Secure Aggregation" protocols come into play. Although various approaches exist in the literature, such as masking [14,15], homomorphic encryption (HE) [27,28], and multi-party computation (MPC) [18,29], achieving the desired performance in practice is challenging due to the high computational requirements of KASs and the limited hardware of IoT devices.

The advantages and disadvantages of these approaches can be summarized as follows:

1. **Masking-Based Methods:**
   Participants add random noise (or a mask) to their updates and transmit them to the server; the collective mask is then removed after a sufficient

number of participants [14,15]. The main advantage of this method is its low computational and communication costs, but there is a risk of mask estimation if the updates of one or a few participants are exposed [26].

2. **Homomorphic Encryption (HE)-Based Methods:**
   HE enables addition (and sometimes multiplication) operations on encrypted data without revealing the data [27,28,31]. FHE (fully homomorphic encryption), while offering the widest functionality, is limited by high computational costs and bandwidth consumption. Partial HE applications still require careful optimization in resource-constrained environments such as IoT [27].

3. **Multi-Party Computation (MPC):**
   Data is stored by dividing it among multiple parties (secret sharing) or as various encrypted shares. The server or intermediary parties perform the aggregation process without revealing the original data [18,29]. However, issues such as connection interruptions and dropout situations on edge devices complicate the implementation of MPC in IoT networks [15].

Each approach offers advantages and disadvantages in different dimensions such as **privacy**, **computation time**, **energy consumption**, and **communication cost** [14,15,27]. This situation indicates that masking is preferred for low latency, while full HE or MPC may be preferred for high privacy. However, the literature has not yet reached a clear solution regarding which protocol should be applied with which parameters in multi-purpose scenarios [22,29].

Optimization approaches are used to find the best solutions in the complex parameter space of FL protocols [16,17]. For example, variables such as masking level, key sizes in HE, number of parties in MPC, and configuration of backup mechanisms can be investigated using meta-heuristic algorithms (Genetic Algorithm (GA), Whale Optimization Algorithm (WOA)) [20,21]. However, most studies focus on a single cryptographic method or optimize only a single performance metric [15,27]. **The gap in the literature** lies in the limited number of approaches that systematically evaluate masking, HE, and MPC approaches within the same framework using multi-objective cost functions (energy, latency, privacy).

This gap can be summarized as follows:

- **Lack of Protocol Diversity and Optimization:** Limited studies on optimizing different protocols with dynamic weights [15,18,27].
- **Realistic Data and Test Platforms:** Pilot tests conducted with limited parameters in environments such as Edge-IIoT [10] and ICSSIM [24,25] [14,15].
- **Active Attack and Dropout Scenarios:** Insufficient research on participant devices disconnecting from the network or active attack scenarios [26,29].
- **Multi-objective Optimization:** The lack of function models that jointly evaluate energy, privacy, and latency [15,27].

This project aims to combine **masking, homomorphic encryption, and MPC protocols** under the umbrella of **meta-heuristic multi-objective optimization** (Whale Optimization Algorithm, Genetic Algorithm) [16,17,20]. In light of Deb's pioneering work on multi-objective evolutionary optimization ([32]), the goal is to determine the Pareto-optimal balance of conflicting objectives such as energy, latency, and privacy. Thus, original contributions will be made both **theoretically (**a new optimization model and protocol design) and **practically (**experimental validation using the Edge-IIoT dataset [23], the ICSSIM virtual test bed [24,25], and a simulated physical test bed to be created).

For critical infrastructures (energy grids, water distribution, transportation, etc.), it is necessary to consider not only the "processing data locally" advantage provided by FL, but also the **security** and **speed** dimensions of the system [24,25]. In a cyberattack scenario, the attacker's ability to predict the system's status or vulnerabilities through model updates poses a significant risk [15,21]. On the other hand, the requirement for real-time management of command and sensor data in environments such as SCADA/PLC (Programmable Logic Controller) renders protocols with high latency impractical [27,31]. Therefore, **establishing an appropriate balance between privacy and performance** stands out in the literature as a difficult problem to solve [28,30].

Additionally, the risk that traditional encryption methods could be broken with the development of quantum computers poses an additional threat **in terms of post-quantum attack scenarios** [22,29]. In this context, the fact that HE and MPC-based methods require larger key sizes and stronger encryption levels to increase security can exacerbate latency and energy consumption issues.

In terms of applicability, experiments conducted on the Edge-IIoT dataset (IoT/IIoT attack variety), ICSSIM, and simulated physical test environments to be created **will strengthen the** study's **connection to the real world**; for example, developing FL-based cyber attack detection or prediction models in infrastructures such as electricity

grid, water distribution, and transportation, will form the backbone of future smart city and Industry 4.0 applications [15,27].

Finally, the project aims to achieve **a 15% energy savings** and a **30% delay reduction**, thereby extending the lifespan of edge devices and maintaining the system's real-time control capabilities [15,31]. Thus, our project aims to make significant contributions to critical infrastructure security and federated learning both nationally and internationally.

This study **integrates** masking, homomorphic encryption, and MPC protocols with meta-heuristic multi-objective optimization models **to establish a secure, energy-efficient, and low-latency federated learning framework**. By providing a systematic and dynamic answer to the question, "How should each protocol be configured in which environment?", we aim to fill a significant gap in the literature.

1.2. Research Question and/or Hypothesis:

Research Question:
"In federated learning scenarios requiring high security and performance in critical infrastructures (EKS/IoT), how can secure aggregation protocols be optimized to ensure both privacy and energy/computational efficiency?"

Hypothesis:

"By improving secure aggregation protocols based on homomorphic encryption/masking using multi-objective optimization methods such as the Whale Optimization Algorithm and Genetic Algorithms, at least 15% energy savings and 30% latency reduction can be achieved; data integrity and privacy risks in critical infrastructures are minimized."

The research question above consists of two main components in line with the operational needs **of critical infrastructures**: **(i) privacy**, **(ii) performance**. The need to minimize cyberattack and privacy risks associated with data transmission in EKS/IoT systems necessitates an in-depth examination of current technological approaches **in secure collection protocols (**masking, HE, MPC) [15,18]. However, the implementation of different protocols often brings critical factors such as **computational power, energy consumption, and communication latency**. Therefore, in our project, we ask the question, "How can we optimize these protocols?" and seek answers.

The target of **at least 15% energy savings and 30% latency reduction**, defined in the **hypothesis** section, represents a significant improvement compared to typical masking/HE/MPC-based FL experiments reported in the literature [15,27,28]. It is estimated that the mechanisms to be developed to reduce the computational load of homomorphic encryption and decrease the number of packets, especially in mobile or embedded IoT devices, will be more efficient than current attempts. Furthermore, meta-heuristic methods such as WOA and GA to be used in the project have been reported to provide improvements of up to 10–20% in the literature [16,17,20]. Based on this, it is assumed that a higher rate of improvement may be possible for critical infrastructures.

Furthermore, the project will not only claim a theoretical increase in security but will also measure its effectiveness **in practice** with a cyberattack detection scenario on the Edge-IIoT dataset [23] and industrial control scenarios on ICSSIM [24,25] and a physical simulation test bed. This application dimension ensures the testability of the predictions in the hypothesis.

1.3. Objectives and Goals:

Objective:
To develop new protocols and optimization models that will make federated learning real-time and energy-efficient while protecting it from privacy violations in EKS and IoT devices used in critical infrastructures.

