

# EMERGING COMPUTER TECHNOLOGIES

*Edited by*  
**Dr. Muhammet DAMAR**  
**Dr. Ömer AYDIN**

There is rapid development and change in the field of computer science today. These affect all areas of life. Emerging topics in computer science are covered in this book.

In the first chapter, there is a log data analysis case study that aims to understand Hadoop and MapReduce by example. In the second chapter, encrypted communication has been tried to be provided on IoT devices performing edge computing. In this way, it is aimed to make communication secure. Arduino is used as an IoT device. In the encryption process, AES encryption is used with 128-bit and 256-bit key lengths. The third chapter presents a more secure connection between the NodemCU and Blynk, using the AES algorithm. it is aimed to prevent a vulnerability in the connection of Google Home devices with IoT during the Blynk IoT connection phase. The next chapter presents a methodological tool based on an evaluation framework for integration of digital games into education (MEDGE), expanded by adding additional information from the students, MEDGE+. The fifth and last chapter proposes a disaster management system utilising machine learning called DT-DMS that is used to support decision-making mechanisms.

[www.izmirakademi.org](http://www.izmirakademi.org)

Yayınevi | Akademik Dergiler | STK  
Akademik Etkinlikler | Proje | Eğitim

@izmirakademidernegi @35izmirakademi @izmirakademidernegi



İZMİR AKADEMİ  
DERNEĞİ



EMERGING COMPUTER TECHNOLOGIES

Dr. Muhammet DAMAR & Dr. Ömer AYDIN

# EMERGING COMPUTER TECHNOLOGIES



*Edited by*  
**Dr. Muhammet DAMAR**  
**Dr. Ömer AYDIN**



İZMİR AKADEMİ  
DERNEĞİ

# ***EMERGING COMPUTER TECHNOLOGIES***

---

Editor  
Dr. Muhammet DAMAR  
Dr. Ömer AYDIN



**İZMİR AKADEMİ  
DERNEĞİ**

# EMERGING COMPUTER TECHNOLOGIES

**EDITOR** : Dr. MUHAMMET DAMAR  
**ORCID** : 0000-0002-3985-3073  
**EDITOR** : Dr. ÖMER AYDIN  
**ORCID** : 0000-0002-7137-4881  
**E-ISBN** : 978-605-71029-2-8

**DATE OF PUBLICATION:** NOVEMBER 2021

©Copyright 2021, İZMİR AKADEMİ DERNEĞİ. CERTIFICATE NO: 51876

All rights of this edition belong to İzmir Akademi Derneği. Without the written permission of the publisher, the whole or part of the book may not be printed, published, reproduced or distributed electronically, mechanically or by photocopying.

İzmir Akademi Derneği has started to publish regularly at the national and international levels since 2021.

**REDACTION** : Emel AYDIN  
**PAGE DESIGN** : Dr. Ömer AYDIN  
**COVER DESIGN** : grafiklite ( <https://bionluk.com/grafikalite> )  
**COVER PICTURE** : canva.com (Free pictures)

## LIBRARY INFORMATION CARD:

**Eds:** Muhammet DAMAR, Ömer AYDIN  
Emerging Computer Technologies / **Eds:** Muhammet Damar, Ömer Aydın  
Date of Publication: November, 2021  
1. Edition. VIII + 67 s. 16x24 cm. There is a bibliography, no index.  
E-ISBN: 978-605-71029-2-8  
1. Hadoop and MapReduce 2. Edge Computing 3. Secure Connection 4. Practical Evaluation 5. Serious Games 6. Digital Twin 7. Disaster Management System

## DISTRIBUTION

**Distribution** : İzmir Akademi Derneği (Certificate No: 51876)  
**Address** : Buyukkale Mh. Kerimoglu Sk. No:40-2 Tire/İzmir, TURKEY  
**Email** : izmirakademidernegi@gmail.com, info@izmirakademi.org  
**Phone** : +90 555 597 19 84  
**Web** : [www.izmirakademi.org](http://www.izmirakademi.org)



# EMERGING COMPUTER TECHNOLOGIES

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned. Nothing from this publication may be translated, reproduced, stored in a computerized system or published in any form or in any manner, including, but not limited to electronic, mechanical, reprographic or photographic, without prior written permission from the publisher İZMİR AKADEMİ DERNEĞİ. The individual contributions in this publication and any liabilities arising from them remain the responsibility of the authors. The publisher and the editor is not responsible for possible damages, which could be a result of content derived from this publication.

The authors of the relevant chapter are fully responsible for the content in this publication. When the authors send the book chapter for publication, they are deemed to have committed to abide by all ethical rules. All material and moral responsibility that may arise due to the chapters in this book belongs to the relevant chapter authors. The authors accepted these conditions in advance by participating in this book as chapter authors. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of going to press, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.



# **CHAPTERS and AUTHORS**

---

## **1. CHAPTER**

Understanding Hadoop and MapReduce by example: log data analysis case study  
Gligor Risteski, Mihiri Chathurika, Beyza Ali, Atanas Hristov

## **2. CHAPTER**

Edge Computing Security with an IoT device  
Beyda Nur Kars

## **3. CHAPTER**

Secure Connection between Google Home and IoT Device  
Ekrem Yiğit

## **4. CHAPTER**

Practical Evaluation on Serious Games in Education  
Slavica Mileva Eftimova, Ana Madevska Bogdanova, Vladimir Trajkovik

## **5. CHAPTER**

Digital Twin Based Disaster Management System Proposal: DT-DMS  
Özgür Doğan, Oğuzhan Şahin, Enis Karaarslan

