

For use by the Project lecturer	Approved	Revision required
Feedback <div>For use by the Project module lecturer only</div>		

To be completed by the student					Language editor details	Language editor signature
PROJECT PROPOSAL 2024			Project no	TG4	Revision no	0
Title	Surname	Initials	Student no	Study leader (title, initials, surname)		
MR	BECKER	JM	20426471	Mr. T. Green		
Project title (the title on the project concept note)						
Intelligent Web Application Firewall using JSON Web Token Inspection						
				Student declaration I understand what plagiarism is and that I have to complete my project on my own.		Study leader declaration This is a clear and unambiguous description of what is required in this project. <u>Approved for submission (Yes/No)</u>
				Student signature		Study leader signature and date

1. Project description

What is the problem to be solved with your project? What is your project about? What does your system have to do?

Web applications are increasingly susceptible to malicious attacks from threat actors. Web application firewalls (WAF's) are systems that intercept and inspect bidirectional web responses between the web application and web clients to prevent malicious requests from reaching the web application. Current WAF's protect against vulnerabilities like cross-site forgery & scripting (XSS), SQL- and OS-injection and file inclusion. Next generation WAF's, or intelligent WAF's, are needed to protect against emerging threats and zero-day attacks like authentication bypasses and business logic flaws which existing WAF's do not cover.

The system will have to use anomaly detection through machine learning to identify and block malicious requests by inspecting JSON web tokens (JWT's) to prevent authentication bypass and business logic flaws. The system will also have to provide typical WAF functions like preventing cross-site forgery & scripting (XSS), SQL- and OS-injections and file inclusions. Current WAF's like the AWS-, Azure-, Cloudflare-, F5-, NGINX- and open-appsec WAF's have competitive specifications but lack on advanced features needed by next-generation WAF's to block emerging threats and zero-day attacks. WAF's provide a higher layer of security on the application level that specifically protect the web application from threats. JWT's contain the payload and user information needed by next-generation WAF's which can be leveraged to train WAF's, giving it an advantage over exist

2. Technical challenges in this project

Describe the technical challenges that are *beyond* those encountered up to the end of third year and in other final year modules.

2.1 Primary *design* challenges

Which aspects of the design of the system do you expect to be the most challenging?

- The first challenge is to achieve a balance between the **True Positive** (malicious traffic)- and True Negative (intended traffic) Filtering Accuracy is needed to assure the system does not filter out the legitimate requests together with the malicious requests.
- The second challenge is to achieve an accurate WAF where vulnerabilities like Cross-Site Scripting & Forgery, File Inclusion, SQL Injection and JWT Authentication Bypass need to be prevented by efficiently and accurately implementing WAF tools like anomaly detection using JWT payloads.
- Lastly, the WAF needs to effectively build a baseline for a web application which can then be applied as a profile to different web applications where the baseline can be trained per website in order to provide security to a range of applications

2.2 Primary *implementation* challenges

Which aspects of the implementation to you expect to be the most challenging?

- The use and integration of different Web Applications will be a major challenge to implement on limited hardware due to the processing and memory constraints.
- The use and integration of client testing software like **Selenium on a embedded host** device is expected to be resource heavy on an embedded device.
- The integration of the WAF on an embedded device will prove to be difficult to implement on a system with limited hardware speed.
- The integrated system is expected **to drop** requests if the system is overwhelmed and this needs to be mitigated by utilising FIFO queues with sufficient overflow buffers.