Objectives:

1. To analyze existing secure aggregation (masking, homomorphic encryption, MPC) methods in the EKS/IoT context.
2. Using the work of Assoc. Prof. Dr. Oğuzhan Ceylan, who previously solved multi-objective optimization problems in electrical power systems, to define a multi-objective function and integrate algorithms (WOA, GA).
3. Edge-IIoT data set, ICSSIM virtual test bed, and the developed physical simulated test bed are secure.

FL prototype.

4. To produce at least 2 SCI articles, 2 conference papers, and sectoral cooperation reports by the end of the project.

These objectives and goals make the project's outputs **measurable** and concrete. For example:

1. **Requirement analysis**: By examining both academic and industrial reports on existing protocols, it will be determined which parameters (masking parameters, homomorphic encryption key size, number of MPC parties, number of communication rounds, number of edge devices, data size) are critical in EKS/IoT environments [14,15,22]. This analysis will reveal, based on quantitative data, which protocol has advantages or disadvantages under which conditions.

2. **Multi-objective optimization problem**:

Delay time: $T(x) = T_{comm} + T_{(comp)}$ *(minimize)* Energy consumption: $E(x) = \sum(P(i_j) \times t(i_j) + E_{(comm)}$ *(minimize)* Privacy level: $P(x) = \log(2)_{(M)} + \alpha \cdot k + \beta \cdot t$ *(maximize)*	(1)

These three objectives will be optimized within a **Pareto-priority** framework using **the Whale Optimization Algorithm (WOA)** and/or **Genetic Algorithm (GA)** [16,17,20,32]. Instead of a single weighted function, **different** trade-offs will be discovered using the dominance principle.

3. **Prototype validation**: In the project, the prototype, which will be implemented using Python/C++-based modules, will be applied both to cyber attack detection on **the Edge-IIoT dataset** ([23,27]) and to SCADA-like industrial process controls virtually on **ICSSIM** ([24,25]) and physically on the physical simulated test bed to be created. Success metrics will include "total latency," "energy consumption" (e.g., measured on Raspberry Pi-like devices), and "privacy resilience against possible attack scenarios."

4. **Publication and collaboration**: The academic contribution of the project will be realized through at least 2 SCI articles and 2 conference papers. Additionally, workshops or demo sessions with the industry will open doors to **real-world** applications; the developed prototype can be tested in the test environments of different institutions (e.g., electricity distribution companies).

The above **clear, measurable**, **and achievable** objectives are planned to be realized within the project's 36-month work schedule. Furthermore, the project's objectives are aligned with **the secure digital transformation of critical infrastructures** and **national cybersecurity** strategies [15,27].

2. METHOD

2.1. Research Design and Phases

The project has a research design consisting of **four main phases**. These phases comprise both theoretical and applied methods in line with the objectives and goals **stated in the introduction**. Throughout the project, **the dependent variables** are:

1. **Privacy Level** (e.g., the attacker's ability to leak data from model updates or a mathematical privacy metric),
2. **Latency** (communication and processing delay),
3. **Energy Consumption** (total power consumption for end devices)

have been determined. In contrast, the independent variables, i.e., the variables we control or modify in the project, are:

- Cryptographic parameters of the protocols,
  - Masking parameters
  - Homomorphic encryption key size
  - Number of MPC parties
- Number of communication rounds,
- Number of end devices,

- Data size,

have been determined [15,27,28].

Figure 3 below summarizes the general flow of the method.

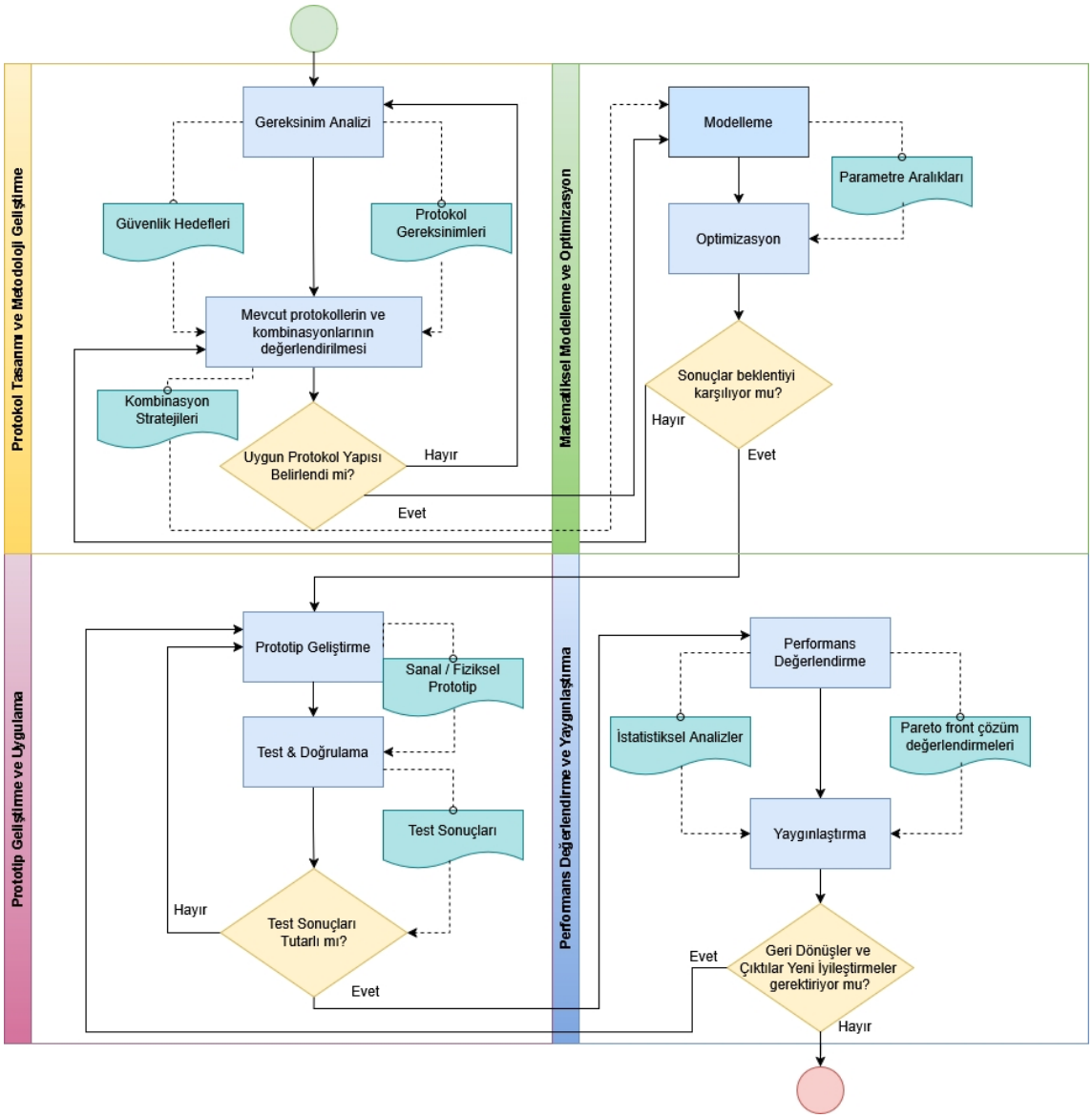

Figure 3 *- Method Flow Chart*

In this diagram, the output(s) of each phase provide input for the next phase. For example, the parameter ranges and security scenarios determined as a result of **Protocol** Design will guide the methods and values to be used in the **Mathematical Modeling** phase.