## SCIENTIFIC BOARD

### EDITORS

**Dr. Muhammet DAMAR**

Dokuz Eylul University  
ORCID: 0000-0002-3985-3073

**Dr. Ömer AYDIN**

Manisa Celal Bayar University  
ORCID: 0000-0002-7137-4881

### BOARD

Prof. Dr Ahmet KAYA  
Ege University, Turkey

Assoc. Prof. Dr Mehmet Hilal OZCANHAN  
Dokuz Eylul University, Turkey

Prof. Dr Vahap TECIM  
Dokuz Eylul University, Turkey

Assist. Prof. Dr Arvind Dhaka  
Manipal University Jaipur, India

Prof. Dr Yalcin CEBI  
Dokuz Eylul University, Turkey

Assist. Prof. Dr Enis KARAARSLAN  
Mugla Sitki Kocman University, Turkey

Assoc. Prof. Dr Amita Nandal  
Manipal University Jaipur, India

Assist. Prof. Dr Mehmet Cem BÖLEN  
Ataturk University, Turkey

Assoc. Prof. Dr Atanas Hristov  
University for Information Science  
and Technology, North Macedonia

Dr Ilker Ali  
International Vision University, North  
Macedonia

Assoc. Prof. Dr Aybeyan SELIM  
International Vision University, North  
Macedonia

Dr Muhammet DAMAR  
Dokuz Eylul University, Turkey

Assoc. Prof. Dr Gokhan DALKILIC  
Dokuz Eylul University, Turkey

Dr Ramiz YILMAZER  
Aviarsis Technology, Turkey

## FOREWORD

---

There is rapid development and change in the field of computer science today. These affect all areas of life. Emerging topics in computer science are covered in this book.

In the first chapter, there is a log data analysis case study that aims to understand Hadoop and MapReduce by example. In the second chapter, encrypted communication has been tried to be provided on IoT devices performing edge computing. In this way, it is aimed to make communication secure. Arduino is used as an IoT device. In the encryption process, AES encryption is used with 128-bit and 256-bit key lengths. The third chapter presents a more secure connection between the NodemCU and Blynk, using the AES algorithm. it is aimed to prevent a vulnerability in the connection of Google Home devices with IoT during the Blynk IoT connection phase. The next chapter presents a methodological tool based on an evaluation framework for integration of digital games into education (MEDGE), expanded by adding additional information from the students, MEDGE+. The fifth and last chapter proposes a disaster management system utilising machine learning called DT-DMS that is used to support decision-making mechanisms.

Dr. Muhammet DAMAR  
Editor

Dr. Ömer AYDIN  
Editor

# CONTENT

---

| <i>Title/Author</i>                                                                                                                                                       | <i>Page</i> |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| <b><i>Understanding Hadoop and MapReduce by example:<br/>log data analysis case study</i></b><br><i>Gligor Risteski, Mihiri Chathurika, Beyza Ali, Atanas<br/>Hristov</i> | 1           |
| <b><i>Edge Computing Security with an IoT device</i></b><br><i>Beyda Nur Kars</i>                                                                                         | 17          |
| <b><i>Secure Connection between Google Home and IoT<br/>Device</i></b><br><i>Ekrem Yiğit</i>                                                                              | 31          |
| <b><i>Practical Evaluation on Serious Games in Education</i></b><br><i>Slavica Mileva Eftimova, Ana Madevska Bogdanova,<br/>Vladimir Trajkovik</i>                        | 41          |
| <b><i>Digital Twin Based Disaster Management System<br/>Proposal: DT-DMS</i></b><br><i>Özgür Doğan, Oğuzhan Şahin, Enis Karaarslan</i>                                    | 53          |



# Understanding Hadoop and MapReduce by example: log data analysis case study<sup>1</sup>

---

## *Gligor Risteski*

---

- ❖ **Affiliation:** Faculty of Communication Network and Security University of Information Science and Technology, Ohrid, North Macedonia
- ❖ **Email:** gligorr@gmail.com

---

## *Mihiri Chathurika*

---

- ❖ **Affiliation:** Faculty of Communication Network and Security University of Information Science and Technology, Ohrid, North Macedonia
- ❖ **Email:** mihiritxla@gmail.com

---

## *Beyza Ali*

---

- ❖ **Affiliation:** Faculty of Communication Network and Security University of Information Science and Technology, Ohrid, North Macedonia
- ❖ **Email:** beyzaali34@gmail.com

---

## *Atanas Hristov*

---

- ❖ **Affiliation:** Faculty of Communication Network and Security University of Information Science and Technology, Ohrid, North Macedonia
- ❖ **Email:** atanas.hristov@uist.edu.mk

---

<sup>1</sup> This study has been published in the Journal of Emerging Computer Technologies.

## **INTRODUCTION**

The term “Big Data” is gaining more popularity every day. The first thing we should know about it is that it does not have a commonly held definition. Basically, as one can understand from its name, Big Data means a big amount of data. Sethy, R. [1] in his article defines "Big Data describes any massive volume of structured, semi-structured and unstructured data that are difficult to process using a traditional database system."

Researches show that data volumes are doubling every year. Although there is not a specific reason behind this rapid growth rate, the new data sources, contribute to that growth highly. Smartphones, tablet computers, sensors, and all other devices that can be connected to the internet generate a vast amount of data. Enterprises improve their technological infrastructures and adopt more powerful platforms, which play an important role in the growth rate of the data that is generated [2].

## **HADOOP**

Hadoop is a collection of open-source utilities which allows the use of network to deal with the problems which include big amounts of data. Hadoop provides framework for distributed storage and framework for processing big data with MapReduce programming model. The core of Apache Hadoop framework contains the following parts:

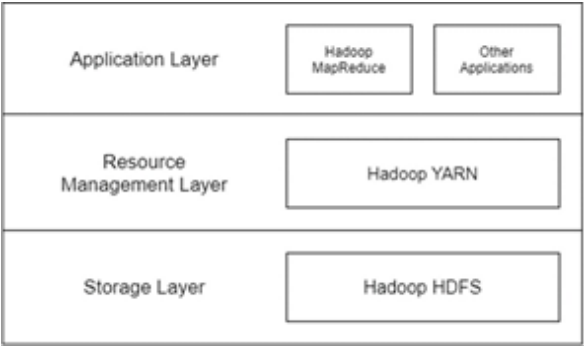
- Hadoop Common
- Hadoop Distributed File System (HDFS)
- Hadoop YARN
- Hadoop MapReduce

Hadoop common represent libraries and utilities which are needed by other Hadoop parts in order to operate. Hadoop distributed File System (HDFS) is a distributed file-system that stores data on commodity hardware, allowing very high bandwidth across the cluster. Hadoop YARN is a platform responsible for managing computing resources in clusters and uses them for scheduling users' applications. Lastly, Hadoop MapReduce is a programming model for big data processing in the cluster.