3. Functional analysis

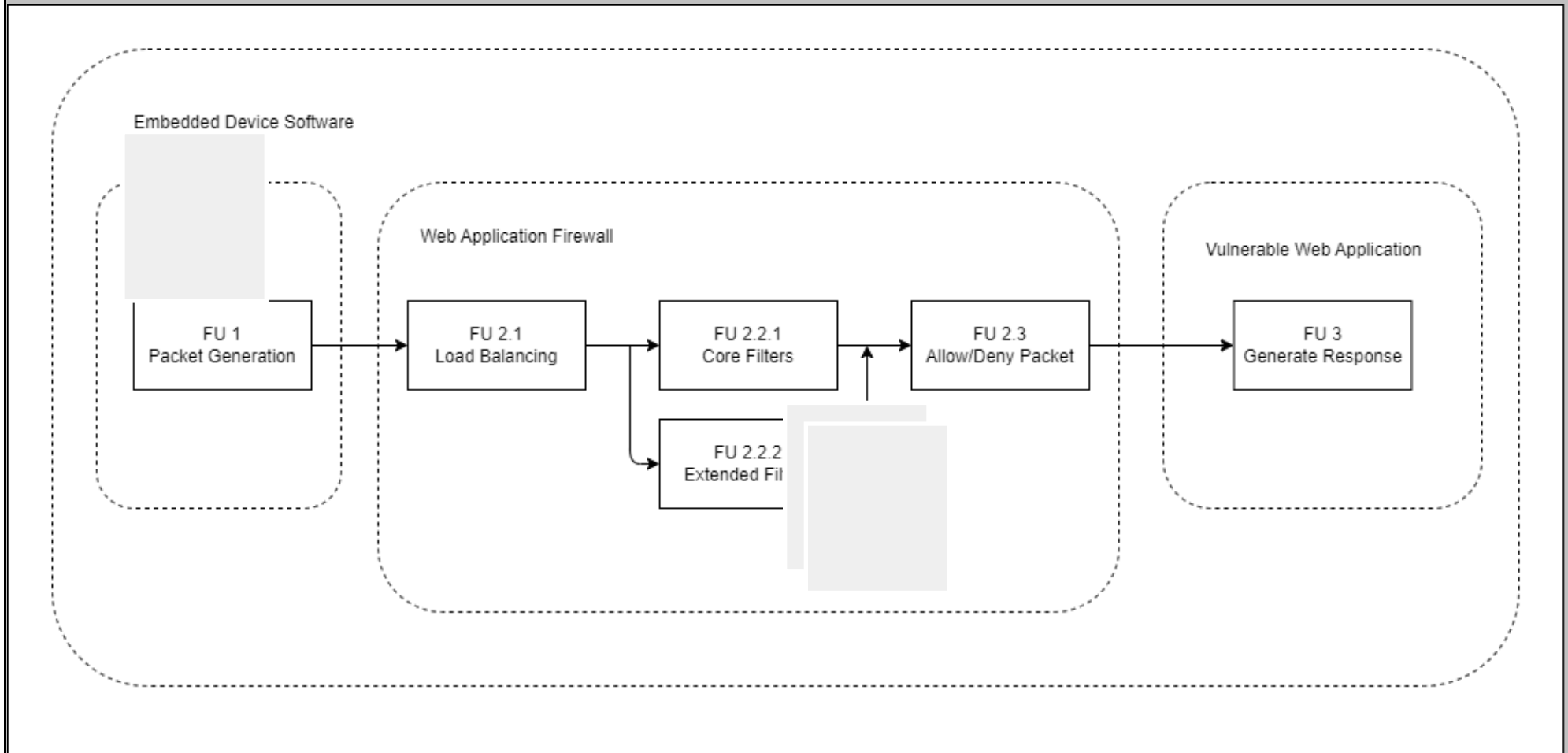
3.1 Functional description

Describe the design in terms of system functions as shown on the functional block diagram in section 3.2. This description should be in *narrative format*. **DO NOT** use a bullet list.

The system has the overall goal of filtering out malicious packets from an infected host trying to cause damage to a series of vulnerable web applications. The system starts by generating either malicious or intended traffic as a client (FU 1). The traffic then travel to the WAF (FU 2) where it reaches a load balancer (FU 2.1) where the traffic is then distributed to a series of WAF filters where core WAF functions like IP Fencing, Geo-fencing and Geo-blocking, Requests Inspection, Response Inspection, Security Rules, Anomaly Scoring, DDOS Rate Limits, Bot Mitigation and Threat Intelligence (FU 2.2.1) as well as next-generation extended functions like JWT Anomaly Detection with machine learning algorithms (FU 2.2.2) are implemented to give a request a threat score. The system needs to generalise well on different Web Applications, profiling each Web Application to keep the requests and responses of the different applications separate in order to more accurately filter the traffic. Its goal is to capture packets and analyse the payload in order to determine if the request is an anomaly, and then either blocking, allowing or further investigating it based on the the anomaly score (FU 2.3). If the packet is blocked, its response is altered and returned to the client as a 406 error. However, if the packets are allowed it reaches the vulnerable web application (FU 3) where it is processed and the appropriate response is generated for the client. This keeps the vulnerable web application safe from threat actors attempting to launch malicious attacks on the web server.