2.2. Protocol Design and Methodology Development

In this phase, the current state of secure aggregation protocols underlying the project is examined. Specifically:

- **Masking**: Most existing studies propose hiding data with random noise or masking keys. However, there are certain risks against passive attacks, and in the case of active attacks, there is complexity in multi-mask sharing strategies [14,15,28].
- **Homomorphic Encryption (HE)**: While fully homomorphic encryption works well in terms of privacy, it has high computational costs . computational cost. It is difficult to overcome the bandwidth

bandwidth and processing time constraints [27,28,31].

- **MPC (Multi-Party Computation)**: It can be implemented in synchronous or partially asynchronous models. However, the need for multiple parties to be constantly online and for reliable channels poses challenges in IoT networks [18,29].

Within the scope of **requirements analysis**, in EKS/IoT environments:

- **Real-time capability**: Low latency thresholds for decision-making mechanisms in SCADA systems and IoT devices,
- **Fault tolerance**: Some end devices dropping out of the network or communication errors occurring.
- **Low energy consumption**: Scenarios where thousands of IoT nodes operate on batteries at a large scale.
- **Data privacy and security:** The need for encrypted data processing in critical infrastructures and minimizing security risks,
- **Compatibility with resource constraints:** Developing optimized solutions for edge devices with limited memory, processing power, and bandwidth.

,

especially during the integration of data security mechanisms such as homomorphic encryption (HE) and secure multi-party computation (MPC).

The Edge-IIoT dataset [23] and **ICSSIM** [24,25] platforms can be given as examples representing the basic requirements here. The Edge-IIoT dataset consists of IoT and IIoT (Industrial IoT) attack types and normal traffic samples, making it applicable to the protocol's **cyber attack detection** scenario. ICSSIM, on the other hand, enables protocol testing in EKS environments by simulating SCADA and industrial control components in a microservices-based manner.

The critical parameters to be determined at this stage (masking parameters, homomorphic encryption key size, number of MPC parties, number of communication rounds, number of end devices, data size) and **performance metrics** (privacy level, latency, energy) will be used in subsequent stages [15,22].

During the **protocol design** phase, it will be defined how masking, HE, and/or MPC approaches **will be applied together** or **in a dual combination**
For example:

1. **Masking + Homomorphic Encryption (HE)**: Gradients are first concealed with a light-level mask, then partial homomorphic encryption is applied [27,28]. This aims for lower computational cost by using partial HE (e.g., BFV, BGV) instead of full HE.

2. **MPC + Masking**: Model parameters are shared among different parties (secret sharing), while additional noise is added against passive attack risks [18,29].

3. **Homomorphic Encryption Only or Masking Only**: If full cryptographic processing power is insufficient on some edge devices, only masking can be used. This situation can be designed as a "heterogeneous protocol" scenario.

Protocol Steps and Dropout Management

The protocol must define how it will operate against passive and active attacks, while also taking into account that edge devices (e.g., IoT sensors) **may not be online during every computation cycle (round)**. In scenarios such as **Federated Learning (FL)** or **secure multi-party computation (MPC)**, edge devices may be unable to join the network or share data in certain rounds for various reasons. This situation makes **dropout resilience mechanisms** critical for model updates and security.

For example, **dropout-resilient approaches** ([26]) **include strategies that preserve data integrity and privacy, such as sharing masking keys.** If an endpoint goes offline during a specific round, the protocol must clearly define how it will manage this loss and ensure security. Such mechanisms ensure system continuity and prevent security vulnerabilities by preventing computation from being disrupted due to missing data.

Table 1 below contains the items of an example protocol flow:

Table 1 *- Protocol Flow Table*

Step	Description	Reference
1. Key/Noise	End devices generate random noise values for masking or keys for HE.	[2], [14],

Generates	pairs.	[15]
2. Starting the Learning Round	Devices begin preparing update parameters after local training (e.g., mini-batch gradient descent).	[1], [6], [13]
3. Encryption / Masking	Updates are encrypted or masked according to the selected protocol. (Partially homomorphic / fully HE)	[14], [18]
4. Aggregation	The server or MPC participants aggregate the received encrypted/masked parameters. Partial decryption or mask removal is performed in a manner that does not compromise privacy.	[5], [16]
5. Model Update	The server/hubs that redistribute the aggregated result announce the new global model. Devices that dropped out can continue from where they left off in the next round.	[2], [13]
6. Check and Repeat	The round is completed, and the protocol parameters (e.g., mask refresh frequency, homomorphic encryption key validity) are updated and continued for the desired number of rounds.	[1], [2], [14]

Security Analysis: A passive attacker can listen to the data sent by the endpoint device or server but does not manipulate it. An active attacker, on the other hand, poses additional threats such as manipulation or insertion of fake data. Taking these differences into account in protocol design, **public key exchange, random mask reset** intervals, and **redundant management of MPC parties** will be planned [15,29].

2.3. Mathematical Modeling and Optimization

In this project, the cost items and performance metrics defined for **the Secure Aggregation protocol** will be addressed using **a multi-objective optimization** approach. Although some studies in the literature [16,17,20] use **a single weighted objective function**, a **Pareto-priority multi-objective optimization** approach based on **Deb (2001)** principles [32] has been adopted. This will enable the simultaneous evaluation of **three different metrics—energy consumption (E), delay time (T),** and **privacy (P)**—and the discovery of **different** trade-offs.

2.3.1. Cost and Performance Functions

The delay (T), **energy (E)**, and **privacy (P)** metrics are defined below, along with an explanation of their importance in the project. Each metric is treated as **a separate objective function**.

(a) Latency (T)

Total delay can be expressed as the combination of **communication** ($_{Tcomm}$) and **processing** ($_{Tcomp}$) delays [15,27]:

$$T = {}_{\text{Tcomm}} + {}_{\text{Tcomp}} \tag{2}$$

- $_{Tcomm}$: The time it takes to send/receive data from end devices to the server (or MPC sides).
- $_{Tcomp}$: The processing time incurred during encryption, masking, and aggregation operations.

Objective: Minimize the delay T value, as **low latency** is critical in real-time decision-making environments such as EKS/IoT.

is of critical importance.

(b) Energy Consumption (E)

Energy consumption in IoT devices can be modeled based on processor (CPU/GPU) usage and wireless communication costs [14,19]:

$$E = \sum_{i=1}^{N} (P_i \times t_i) + E_{comm} \quad (3)$$

- $_{Pi}$: Average power consumption of the relevant device (Watt).
- $_{ti}$: Time spent on the relevant operation (seconds).
- $_{Ecomm}$: Energy expended for data transmission/reception.

Objective: To minimize the Energy E value, as **energy savings** are critical, especially in large-scale IoT scenarios where **thousands of end devices** operate on batteries.

(c) Privacy Metric (P)

The privacy metric can be defined in terms of the protocol's susceptibility to passive/active attacks or **the difficulty** of an attacker recovering data [18,22,29]. The literature contains various formulas that quantify this metric. For example [27,30]:

$$P = f(M, k, t) = \log_2 (M) + \alpha \cdot k + \beta \cdot t \quad (4)$$

- M: Size of the encryption/masking space (number of possible key or mask combinations).
- k: Cryptographic parameter (e.g., key size, number of parties).
- t: Time-related security variable (e.g., key renewal period).
- α, β: Contribution coefficients of each parameter to confidentiality.
- $_{log2}$(M): Reflects the size of the key space in bits; the number of combinations the attacker must solve.