Java is a programming language that is mostly used for writing MapReduce programs but Hadoop allows the use of any programming language to write MapReduce programs.

To exploit the parallel processing that Hadoop gives, we have to express our query as a MapReduce job. After some local, little scope testing, we can have the option to run it on a cluster of machines.

The Hadoop structure is demonstrated in Figure 1.



**Fig 1.** Hadoop Structure

**Hadoop Distributed File System (HDFS)**

Hadoop Distributed File system is one of the most reliable storage systems designed to store a smaller number of large files rather than a greater number of small files. Among many of the features HDFS provides, the fault-tolerant storage layer can be mentioned as one of the most important features.

Replication of data in the Hadoop file system helps the user to attain this feature. Even in situations where hardware failure happens, the data reliability is still high [3].

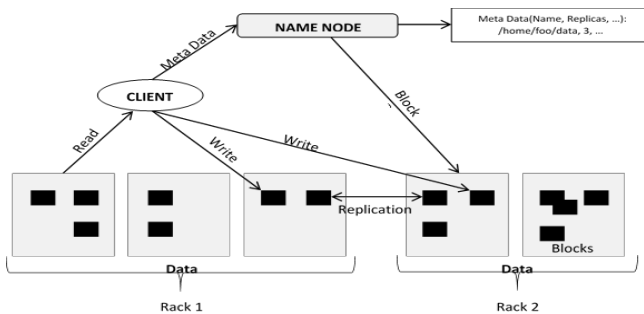
Apache Yarn, introduced in Hadoop 2.x, is the resource management layer of Hadoop, which is also used for job scheduling and data operating system. It allows different data processing engines to run and process data stored in HDFS. Some of these processing engines are graph processing, interactive processing, stream processing, and batch processing. Data processing platform Yarn has the functionality named MapReduce, which empowers Hadoop by allowing processing numerous different frameworks on the same hardware where Hadoop is deployed [4].

HDFS was developed using distributed file system design and is designed using low-cost hardware. It is more fault-tolerant than other distributed systems. HDFS is capable of holding larger amounts of data providing easy access and parallel processing. In HDFS files are stored across multiple machines in order to prevent possible data losses in case of system failure [3]. Features of HDFS are: distributed storage and processing, command interface, checking the status of clusters, streaming access to the file system, file permission, and authentication.

- Distributed storage and processing
- Command interface to interact with HDFS
- Checking the status of cluster easily
- Streaming access to file system data
- File permissions and authentication

### ***HDFS Architecture***

HDFS follows the master-slave architecture. It has elements such as namenode, datanode, and blocks, where the built-in servers of the first two elements enable users to easily check the status of the cluster. Similarly, they are commodity hardware [point]. In Figure 2 the architecture of the Hadoop file system is demonstrated.



**Fig 2.** Architecture of a Hadoop File System

- Namenode

The namenode, containing GNU/Linux operating system and the namenode software, can act as the master server running on commodity

hardware. Renaming and opening/closing files/directories are executed in this element. It also manages the file system and regulates clients' access to files.

- Datanode

Similar to namenode, the datanode too, contains the GNU/Linux operating system and datanode software. For every node in a cluster, there exists a datanode, which is able to manage the data storage of the system. Depending on the instructions of the namenode, datanode performs block creation, block deletion, and block replication. It also performs read/write operations on the file system when requested by a client.

- Block

. Usually, the user data are stored in files such that a file in a file system is divided into segments and then are stored in individual data nodes. Those file segments are called blocks. A block is the minimum amount of data that HDFS can read/write. The block size can be increased if needed but the normal size is 64MB.

## **Apache Hadoop Yarn**

To create a split between Resource Manager (RM) and Application Master (AM), YARN separates the functions of resource management and job scheduling into separate daemons. An application can be an individual job or a DAG of jobs.

- ResourceManager (RM)

Resource Manager (RM) together with the Node Manager (NM) comprise the data-computation framework. The RM adjudicates the resources in the system and NM is responsible for containers and monitoring resource usage. In other words, RM is the ultimate authority and NM is the framework agent. RM has two main components known as Scheduler and Application Manager.

- Scheduler

Scheduler controls the allocation of resources to the several running applications. It performs the scheduling function depending on the resource requirements of applications, but it does not take any responsibility for

performing monitoring, tracking the status of applications or application or hardware failures.

- ApplicationsManager

Applications Manager: responsible for restarting the AM container in case of a failure. It also accepts job submissions.

- ApplicationMaster (AM)

The AM is responsible for requesting resources from the RM and then executing and monitoring the tasks. It works together with NM when executing the tasks.

The main idea behind using YARN on Hadoop is the notion of resource reservation via the Reservation System. The Reservation System tracks the resources, performs admission control for reservations, and reserves resources to ensure the execution of important jobs [4].

YARN supports the notion of Federation via the YARN Federation feature. The idea behind this is to scale Yarn up to very large amounts of nodes by wiring YARN clusters and sub-clusters. The Federation feature makes this transparent wiring of clusters appear as a single big cluster [5].

## ***THE ROLE OF MAPREDUCE IN HANDLING BIG DATA***

Big Data means "big power" when handled efficiently. It can give new aspects to the enterprises, like which strategy will increase the profitability, which customers buy which products, the current situation of the company versus the situation of the competitors, and so on.

As the data comes from different sources and different structures it is important to categorize it with respect to some characteristics of the data. The most important and the most known characteristics of Big Data are known as the "3Vs of Big Data" where the Vs stands for volume, variety, and velocity. Volume refers to the amount of data, variety refers to the type of data, i.e. text, image, video, etc. and velocity refers to the speed at which the data comes from different sources [6].