The client, the WAF and the vulnerable Web Application are to be implemented on a single embedded device for simplicity and to maximise system resources. The system will provide the same efficiency provided by existing solutions while providing an extra layer of security by blocking JWT Authentication Bypass and business logic flaws, something that is not offered in existing WAF's.

3.2 Functional block diagram (this should not be a flow diagram)



4. System requirements and specifications

These are the core requirements of the system or product (the mission-critical requirements) in table format **IN ORDER OF IMPORTANCE**. Requirement 1 is the most fundamental requirement.

	Requirement 1: the fundamental functional and performance requirement of your project	Requirement 2 (Number 2 in the order of importance)	Requirement 3 (Number 3 in the order of importance)
1. Core mission requirements of the system or product. Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.	The core requirement is to produce a next-generation WAF that performs on par with current solutions.	The secondary requirement is to produce a next-generation WAF which can protect against the same threats as existing solutions.	To produce a next-generation WAF which can protect against the next-generation threats with the same efficacy as already common threats.
2. What is the target specification (in <i>measurable</i> terms) to be met in order to achieve the requirement in 1. above?	The target specification is to achieve a 90% Overall Balanced Filtering Efficacy, i.e. the  unt of total requests returned to the client the total amount of requests sent.	The target specification is to achieve a 90% Balanced Filtering Efficacy, i.e. the amount of total requests returned to the client over the total amount of requests sent.	The target specification is to achieve a 90% Balanced Filtering Efficacy, i.e. the amount of total requests returned to the client over the total amount of requests sent.
3. Motivation: Defend the <u>specific</u> target specification, i.e. the <u>value that you selected</u> . I.e., <i>why</i> will meeting the specification given in point 2 above <i>solve the problem</i> ?	The Overall Balanced Filtering Efficacy is chosen to be 90% as this is on par with current products in industry . This implies a 95% True Positive Filtering Efficacy and an 85% True Negative Filtering Efficacy rate.	The Balanced Filtering Efficacy for the current features is chosen to be 90% as this is on par with current products in industry. This implies a 95% True Positive Filtering Efficacy and an 85% True Negative Filtering Efficacy.	The Balanced Filtering Efficacy for the next-generation features is chosen to be 90% as this is on par with current products in industry. This implies a 95% True Positive Filtering Efficacy and an 85% True Negative Filtering Efficacy.
4. How will you demonstrate at the examination that this requirement and specification (points 1 and 2 above) have been met? Be explicit about how you will <i>prove</i> these were met.	The client testing software will send requests to the Web Application where the trained WAF will filter the requests based on whether it classifies as malicious or not. The results will be displayed to the WAF administrator.	The client testing software will send requests with common threats to the Web Application where the trained WAF will filter the malicious requests based on whether it classifies as a threat or not.	The client testing software will send requests with next-generation threats to the Web Application where the trained WAF will filter the malicious requests based on whether it classifies as a threat or not.
5. Your own design contribution: what are the aspects that <i>you will design and implement yourself</i> to meet the requirement in point 2? <u>If none, remove this requirement.</u>	My own contribution would be the JWT decryption and inspection .	My own contribution would be the profiling per web application where Machine Learning algorithms for training the model is used to filter out most malicious traffic	My own design contribution would be by the baseline algorithms such as anomaly detection for next-generation WAF functions.
6. What are the aspects to be taken off the shelf to meet this requirement? If none, indicate "none". Clearly specify for what tasks library functions will be used (if relevant to the project).	The aspects that would be taken off the shelf are the baseline algorithms for existing WAF features as well as the embedded platform and operating system.	The aspects that would be taken off the shelf are the filtering specifications for existing WAF functions.	None