Objective: To maximize the privacy P value.

2.3.1.1. Pareto-Based Multi-Objective Optimization

In the project, **T** and **E** are treated as **three targets** that must be **minimized**, while **P** must be **maximized**. According to **Deb**'s **(2001)** approach [32], instead of using **a single composite function** (e.g., F(x)=αE+βT−γP), **Pareto optimization** is applied. This ensures that:

- D**ominance Principle**: A solution **x** can be **Pareto optimal** as long as **it is not completely dominated** by another solution **y**.
- **Pareto Front**: The set of **non-dominated** solutions found by the system. Here, **different equilibrium points** (e.g., low E–low T, medium P, etc.) can be found simultaneously.
- **Selection Based on Scenario**: Critical infrastructure managers can choose from solutions on the Pareto front based on **current needs** (battery status, attack risk, etc.).

This approach better adapts to **the dynamic** and **multidimensional nature** of environments with **limited resources** (IoT) and **high security requirements** (EKS).
nature of environments with limited resources (IoT) and high security requirements (EKS).

When the three separate objectives addressed in the project—for example, latency (T), energy consumption (E), and privacy (P)—need to be optimized in different directions (minimizing the value in two metrics, maximizing the value in one metric, etc.), **Pareto optimization** emerges as the most appropriate approach. The fundamental idea behind this method is to find the most "balanced" solutions (**Pareto optimal solutions)** by considering conflicting objectives **simultaneously**, rather than defining a single aggregate cost function [32].

1. **Dominance Principle**
   Any solution x is considered non-dominated if no other solution y is *worse (or equal and worse in one objective)* than x in both delay and energy. Therefore, x is dominant over any solution that "satisfies" the system's different needs as a whole to a lesser extent, and is considered *Pareto optimal* if it does not encounter a solution that is the opposite.

2. **Pareto Front**
   The set of non-dominated solutions obtained as a result of the domination analysis, where no solution can completely outperform (or "outperform") another, is called the *Pareto front*. Each solution on the Pareto front represents **a** different trade-off:

   - **If you want to reduce energy consumption further,** the delay or privacy cost increases.
   - **If you want to reduce latency a little more**, you may have to accept an increase in energy consumption.

3. **Selection Based on Scenario**
   Critical infrastructure managers (or decision-makers) can select any solution on **the Pareto front** based on the current
   They can choose **according to their needs**. For example:

   - **During a period of high attack risk**, they can shift toward a solution that maximizes privacy (in which case latency and energy consumption may increase).
   - In a scenario involving **IoT devices with limited battery life**, they may prefer the Pareto point that minimizes energy consumption.

This flexibility plays a critical role in the project because the Pareto-based approach aligns with **the dynamic and multidimensional nature of environments with limited resources (IoT) and high security requirements (EKS)**.

Pareto Front Example: Delay (T) - Energy (E) Relationship

Figure 4 below presents a Pareto solution example for a **two-objective** optimization problem involving delay and energy. The blue dots represent potential solution points obtained randomly or experimentally. The vertical axis shows the delay time (ms, seconds), while the horizontal axis shows the energy consumption (Joule, Watt-hours).

- **Blue Points (All Solutions):** Solutions or scenarios obtained with different parameters.

- Red **Points (Pareto Optimal Solutions):** *Each of* these points is not dominated *by* any other solution because it is not worse than any other solution in terms of both energy and delay.

- **Red Line (Pareto Front):** The curve formed by connecting these red points together shows the "best equilibrium points" collectively. The Pareto front can be considered in 3D (or even higher dimensions) rather than 2D when multiple objectives such as T, E, and P are addressed together in a project; here, T and E are shown in two dimensions for illustrative purposes.

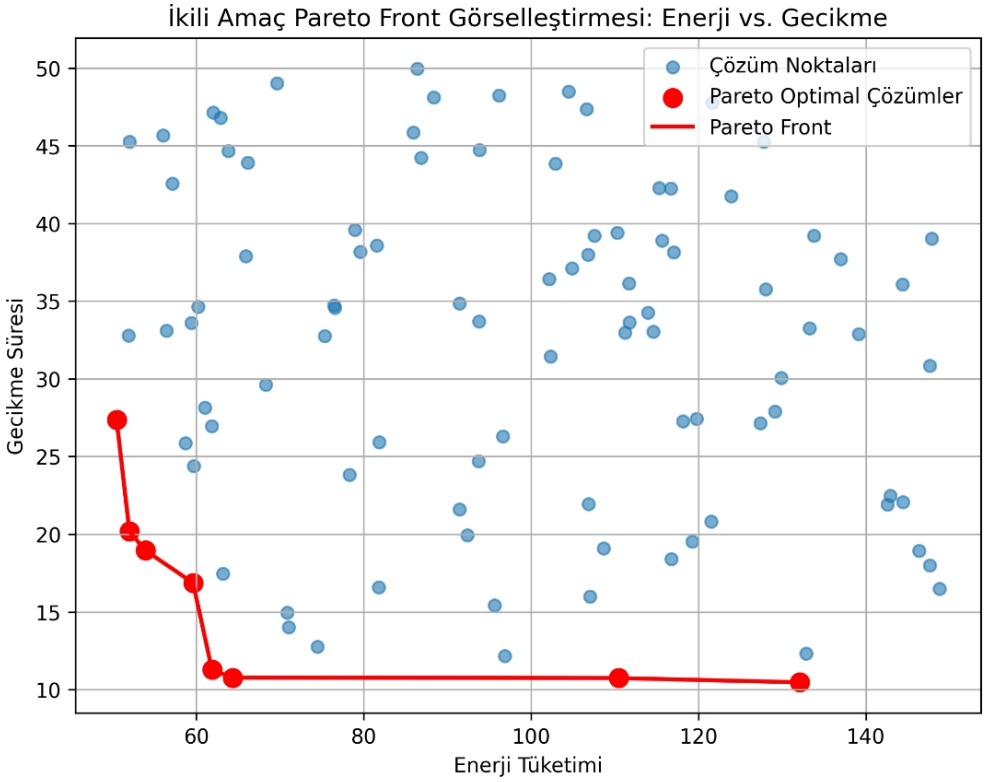

Figure 4 - *Visualization of the Binary Objective Pareto Front: Energy vs. Delay*

The following information can be derived from this graph:

- The Pareto points in the lower left corner can achieve both low energy and low latency objectives **quite well** ; however, **other metrics** (e.g., privacy, attack resistance) may be neglected at this point.
- As one moves along the Pareto front, improving one metric may **inevitably** result in a loss in another metric (e.g., higher energy, longer delay).
- When working with real data (e.g., the Edge-IIoT dataset in our project, test bed environments), decision-makers can select the ideal point based on system requirements using this Pareto front obtained from optimization algorithms (WOA, GA) can select the ideal point based on **system requirements**.

Thus, Pareto-based multi-objective optimization contributes significantly to the **real-time**, **resource-efficient**, and **secure** operation of critical infrastructure and IoT devices by balancing conflicting objectives such as **energy, latency, and privacy** within the scope of the project.