Big Data comes with its own set of problems that need to be resolved. Processing power, storage, data issues, and cost are the most important problems. The old techniques for working with or analyzing information are

not enough to deal with Big Data. Therefore, new technologies are needed and this is where MapReduce comes into the picture [2].

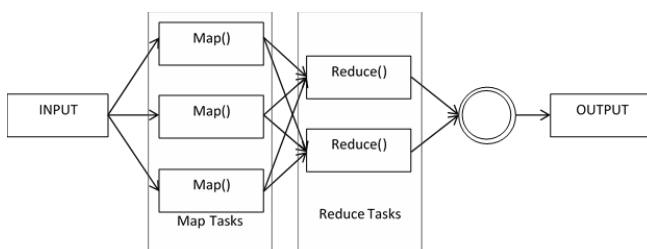
## **MapReduce**

MapReduce is a programming model that is used for accessing and processing big data stored in HDFS. Programs written in MapReduce are executed on a distributed system where big data is split into smaller chunks of data and are executed in parallel [7]. MapReduce has two functionalities, Map () and Reduce (). This model has been used in Google's search index, machine learning, and statistical analysis [8]. Implementation of MapReduce is highly scalable and easy to use. The run-time system allows programmers with no prior knowledge or experience with parallel processing to utilize the resources of distributed systems easily, by handling details like partitioning the inputs, scheduling the program's execution, handling failures, and managing inter-machine communication [7, 9].

Although it is impossible to prevent failures, the objective is to minimize the probability of failure to a level that will not harm the overall process. Two methods that would help to increase the "fault tolerance" in Big Data are the following [6]:

- First divide the whole computation into smaller tasks and then assign each task to a different node.
- Assign a node to observe if the other nodes are working properly. In case a node fails to complete its task, the task is restarted. But this may cause a complication in the process if some tasks are recursive.

MapReduce is one of the core building blocks of processing in the Hadoop framework. Hadoop uses the MapReduce algorithm to run the applications in parallel. It provides the necessary solution to keep the process going since it can survive failures without losing data. In 2004 Google published about MapReduce technology [7]. MapReduce comprises of two distinct tasks: Map and Reduce. Mapping is the first phase and Reducing happens after the Mapping phase is completed. In Map phase data is processed and key-value pairs are produced. This is known as the map job. Then the produced key-value pairs are fed into the Reducer. After collecting all the key-value pairs from all of the map jobs the Reducer groups the pairs into a smaller set of key-value pairs, producing the final output [7, 8].



**Fig 3.** How does MapReduce work?

## Algorithm of MapReduce

MapReduce program has three stages of executing: map, shuffle, and reduce.

- **Map stage:** at this stage, the map job processes the input data which is in the form of a file or directory. As an output of the map job several small chunks of data are produced.
- **Shuffle stage:** at this stage, the output of the map stage is accepted and the relevant records are consolidated.
- **Reduce stage:** reduce stage acts together with the shuffle stage. The reducer processes the data that comes from the map stage and produces a new output which is then stored in the Hadoop file system.

During the process, a task from Map and Reduce phases are sent to the appropriate servers in the cluster. All the data passing details like issuing tasks, verifying task completion, copying data between nodes, are managed by the MapReduce framework. Because the computing is performed on nodes using the data from local disks the network traffic is reduced significantly. After the tasks are completed the results are sent back to the Hadoop server [10].

## MapReduce with Python

‘mrjob’ is a Python library for MapReduce for writing and running Hadoop streaming jobs. It is created by Yelp. When a MapReduce job is written using ‘mrjob’, it can be tested locally and run on a Hadoop cluster or in the cloud. Using ‘mrjob’ for writing MapReduce applications has many advantages. Some of them are:

- It is a dynamically developing framework.

- It has extensive documentation.
- Installing Hadoop is not enforced. Applications written using 'mrjob' can be executed and tested without installing Hadoop.
- It allows the MapReduce applications to be written in a single class rather than writing separate programs for Map and Reduce phases.

Although it provides a great solution, 'mrjob' has its disadvantages. The most important disadvantage is that it does not provide the level of access to Hadoop that other APIs provide. This is because it is a simplified library [11].

### ***Log Analyzer with Hadoop***

Log analysis is both art and science which aims to make sense out of computer-generated records. These records are called log or audit trail records. The process of creating these records is known as data logging [12]. Some of the most important reasons for performing log analysis are:

Understanding user behavior

- System troubleshooting
- Proper resource allocation
- Improved business operations
- Improved security

Achieve compliance

Data centers generate thousands of terabytes or petabytes of log files every single day. It is very challenging to store and analyze these data not only because of its large volume but also because of the different structure of log files. Due to not being able to deal with a large volume of log files efficiently, conventional database solutions run short for the needs in log analysis. As a result of the comparison of SQL DBMS and Hadoop MapReduce in [13], Hadoop MapReduce overperforms DBMS in the means of tuning up with the task and loading data. As it can be seen from this result, with the unprecedented increase in the data generated traditional methods fall short with providing a solution for data analysis. This is, exactly, the point where the new technologies stepped in [8]. Hadoop MapReduce has a wide area of

applications for Big Data analysis [3], [9],[11]. The true power of Hadoop lies in its ability to scale up to a great number of computers, where each computer has several processor cores, by connecting commodity computers to work in parallel. This plays an important role in log analysis as it can benefit thousands of nodes which will store multiple blocks of log files.

In this paper, we propose an idea on how Hadoop can be used to analyze web server logs, in our case Nginx access log. Web server access logs are generated by the web servers all the time, recording all accesses on the hosted web pages. This means that the access logs can be very big. The web access log contains information about time, IP addresses, browser type, etc. All of this information is important for the system administrators as it provides information about system usage, security, and system troubleshooting.

This idea is proposed for analyzing one kind of logs only, but Hadoop can be used in every situation where big log files are generated, such as system logs, logs from some business application, etc.

The real-world usage (practical usage) of this system can be implementing it as a base of a larger system used by many users for log analyzing.