System requirements and specifications page 2

	Requirement 4	Requirement 5	Requirement 6
1. Core mission requirements of the system or product. Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.			
2. What is the target specification (in measurable terms) to be met in order to achieve the requirement in 1. above?			
3. Motivation: Defend the <u>specific</u> target specification, i.e. the <u>value that you selected</u> . I.e., <i>why</i> will meeting the specification given in point 2 above solve the problem?			
4. How will you demonstrate at the examination that this requirement and specification (points 1 and 2 above) have been met? Be explicit about how you will <i>prove</i> these were met.			
5. Your own design contribution: what are the aspects that <i>you will design and implement yourself</i> to meet the requirement in point 2? <u>If none, remove this requirement.</u>			
6. What are the aspects to be taken off the shelf to meet this requirement? If none, indicate "none". Clearly specify for what tasks library functions will be used (if relevant to the project).			

System requirements and specifications page 3

	Requirement 7	Requirement 8	Requirement 9
1. Core mission requirements of the system or product. Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.			
2. What is the target specification (in measurable terms) to be met in order to achieve the requirement in 1. above?			
3. Motivation: Defend the <u>specific</u> target specification, i.e. the <u>value that you selected</u> . I.e., <i>why</i> will meeting the specification given in point 2 above solve the problem?			
4. How will you demonstrate at the examination that this requirement and specification (points 1 and 2 above) have been met? Be explicit about how you will <i>prove</i> these were met.			
5. Your own design contribution: what are the aspects that <i>you will design and implement yourself</i> to meet the requirement in point 2? <u>If none, remove this requirement.</u>			
6. What are the aspects to be taken off the shelf to meet this requirement? If none, indicate "none". Clearly specify for what tasks library functions will be used (if relevant to the project).			

System requirements and specifications page 4

	Requirement 10	Requirement 11	Requirement 12
1. Core mission requirements of the system or product. Focus on requirements that are core to solving the engineering problem. These will reflect the solution to the problem.			
2. What is the target specification (in measurable terms) to be met in order to achieve the requirement in 1. above?			
3. Motivation: Defend the <u>specific</u> target specification selected, i.e. the value. <i>Why</i> will meeting the specification given in point 2 above <i>solve the problem</i> ?			
4. How will you demonstrate at the examination that this requirement and specification (points 1 and 2 above) have been met? Be explicit about how you will <i>prove</i> these were met.			
5. Your own design contribution: what are the aspects that <i>you will design and implement yourself</i> to meet the requirement in point 2? If none, <i>remove this requirement</i> .			
6. What are the aspects to be taken off the shelf to meet this requirement? If none, indicate "none". Explicitly indicate what tasks library functions will be used for (if relevant to the project).			

5. Field conditions

These are the REAL WORLD CONDITIONS under which your project has to work and has to be demonstrated.

	Real world field condition 1	Real world field condition 2	Real world field condition 3
Field condition requirement. In which field conditions does the system have to operate? Describe the one, two or three most important field conditions.	The system must be able to operate in an enclosed and air conditioned room at room temperature to simulate data center conditions close to servers.		

6. Student tasks

6.1 Design and implementation tasks

List your primary design and implementation tasks in bullet list format (5-10 bullets). These are *not* product requirements, but *your* tasks.

- The client, vulnerable web server and WAF needs to be designed and implemented on an embedded platform.
- The system needs to be simulated on a PC using Python where hardware constraints are not a problem and then scaled to fit on the embedded platform.
- The system needs to effectively prevent OS- and SQL-Injection, Cross-Site scripting, File Inclusion and JWT authentication bypass attacks to protect the vulnerable web server.
- The system will display the result of the tests to the WAF administrator.

6.2 New knowledge to be acquired

Describe what the theoretical foundation to the project is, and which new knowledge you will acquire (*beyond* that covered in any other undergraduate modules).

- The student needs to learn what JWT's are and how they are used to bypass authentication measures
- The student needs to learn what WAF's are and how security vulnerabilities in web applications are prevented.
- The student needs to learn what classifier based algorithms are and how they can be used to model user behaviour and prevent out of the ordinary behaviour.
- The students needs to learn how web application servers and their clients communicate in order to intercept the communication and how that is exploited.
- The student needs to learn how packet filtering works and how it can be utilised in machine learning in order to predict user behaviour.
- The students needs to learn what packet inspection is and how it is utilised by WAF's to determine the intent of the package.