2.3.2. Optimization Approaches

Meta-heuristic methods such as **the Whale Optimization Algorithm (WOA)** and **Genetic Algorithm (GA)** [16,17,20,21] will be used within the scope of the project to solve the multi-objective optimization problem. In the literature, both algorithms have successful applications on **single-objective** or **multi-objective** functions. In this study, **Pareto-based** versions will be preferred, and **the Pareto front** will be obtained using concepts such as **domination, non-dominated sorting, and crowding distance** [32].

2.3.2.1. Whale Optimization Algorithm (WOA)

The Whale Optimization Algorithm is a method that mimics **the "bubble-net" hunting behavior of humpback whales** [3], [4], [7]. The basic steps are as follows:

1. **Encircling Prey**:
   Whales approach the solutions considered as "prey" by drawing circles around them.
2. **Bubble-Net Strategy**:
   The position is updated using a combination of **circular** and **radial** components with a logarithmic spiral movement.
3. **Exploration**:

When a good solution has not yet been found or when escaping a local minimum is desired, movement to different positions is provided.

Pareto-Based WOA:

- Instead of searching for a single "best" solution, solutions that **are not dominant** within the population are stored as a pioneer (elite) set.
- In each iteration, whales update their positions based on different solutions on this **Pareto front**, **discovering solutions that perform well in terms of different**
  objectives (T, E, P).

In the context of the project, a **whale** (i.e., candidate solution) may include the following parameters:

$x=[\ \text{MaskingLevel}\ ,\ \text{HEKeySize}\ ,\ \text{MPCParties}\ ,\ \text{CommunicationRoundCount}\ ,\ldots]$	(5)

In each iteration, thanks to the **Pareto-optimal** approach, energy and latency are reduced while privacy is also maximized as much as possible [32].

2.3.2.2. Genetic Algorithm (GA)

Genetic Algorithm (GA) is a meta-heuristic method that models the **genetic evolution** process **in nature** (selection, crossover, mutation) [16,21].

- **Selection**: Solutions with high fitness values are transferred to the next generation **with a higher probability**
  to the next generation.
- **Crossover**: New "individuals" (solutions) are created by combining the parameters of the parent solutions.
- **Mutation**: A random or controlled change (perturbation) is added to some of the new solutions
  is added to some new solutions; **diversity** is maintained in the population.

Pareto-based GA (e.g., **NSGA-II**, MOEA [32]):

- Instead of a single function, "non-dominated" solutions are selected using the **domination** principle (non-dominated sorting).
- The **crowding distance** metric is used to preserve the diversity of solutions on the Pareto front.
- In each generation, **non-dominated** individuals are preserved with **elite protection**, thus capturing different equilibrium points for ***minT***, ***minE***, ***maxP*** at the same time.
  different **equilibrium points** are captured simultaneously for minT, minE, and maxP.

2.3.2.3. WOA and GA in the Project Scope

- **Parameter Vector**: Masking parameters, Homomorphic encryption key size, MPC party count, Communication round count, Edge device count, Data size
- **Population / Whales**: "Whales" x_i in WOA, "Individuals" x_i in GA.
- **Iteration / Generation**: At each stage, the values of the solutions (T, E, P) are calculated and the **Pareto front is updated** according to the **dominance** analysis
  the Pareto front is updated [32].

If the solution space is large (e.g., hundreds of variables such as **bit size, number of rounds, masking level**), WOA or GA will be run in parallel on **HPC servers**. In terms of speed-quality balance, both **methods** will be tested and a **comparative** analysis will be performed. If necessary, **ensemble** approaches ([14,17]) can also be used (e.g., using WOA output as the initial population for GA or vice versa), thereby increasing the chance of **overcoming local minima**.

Dynamic Scenarios and Constraints:

- A minimum threshold can be set for the privacy metric **P** in different scenarios (e.g., **high attack risk**).
- **Under low battery** conditions, solutions can be eliminated when the energy **E** upper limit is exceeded.
- Such **dynamic constraints** will determine which solutions from the Pareto front will be selected in practice.

Consequently, **WOA** and **GA** will be used with a **multi-objective Pareto optimization** approach in the project context to find the **most suitable** parameter sets for **federated learning protocols**. **Appropriate** solutions can be selected on **the Pareto front** in cases such as different network conditions, attack levels, or latency requirements, thus creating an **intelligent and dynamic** parameter selection mechanism.

2.4. Prototype Development and Implementation

In this phase, secure aggregation protocol modules will be developed in Python/C++ and comprehensively evaluated using both virtual and physical test environments. The application environment will be divided into three main components:

1. **Prototype Architecture**

   - **End Devices (Client):**
     - Arduino-Based Boards: Used for low-power sensor and data collection operations.
     - Raspberry Pi: Will be used for applications requiring higher processing power; functionality will be enhanced to simulate real-world scenarios for IoT devices.
     - These devices will be configurable in a controllable manner to measure the effect of different cryptographic parameters (masking parameters, homomorphic encryption key size, MPC party count, Communication round count, Endpoint device count, Data size) will be configurable in a controlled manner.
   - **Server (Aggregator/Coordinator):**
     - The HPC server or cloud-based platform will provide integration and coordination of protocol modules; central management of the secure aggregation process and Pareto optimization processes will be performed here.
   - **Network Layer:**
     - Data exchange between devices will be enabled using MQTT, TCP/IP, or similar communication protocols.
     - To enable real-time monitoring of network performance, parameters such as communication round count and data size will also be monitored.

2. **Test Environments**

   - **Virtual Test Environment – ICSSIM:**
     - ICSSIM, which models SCADA/PLC systems based on microservices, will provide a simulation environment that includes industrial sensor data, virtual PLCs, and system parameters.
     - This environment will enable the protocol to be tested with near real-time data and will allow for the measurement of metrics such as communication, latency, energy consumption, and protocol stability.
   - **Physical Simulation Test Bed:**
     - Arduino and Raspberry Pi-Based Test Bed:
       - Arduino and Raspberry Pi boards will be integrated to simulate the functions of IoT devices on real hardware.
       - This test bed will be set up in a laboratory environment and will simulate scenarios such as data exchange, protocol implementation, and dropout (temporary device disconnections) management.
     - This structure will enable the comparison of results in the virtual environment with the performance of physical devices and adapt the model to real-world conditions.

3. **Metric Measurement and Evaluation**

   The prototype's performance will be evaluated based on the following metrics:

   - **Latency (T):**
     - Communication delay (T_comm) and processing delay (T_comp) occurring during encrypted or masked data transfer will be measured.
     - The average cycle time obtained both in the virtual environment (ICSSIM) and on the physical test bed will be evaluated.
   - **Energy Consumption (E):**
     - The energy consumed during the operation of protocol modules on edge devices such as Arduino and Raspberry Pi will be calculated using actual measurement devices or software estimation libraries.

- Energy consumption during communication and processing phases will be analyzed separately.

- **Privacy (P):**
  - The protection levels of the protocol in passive and active attack scenarios (e.g., data eavesdropping or message manipulation) will be evaluated.
  - Integration of cryptographic methods such as (masking, homomorphic encryption, MPC) will be examined in terms of data integrity and privacy levels based on defined criteria.

- **Protocol Stability and Dropout Management:**
  - The protocol's flexibility in the event of certain endpoints going offline during specific rounds will be thoroughly tested, along with the operation of dropout-resistant mechanisms.
  - Both in simulation and physical test environments, the system's continuity and security in the system during dropout situations.

This structure will enable comprehensive testing of the developed secure aggregation protocol under both virtual and real hardware conditions, allowing for necessary improvements to be made to the system based on the measured metrics.