### ***Environment Setup***

We have installed Hadoop for demonstrating purpose on two virtual machines hosted on Digitalocean. Each virtual machine has 2 cores CPU and 8GB of RAM memory (installing is explained in details in Installing Hadoop section) with installed Ubuntu 16.04. Also, for testing purpose we have installed Hadoop on one local virtual machine with 1 core CPU and 7GB of RAM memory on hosted hypervisor VMware Workstation Pro 15 using Bitnami Hadoop Stack image. In this project the log generated from Python Fake Logs script is used [14].

### ***Results***

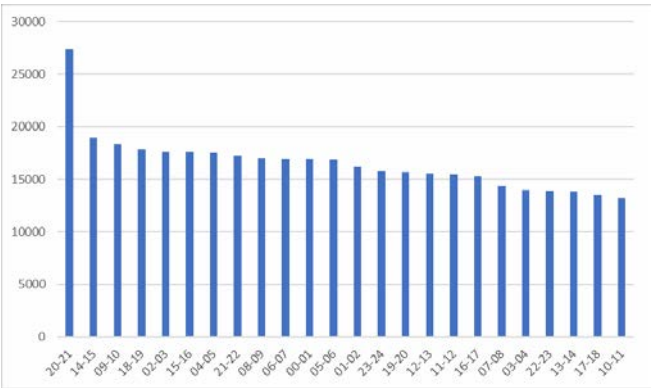
Our aim was to validate the role and importance of Hadoop in meeting the challenge of dealing with Big Data, by an analysis of log files in order to evaluate the system activities. We benefitted from Hadoop's ability to scale up to as many computers as needed. Firstly, log files are broken up into blocks with MapReduce class created with mrjob python library (dividing into blocks

is explained in section 3.1). Blocks are then distributed over the nodes in a Hadoop cluster. With parallel computation, the job is divided into a number of tasks, which in return improves the performance. In our log analysis, we used Python's 'mrjob' library.

We use a dummy log file only for testing. Our log file is the Nginx access log file. The log file contains information about:

- visitor's IP address,
- date and time of the access,
- visited page,
- type of request,
- type of the user's web browser,
- type of the user's operating system.

In our case, two reports are extracted from the Nginx access file as an example. These two reports are only an example of what can be achieved. The following charts demonstrate the results of our experiments. In Figure 4 the number of visits is shown based on the hours the visitor visited the websites. We can see that the number of visits is the highest at 20:00 – 21:00. The number of visits is the lowest at 10:00 – 11:00.



**Fig 4.** Number of visits by hour

The second report analyzes which are the most visited pages. It is possible to make many other reports depending on the user's needs. For example, the most used browser, most used operating system, detecting suspicious behavior, etc.

In Figure 5 the distribution of the visits is displayed with regard to the most visited 10 sites.

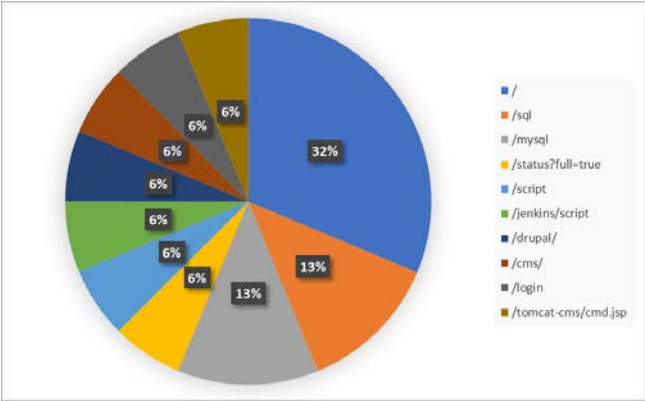


Fig 5. Top 10 most visited sites

**DISCUSSION**

Log data grows very quickly as they are generated at a record rate. Processing power, storage, and cost are some of the biggest challenges that come with this Big Data. Using conventional methods to deal with such data is too costly in terms of both time and money. Hadoop provides frameworks

for distributed storage and processing for Big Data, which makes it possible to use a network for processing and analyzing big amounts of data. In this way, companies save time and money when trying to diagnose problems, solve issues, or obtain the knowledge they wouldn't be able to obtain otherwise. The ability to store as well as distribute large datasets across numerous servers in a cost-effective way makes Hadoop very advantageous with regard to the traditional relational database management systems. Moreover, Hadoop MapReduce processes terabytes of data in minutes and by executing tasks in parallel shortens the processing of data by a considerable amount.

## **CONCLUSION**

Log analysis is important for businesses and system administrators in many aspects, as they give information about malfunctions, defects, security-related issues, and so on. On the other hand, being generated constantly provides a good example of Big Data, therefore a good example of our experiments. Our analysis shows that analyzing logs with Hadoop MapReduce makes the detection of malfunctions, defects, and so on faster and simpler. The results also show that there is not a big difference if we run the code in a real cluster or in the test environment.

## REFERENCES

- [1] Sethy, R. et al. Big Data Analysis using Hadoop: A Survey. International Journal of Advanced Research in Computer Science and Software Engineering 5(7), 2015, pp. 1153-1157.
- [2] Schneider, R.D. Hadoop For Dummies, Special Edition. John Wiley & Sons Canada, Ltd. 2012.
- [3] Borthakur, D. HDFS architecture. Document on Hadoop Wiki. <http://hadoop.apache.org/common/docs/r0.20.2010/>.
- [4] Vavilapalli, V. K.; et al. Apache hadoop yarn: Yet another resource negotiator. Proceedings of the 4th annual Symposium on Cloud Computing. ACM, 2013.
- [5] Hadoop, Apache. Hadoop Archives Guide. The Apache Software Foundation, <http://hadoop.apache.org/docs/current/hadoop-yarn/hadoop-yarn-site/YARN.html> (2019). Retrieved Oct. 15, 2019.
- [6] Kaur, I. et al. Research Paper on Big Data and Hadoop. IJCST, 7(4), 2016, pp. 50-53.
- [7] Dean, J. and Ghemawat, S. MapReduce: Simplified data processing on large clusters. Proceedings of Operating Systems Design and Implementation, 2004.
- [8] Yang, H. et al. Map-reduce-merge: simplified relational data processing on large clusters. Proceedings of the ACM SIGMOD international conference on Management of data. ACM, 2007.
- [9] Rohloff, K. and Schantz, R.E. High-performance, massively scalable distributed systems using the MapReduce software framework: the SHARD triple-store. Programming support innovations for emerging distributed applications. ACM, 2010.
- [10] Point, Tutorials. Retrieved Oct. 15, 2019 from Internet Site <https://www.tutorialspoint.com.html>. Tutorials Point.
- [11] Miner, D. and Radtka, Z. Hadoop with Python. O'Riley Media. 2016.
- [12] Log analysis [https://en.wikipedia.org/wiki/Log\\_analysis](https://en.wikipedia.org/wiki/Log_analysis)
- [13] Sayalee Narkhede and Tripti Baraskar - Hmr Log Analyzer: Analyze Web Application Logs Over Hadoop MapReduce International Journal of UbiComp (IJU), Vol.4, No.3, July 2013.

[14] “Python Fake Logs.” Internet:  
https://github.com/s4tori/fake“Python Fake Logs.” Internet:  
https://github.com/s4tori/fake-logs.-logs.



# Edge Computing Security with an IoT device<sup>1</sup>

*Beyda Nur Kars*

---

- ❖ **Affiliation:** Dokuz Eylul University Department of Computer Engineering  
İzmir, TURKEY
- ❖ **Email:** beyda.kars@ceng.deu.edu.tr

---

<sup>1</sup> This study has been published in the Journal of Emerging Computer Technologies.

## INTRODUCTION

Edge Computing is an IT system designed to bring applications and computing capability as close as possible to the users or "objects" who need them. Edge Computing; mobile computing is driven by the reduced cost of computer components and the number of networked devices on the Internet of Things (IoT). Depending on the application, time-sensitive data in an Edge Computing can be processed by a smart device or sent to a medium server in a location close to the receiver. Edge computing is the optimization of the application's data with end-to-end encryption. That is, cloud computing turns it into a more sophisticated computing cloud architecture. Edge Computing is expected to play a serious role in the near future, especially in terms of information technologies. Edge Computing is referred to as end calculation in English. As it is known, the applications we use on smartphones have to deliver their existing data to the cloud, which is quite far away. Likewise, data is expected to return from this long journey. Edge Computing prevents this waste of time. Edge Computing, which can carry its computing power even to the extreme parts of the networks, brings serious relief in this sense. it cannot replace cloud computing or directly replace it. Edge Computing aims to process data from the ends to the cloud.

Despite all these positive aspects, Edge Computing can pose certain risks in terms of security. The fact that the architecture is distributed brings an increase in the number of vectors. Some vulnerabilities may also arise in the system. When more and more intelligence that an end customer has become more vulnerable, malware becomes infiltrations and vulnerabilities. On the other hand, licensing can force users to some extent, especially in terms of cost. Because each additional function is licensed separately, which naturally leads to higher costs.

Edge Computing brings advantages beyond classical architecture, as it optimizes resource usage significantly. First of all, the processors used in the related devices provide a relatively low power requirement and provide more efficient hardware security.

Mahadev Satyanarayanan, who is described as the father of Edge computing, explained how edge computing was born in his article [1]. According to Satyanarayanan, the limited resources of mobile devices have created a need for a platform that will perform the calculation process instead of these devices and have a more powerful resource. Cloud Computing, which

promises much stronger resources (computing, memory and storage resources) compared to mobile devices, emerged to solve this problem. However, as latency-sensitive and bandwidth-hungry applications developed, gathering all computing power in the remote cloud environment started to cause problems. The main reason behind this problem is that the application packages have to pass through many routers controlled by many Internet Service Providers (ISPs) at different layers until they reach the cloud environment. Each router through which packets pass manifests itself as an increase in the Round Trip Time (RTT) of the packets of delay sensitive applications. In addition, the end-to-end delay values of the routes provided by routing protocols for packets can change dynamically due to ISPs and network problems and situations. In addition to all these, it is predicted that billions of IoT devices will connect to the internet and transfer data in the future. The fact that billions of devices send data to the cloud environment located at a single point in a remote location will inevitably lead to bandwidth bottlenecks. For all these reasons, it is clear that remote cloud environments cannot be a solution for latency sensitive and bandwidth hungry applications. Edge computing was born to solve this problem. Edge computing aims to prevent applications from being affected by long RTT times and bandwidth bottlenecks by moving cloud resources to the endpoints of networks.

## **RELATED WORKS**

Confidentiality should be able to ensure that the data is only available to authorized users during the process and that it is not interfered by unauthorized people. Privacy is the most important security item in IoT because many devices can be integrated into IoT. It should be ensured that data received with a measuring device does not provide secure information to neighbouring devices. To ensure this privacy, advanced techniques and others, including key management mechanisms, should be developed and used [2].

It is essential to use wireless data transmission and to encrypt the information transferred between the nodes in order to keep it confidential. The most appropriate encryption algorithms and adequate key management systems are required to secure this data [3].

Wireless sensor networks have a large number of trust-based intrusion detection systems (IDS) that are used to defend against attacks. However, the effectiveness of IDS decreases in IoT due to the large amount of data produced in a short time. Meng et al. [4] proposed a Bayesian-based trust management

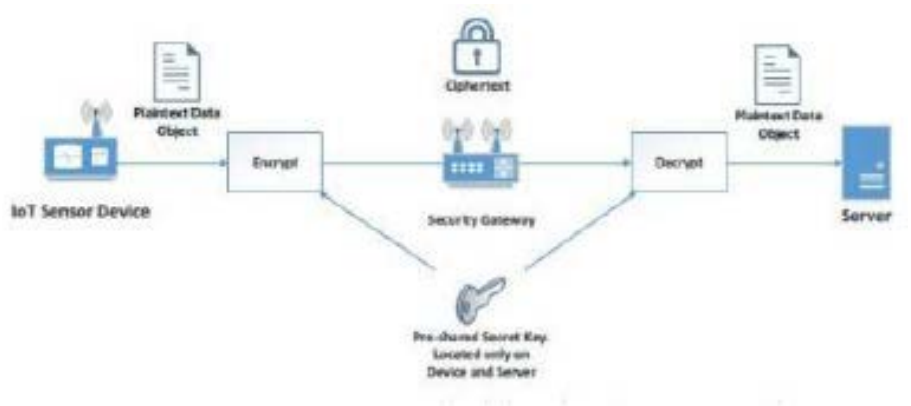
method that incorporates traffic sampling into IDS under a hierarchical structure [5].