2.5. Performance Evaluation and Deployment

The collected **cost** and **performance** data will be compared with methods in the literature that use **only masking** or **only HE** [14,15,27]. **Pareto-based** results aim to confirm the predicted efficiency gains (in terms of energy and latency) **of 15% and above**.

- **Statistical Analysis**:
  - The statistical significance **of different protocol configurations** will be examined using one-way ANOVA or multiple comparison tests (Tukey, Bonferroni).
  - Using time series analysis, the change in delay T and energy consumption E in each round will be examined [16,21].

- **Dissemination**:
  - Project outputs will be disseminated to **the scientific community** and **industrial stakeholders** [15,27].
  - Thanks to the **Pareto front** structure, **different equilibrium points** obtained **in different scenarios** can also reveal different perspectives within the scope of the results they will produce. At this point, this information will also be shared with relevant stakeholders using appropriate channels.

Addendum: In performance evaluation, attack detection metrics such as **ROC (Receiver Operating Characteristic), Precision-Recall,** and **F1-score** will be used, not just average values. This will enable more comprehensive results regarding false positive and false negative rates.

2.6. Preliminary Work and Infrastructure

The project team leader, Assoc. Prof. Dr. Oğuzhan Ceylan's experience in the field of **multi-objective optimization** (e.g., [16,17,20]) already constitutes the preliminary work for this project. Previously, **WOA** and **GA** methods have been reported to yield efficient results in **distributed energy systems** and **IoT resource allocation**. In addition, the project team has **basic-level testing** experience using Edge-IIoT data sets and ICSSIM. Therefore, once the requirements analysis is complete at the start of the project, it will be possible to quickly move on to the protocol design and pilot testing phases.

In summary, the methodology section comprehensively explains **the logic of the research design**, **dependent and independent variables**, **mathematical modeling**, **optimization techniques**, **prototype development** steps, and **performance evaluation** criteria. This approach is structured to enable the achievement of **the project's aims and objectives**, as each stage is fed by inputs from the previous one, and a **comparative** evaluation **against the literature** is presented in the final stage.

3. PROJECT MANAGEMENT

3.1. Management Structure: Work-Time Schedule and Work Packages

Project Team: Prof. Dr. Hasan Dağ (Advisor), Assoc. Prof. Dr. Oğuzhan Ceylan (Project Coordinator) – (IP 1,2,4), Doctoral Fellow 1 (D1) – (IP 2,3,4), PhD Fellow 2 (D2) – (WP 1,2), Master's Fellow (YL1) – (WP 3,4), Undergraduate Fellow 1 (L1) – (WP 1), Undergraduate Fellow 2 (L2) – (WP 3)

3.1.1. Work-Time Schedule

WORK-TIME SCHEDULE (*)

IP No	Work Package Name	Importance to Project Success (%) (**)	Who Will Perform It (***)	Months																																				
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1	Protocol Design and Methodology Development	20	Y, D2, L1	X	X	X	X	X	X	X	X	X	X	X	X																								
2	Mathematical Modeling and Optimization	35	Y, D1, D2						X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X												
3	Prototype Development and Implementation (Edge-IIoT, ICSSIM)	25	D1, YL1, L2																	X	X	X	X	X	X	X	X	X	X	X	X								
4	Performance Evaluation and Dissemination	20	Y, D1, YL1																					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

3.1.2. Work Packages

<table>
<tr><th colspan="2">WORK PACKAGE TABLE (*)</th></tr>
<tr><td>WP No: 1</td><td>IP Name: Protocol Design and Methodology Development</td></tr>
<tr><td colspan="2">IP Objective:

• Design the step-by-step flow of secure aggregation protocols based on masking, HE, and MPC.
• Define critical details such as the number of communication rounds, dropout management, and key exchange.
• Conduct resilience analysis of the protocol against passive/active attack scenarios.</td></tr>
<tr><td>Tasks/Responsibilities to be Performed within the Scope of the IP:

• Schematic diagrams of cryptographic approaches (masking+HE, MPC+masking, and other combinations).
• Determining the communication phases of each protocol (from end device to server, from server to end device).
• Designing a dropout resilient (end-device connection interruption) mechanism.
• Security analysis: Masking decryption in passive attacks, fake parameter attempts in active attacks, etc.</td><td>Person(s) Who Will Implement the IP and Their Contributions to the IP (**)

• Project Leader (Y): Outlining the general methodology and protocol steps, providing guidance for security analysis.
• PhD Fellow 2 (D2): Drafting cryptographic methods (e.g., MPC, HE); technical requirements analysis for SCADA/PLC compatibility.
• Undergraduate Fellow 1 (L1): Conducting literature reviews, preparing small-scale test scenarios (dropout, passive attack), drawing flowcharts.</td></tr>
<tr><td colspan="2">Success Criteria: Qualitative and/or quantitative criteria are specified in a measurable and trackable manner to determine when the relevant work package is considered successful.

• Detailed flowcharts and methodology documentation for at least two different secure protocols will be completed.
• At least one protection strategy will be proposed and reported for each type of passive and active attack (e.g., masking solution, fake parameter attempt).
• In small-scale concept tests, the protocol will be shown to perform its secure collection function with at least a 90% success rate (e.g., data integrity preservation).</td></tr>
<tr><td colspan="2">Interim Deliverables:

• Protocol Design Document (flow diagrams, UML-like diagrams).
• Passive/active attack analysis report (resilience strategies against different types of attacks).
• Concept Test Results: For example, a brief report showing whether the protocol works with a small data set.</td></tr>
<tr><td colspan="2">Risk Management:</td></tr>
<tr><td>Risk Definition</td><td>Measures to be Taken (Plan B)</td></tr>
<tr><td>• The protocol becoming too complex and difficult to implement.</td><td>• Modular design, ability to disable optional components (e.g., using partial HE instead of full HE).</td></tr>
<tr><td>• Inadequate modeling of active attack scenarios or the emergence of new/different threats during the project process.</td><td>• Review all known threat models in the literature. Define additional scenarios for new/different threats.</td></tr>
</table>

WORK PACKAGE TABLE (*)

WP No: 2	**WP Name:** Mathematical Modeling and Optimization

WP Objective:

- Express the parameters of each protocol, such as processing time, bandwidth, energy cost, and privacy level, using mathematical formulas.
- Define a **multi-objective optimization** model based on the (T, E, P) metrics of each protocol and perform a **Pareto front** search in the parameter space using meta-heuristic techniques such as WOA and GA based on **the dominance principle**.
- **Statistically analyze** the optimization results based on outputs such as **energy savings**, **reduction in communication delay**, and **improvement in privacy level**.

Tasks/Responsibilities to be Performed within the Scope of the IP:	**Person(s) Who Will Implement the IP and Their Contributions to the IP (**)**
• Creating functions for the delay (T), energy (E), and privacy (P) metrics. • Defining the **Pareto** approach for **jointly** (multi-objective) optimizing the T, E, P metrics; determining scenario-dependent **constraints** or **priorities** (e.g., minimum privacy, maximum delay). • Adapting or developing Whale Optimization Algorithm (WOA) and Genetic Algorithm (GA) codes. • Conducting initial optimization tests on HPC (High Performance Computing) servers.	• **Project Leader (Y):** Selection and general framework of the Pareto-based multi-objective optimization approach; verification of outputs. • **PhD Fellow 1 (D1):** Adapting Whale Optimization Algorithm (WOA) or Genetic Algorithm (GA) codes, formulating energy-privacy delay functions. • **PhD Fellow 2 (D2):** Model studies on HPC (high-performance computing), parameter fine-tuning, and statistical analysis (e.g., ANOVA, p-values).