James King of Lulea University created an IoT network in his work in 2015 within the local network. Arduino used it as a gateway and collected and encrypted data from different devices and sensors. The encryption algorithm used is the Advanced Encryption Standard (AES) with both 128-bit and 256-bit key length [6]. Mahmudur Rahman, Bogdan Carbunar and Umut Topkara from Florida International University collects instant values from devices such as stopwatch, heart rhythm tracker and moisture meter used in fitness studies and transmits the IoT device to a web server remotely. In this study, FirstBeat is used as an encryption algorithm and Arduino Uno device is used as an IoT device. It also belongs to the IoT device from the webserver by taking the coordinate data, we can track the point we are at via GPS [6].

There is a need for special security mechanisms produced for low-resource devices such as IoT and wireless sensor network devices. It is necessary to use random number generators [7-11], authentication protocols [12-14] and some encryption algorithms that will be built-in in some of these devices.

**PROPOSED WORK**

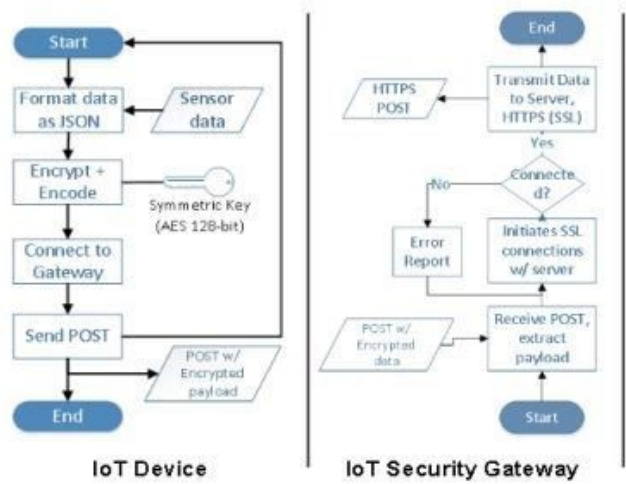
In this section, we will briefly mention the target in the study. We will talk about the hardware materials that will be uses later and the flow of my work.



**Figure 1.** Encrypted data transfer

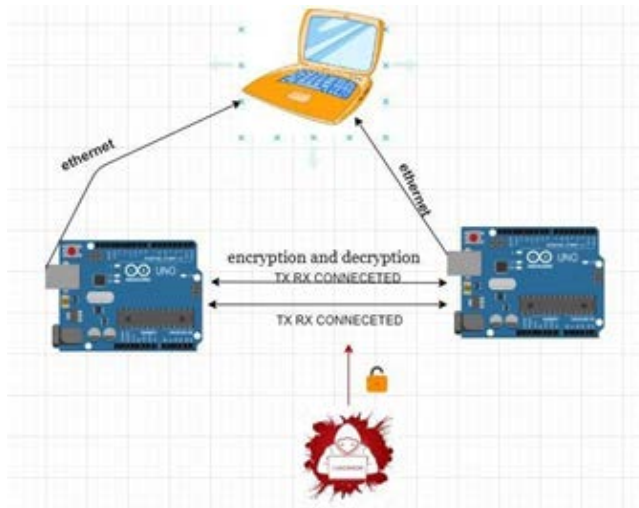
Our goal in this study is to encrypt data sent from IoT devices. This operation is shown in Figure 1. We used two Arduino devices in this study. The first one was used as a gateway device, the second one was used to get sensor data. DHT11 is used as the sensor to get data from the physical environment.

The data received from the sensor will be encrypted with AES 128 and sent to the network gateway. Thus, the data sent by IoT devices will be secured. Arduinos are also communicated among themselves. A wired connection has been made.



**Figure 2.** Workflow

In Figure 2, the workflow of operations of the two Arduino IoT device can be seen.



**Figure 3.** IOT Security Diagram

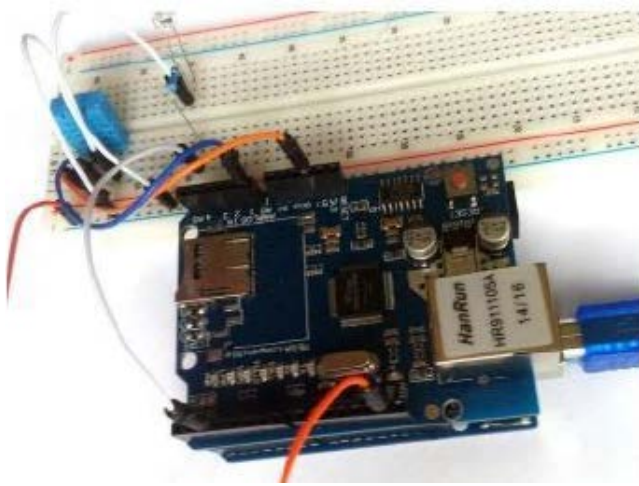
Figure 3 shows the relationship between all the necessary material in the workflow. Sensor data from Arduino 1 is sent to the laptop and encrypted there. And again from Arduino 2 to Arduino 1 (i.e. gateway is sent). Then Arduino 1 is decrypted.

The purpose of encrypting the data is to protect the data from malicious software while it is transmitted from one device to another.

## **MATERIALS**

As seen in the figures, we have two modules. The first one is the module with Arduino and sensor connection. The second one is Arduino, which will be used as gateway.

Figure 4 also has an Arduino module that allows us to receive data from the sensor.