Success Criteria: *Qualitative and/or quantitative criteria for the relevant work package to be considered successful are specified in a measurable and traceable manner.*

- In the multi-objective optimization process to be performed using at least two different optimization algorithms (**e.g., WOA and GA**), **the target criteria** — energy consumption, communication delay, and privacy level — will **each reach at least 80% of their optimal value**.
- As a result of optimization, **energy consumption will be reduced by at least 15%** compared to the reference system. The delay time will be reduced by at least 30% compared to the reference system.
- Model accuracy will be maintained above 90%, and the statistical significance of the results obtained will be demonstrated at a level of $p < 0.05$.

Intermediate Outputs:

- Mathematical Model Document (E, T, P formulas)
- Optimization Algorithms Source Code (Python/C++)
- Performance Report: Best/average success values in test scenarios, parameter combinations.

Risk Management:

Risk Definition	**Measures to be Taken (Plan B)**
• Insufficient computing resources or delays on the HPC server.	• Refer to the cloud infrastructure of the university/partner institutions or additional HPC resources.

• Insufficient Pareto front generation (obtaining very few non-dominated solutions).	• Parameter range , Increasing diversity with additional mutation/exploration steps.
• Optimization algorithms getting stuck in local minima or results not converging for a very long time.	• Try different meta-heuristic methods (PSO, Tabu Search), fine-tune parameters.

WORK PACKAGE TABLE (*)	
WP No: 3	**WP Name:** Prototype Development and Implementation (Edge-IIoT, ICSSIM)

WP Objective:

- Convert the designed secure aggregation protocol(s) into a prototype running on the Edge-IIoT dataset and ICSSIM test bed in a Python/C++ environment.
- Measure metrics such as latency, bandwidth, and energy consumption in realistic scenarios.

Tasks/Responsibilities to be Performed Under the Scope of the IP:	**Person(s) Who Will Implement the IP and Their Contributions to the IP (**)**
• Software development and integration of protocol modules (masking, HE, MPC). • Cyber attack detection scenario in the Edge-IIoT dataset: Participation of edge devices in FL training, implementation of secure collection protocols. • Industrial control scenario on the ICSSIM test bed: Processing SCADA/PLC data with federated learning. • Creation of a physical simulation test bed and application of the method in a physical environment. • Performance measurements: Recording values such as transferred data size, processing time, energy consumption (e.g., on Raspberry Pi).	• **PhD Fellow 1 (D1):** Preparing the backbone of the prototype software, connecting optimization modules in Edge-IIoT data set integration. • **Master's Student (MS1):** Code development (Python/C++), setting up an industrial control scenario in the ICSSIM environment, energy and latency measurements. • **Undergraduate Fellow 2 (L2):** Test automation and debugging of established software; data collection and reporting (e.g., performance graphs).

Success Criteria: *Qualitative and/or quantitative criteria that define when the relevant work package is considered successful are specified in a measurable and trackable manner.*

- The prototype will be deployed to operate seamlessly (with 90% or higher success) in Edge-IIoT and ICSSIM environments.
- It is targeted to reduce average latency by 30% and energy consumption by 15% in at least two scenarios (e.g., energy and attack detection).
- The model will be ensured to complete the FL update (including dropout management) flawlessly between 5 and 10 iterations without the accuracy dropping below 90%.

Intermediate Outputs:

- Working Prototype Software (Python/C++ repository)

- Test Result Reports (Edge-IIoT and ICSSIM): Performance evaluation with tables and graphs.
- Video or Demo Recording (optional): A short prototype demonstration for presentation.

Risk Management:

Risk Identification	**Measures to be Taken (Plan B)**
• Error messages or software incompatibilities in the Edge-IIoT data set and ICSSIM test bed that differ from expectations.	• Version compatibility checks, additional documentation, or transition to similar data sets/test environments (e.g., NSL-KDD, CICIDS in cyberattack data sets).
• High memory or CPU consumption in the software.	• Code and algorithm optimization, lighter cryptographic libraries, HPC support.

WORK PACKAGE TABLE (*)

WP No: 4	**IP Name:** Performance Evaluation and Dissemination

IP Objective:

- Demonstrate efficiency gains by comparing the results with similar protocols in the literature.
- Prepare scientific articles and papers, develop sectoral collaborations, evaluate potential patent applications.

Tasks/Responsibilities to be Performed within the Scope of the IP:	**Person(s) Who Will Implement the IP and Their Contributions to the IP (**)**
• Analysis of project results: Measurements obtained, optimization successes, protocol efficiency. • Comparison with protocols in the literature: Typical reference values for masking, HE, MPC (e.g., Bonawitz protocols [2]). • Scientific publications: At least 2 SCI articles, international conference papers. • Sectoral collaboration meetings, demo presentations, initiation of potential patent/utility model processes.	• **Project Leader (Y):** Comparison with reference methods in the literature, presentation of project results to academic/industrial stakeholders, coordination of articles and papers. • **Doctoral Fellow 1 (D1):** Documenting notable improvements in the analysis of the obtained results (e.g., energy, delay, privacy), preparing SCI article drafts. • **Master's Fellow (MF1):** Compiling graphs and tables for publications, technical preparation for prototype demonstrations, supporting sectoral collaboration presentations.

Success Criteria: *Qualitative and/or quantitative criteria for the relevant work package to be considered successful are specified in a measurable and trackable manner.*

- The developed protocol will provide at least a 15–30% improvement in energy consumption and latency compared to similar approaches in the literature.
- Communication overhead and cryptographic computation costs will be reduced by at least 20% compared to existing reference protocols.

compared to existing reference protocols. • The efficiency gains will be quantitatively proven through technical reports and comparative analyses and shared with industry/academic stakeholders.
Interim Outputs: • Comparison Report: Side-by-side performance evaluation with other protocols (e.g., Bonawitz). • Academic Publications (article, paper): Prepared and submitted versions. • Patent/Utility Model Application File (if applicable).

Risk Management:

Risk Identification	**Measures to Be Taken (Plan B)**
• Experimental results in the literature (e.g., failure to achieve the targeted 15–30% gain).	• Parameter fine-tuning, additional optimization, or protocol simplification steps. Publication of results with a partial improvement rate while remaining faithful to the project's primary objective.
• Timing/coordination issues with industry stakeholders.	• Organize prototype demo days on campus, or remotely online if necessary.

3.2. Research Opportunities

RESEARCH OPPORTUNITIES TABLE (*)

Infrastructure/Equipment Type, Model (Laboratory, Machinery-Equipment, etc.)	**Location Received Executing/Participating Institution/Organization**	**Purpose of Use in the Project**
High High-Performance Desktop Computer	Kadir Has University	Data processing, creation of a virtual test bed, conducting experiments in a virtual environment

4. WIDESPREAD IMPACT

4.1. Expected Outputs

Output Type	**Expected Output(s)**	**Expected Time Frame (*)**
Scientific/Academic Outputs (Paper, Article, Book Chapter, Book, etc.):	**At least 2 SCI/SCI-Expanded articles** will be prepared and the following journals will be targeted: • *IEEE Transactions on Industrial Informatics*	12–36 months

	• *Computers & Security (Elsevier)* • *Sensors (MDPI)* These articles will detail the relationship between **secure federated learning** and **multi-criteria optimization**, as well as **EKS/IoT scenarios**. **At least 2 international conference papers** will be prepared and submitted to conferences in the fields of security, IoT, and industrial control systems within the scope of *IEEE/ACM* (e.g., participation in the Universities Power Engineering Conference, Smart Energy Systems and Technologies conferences, International Computer Engineers Conference, International Management Information Systems Conference). **Domestic/international symposium presentations** will be held and project results will be shared with academic and industry stakeholders.	
Economic/Commercial/Social Outputs (Product, Prototype, Patent, Utility Model, Production License, Registration, Visual/Audio Archive, Inventory/Database/Documentation Production, Spin-off/Start-up Company, etc.):	- Secure and energy-efficient federated learning prototype (Python/C++-based) - Potential patent or utility model application (cryptographic protocol components) - Organizing workshops/demos with potential industry partners such as electricity distribution and automation companies	18–36 months
Researcher Training and New Project(s) Outputs for the Creation of (Master's/Doctoral/Medical Specialization/Art Proficiency Theses and National/International New Projects, etc.):	- Completion of thesis work by 1 doctoral and 1 master's student working on the project - Application for international funding such as EU Horizon Europe, EUREKA, etc. based on project results	24–36 months ( ), Post-Project

4.2. Expected Impacts

4.2.1. Anticipated Application Areas

When the project is successfully implemented, data sharing and machine learning processes for EKS/IoT devices in critical infrastructure sectors (energy, water distribution, transportation, etc.) can be made secure, low-latency, and energy-efficient. (See Twelfth Development Plan 3.2.3.6 Information and Communication Technologies, 3.5.2 Security Services). Institutions that will be directly affected by or benefit from/provide the project:

- **Electricity Distribution Companies:** Real-time analysis of data detected in SCADA systems through secure federated learning (e.g., fault/outage detection, cyberattack prevention) (See Plan 3.2.2.2 Energy; 3.5.2 Security Services).
- **Automation and IoT Manufacturers:** Processing data from sensors using secure collection modules in the prototype, thereby offering "data privacy compliant" solutions (See Plan 3.2.3.3 Science, Technology, and Innovation; 3.2.3.6 Information and Communication Technologies).
- **Public Institutions (Ministry of Energy, Ministry of Transportation, etc.):** Referencing project outputs in determining cybersecurity policies for critical infrastructure (See Plan 3.5.2 Security Services).

- **Defense and Security Sector:** Minimize security risks while increasing the use of FL (Federated Learning) in data-driven decision-making processes at strategic facilities (See Plan 3.2.2.3 Defense Industry; 3.5.2 Security Services).

At this point, the project outputs align with **the objective of "Developing advanced manufacturing technologies, digital transformation, and the cybersecurity ecosystem" included in the Twelfth Development Plan** (See Plan 3.2.3.3, 3.2.3.6, and 3.5.2). In line with the goal of increasing cybersecurity measures in critical infrastructures in Turkey and developing domestic and national technologies (see Plan 3.2.2.3, 3.2.3.3), the secure federated learning infrastructure provided by this project can make a significant contribution.

4.2.2. Socio-economic/Cultural Contribution

- **Quality of life and civil security:** By detecting cyberattacks early and strengthening prevention mechanisms, the project contributes to protecting vital services such as uninterrupted energy, safe transportation, and clean water supply on a societal scale.
- **Economic and Environmental Benefits:** Projected savings of up to 15% in energy consumption create a sustainable framework by reducing the battery and maintenance costs of IoT devices. This supports **clean and circular economy** principles.
- **Combating Climate Change:** Federated learning, which reduces inefficient data transfers and increases edge computing capacity, can partially alleviate dependence on data centers that consume large amounts of energy.
- **Education and Awareness:** Prototypes and publications to be developed may trigger new courses and research projects in **lifelong learning** and **higher education** institutions. It may become possible to develop curricula on FL and security issues in universities' control, computer, electrical-electronics, and energy departments.

All these contributions are directly related to the axes of **"digital transformation, cybersecurity, high value-added production, qualified workforce, and innovation"** specified in the Twelfth Development Plan. At the same time, it can contribute to **national policy** documents in the areas of "civil security" and "sustainable development."

4.3. Activity Plan to be Implemented within the Scope of Dissemination of Project Results and Science Communication

Target Audience: *The target audience expected to benefit from the outputs and results achieved during the project process (academics, policymakers and practitioners, private sector, individuals, specific age groups, etc.) is specified, along with the approach to be followed to reach the relevant target audience and how the target audience will benefit from the planned dissemination activities.*

- **Academics and Researchers:**
  - Academics working in the fields of computer, electronic, control, and energy engineering.
  - Cybersecurity and artificial intelligence research groups, cryptography experts.
- **Policy Makers and Implementers:**
  - Ministries (Energy, Transportation, etc.), relevant regulatory agencies.
  - Local governments (smart city projects), critical infrastructure managers.
- **Private Sector and Industry Representatives:**
  - Electricity distribution companies, automation firms, IoT manufacturers.
- **General Users and Educational Institutions:**
  - Individuals who value data privacy, groups seeking to increase cybersecurity awareness.
  - Undergraduate and graduate students.

Objectives and Expected Outcomes:

- **Raising Awareness:** Emphasizing the importance of secure federated learning in critical infrastructures, raising awareness of the concept of **"data privacy + artificial intelligence"** among the public and within the industry.
- **Sharing Knowledge:** Disseminating project results through academic articles and papers, industry training, and workshops. This supports the development of **qualified human resources**.
- **Technology Transfer: Promoting commercial productization** and **R&D partnerships** in the private sector through prototypes and patent/utility model studies.

partnerships in the private sector.

- **Contribution to Policy Development:** Providing data for policy design related to **digital transformation** and **cybersecurity** goals included in the Twelfth Development Plan based on the findings of the project.

Tools to be Used:

- **Scientific Events:** Conferences, seminars, workshops, and symposiums. Oral/poster presentations at national and international congresses.
- **Media and Publications:** Interviews and press releases for national/local media, articles in academic journals, popular articles in industry journals.
- **Sectoral Meetings:** Trade shows, industry days, demo or pilot demonstrations. Testing the prototype in real hardware environments.

Timing: *The time frame in which planned activities will be carried out and their duration are specified.*

- **0–6 months:** Creation of the project website and social media accounts, simple promotional materials (infographics, posters).
- **6–12 months:** Publication of the first literature report and design findings; pilot meetings with industry stakeholders.
- **12–24 months:** The first version of the prototype is completed, **national/international conference** papers are presented. Draft academic articles are prepared.
- **24–36 months:** Final prototype testing, patent/utility model application evaluation, **large-scale workshops** or demo days. Final versions of articles are published.
- **Post-Project:** Spin-off/start-up opportunities, new project applications (EU Horizon Europe, TEYDEB, etc.), development of industrial collaboration.

OTHER ISSUES YOU WISH TO MENTION

APPLICATION FORM

ATTACHMENTS ATTACHMENT

1: SOURCES

ATTACHMENT 2: BUDGET AND JUSTIFICATION

ATTACHMENT 3: OTHER PROJECTS AND CURRENT PUBLICATIONS OF THE PROJECT TEAM (This is automatically generated by the system based on the information entered into the Project Application System (PBS).)