**Figure 4.** Arduino and Dht11 Connection

## **TEST RESULTS**

Before talking about the test results in this part of the project, we will talk about the AES algorithm. We will give some information about the use of the AES algorithm in the working mechanism of this study.

The AES algorithm is a block cypher algorithm that encrypts 128-bit data blocks with 3 different key options, 128, 192 or 256-bit keys. The difference in key lengths makes the number of AES tour cycles differ. Since it is 128 bits, 10 rounds of processing are required. Since its length is 192 bits, 12 tours of processing is required. Because it is 256 bits, 14 rounds of operation are required [15].

Each round contains 4 different steps. These are as follows: byte replacement, row shift, column shuffling, and addition with a rotary switch. The data entered after 10 rounds is obtained as encrypted. The AES algorithm takes the key and passes it through certain processes, creating as many keys as the number of transactions. This number is 10 for 128-bit length. 10 different keys are created and the last key formed becomes the first key used to decrypt [16].

AES is a symmetric encryption algorithm, so the same key is used for both encryption and decryption. While deciphering the encrypted text, the operations performed while creating the encrypted text are applied in reverse.

The same key was used in encryption and decryption in the project as in Figure 5 and Figure 6.

```
uint8_t key[] = {0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15};  
aes128_enc_single(key, data);
```

**Figure 5.** Code for the usage of the encryption key

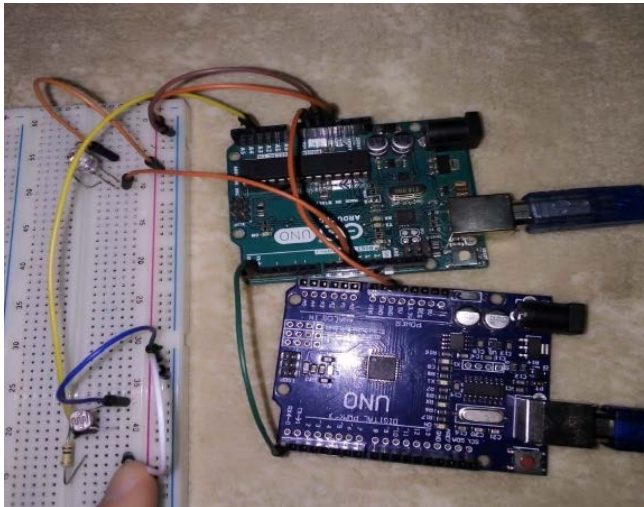
```
uint8_t key[] = {0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15};  
Serial.print("decrypted:");  
aes128_dec_single(key, data);
```

**Figure 6.** Code for the usage of the decryption key

The Arduino's own library was used to encrypt the sensor data. The received data turns into encrypted text with the key we set. But since the sensor data that is received is a numerical value, the type conversion was made.

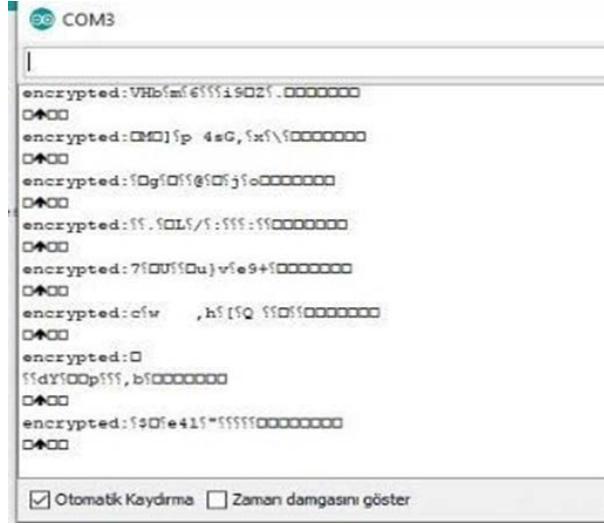
In the test part, we first load our code into our sensor module and continue to supply energy elsewhere by separating the module from the computer. Then we load the code into the module we will use as a gateway and provide the connection between the two Arduino devices. We ensure that encrypted data is sent to the gateway.

In Figure 7, circuits have all connections were shown.



**Figure 7.** Hardware





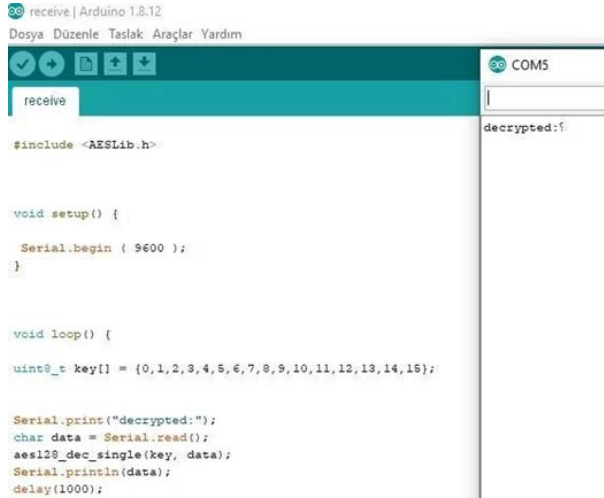
**Figure 9.** Receive and Decrypted

```

#include <AESLib.h>
int ledPin = 10;
int LDRPin = A3;
void setup() {
  pinMode(ledPin, OUTPUT);
  Serial.begin ( 9600 );
}
void loop() {
  int isik_degeri = analogRead(LDRPin);
  int dat = isik_degeri; //16 chars == 16 bytes
  delay(1000);
  char data[4];
  String str;
  str=String(dat);
  str.toCharArray(data,4);
  uint8_t key[] = {0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15};
  aes128_enc_single(key, data);
  Serial.print("encrypted:");
  Serial.print(data);
  Serial.println();
  delay(1000);
}

```

**Figure 10.** Arduino code for encryption



**Figure 11.** Receive and Decrypted

## CONCLUSION

In this study, an edge computing system was established with two Arduino devices. One of the devices was used as a gateway. The other is used to collect the data obtained from the sensor. The collected data is encrypted with AES encryption algorithm to ensure secure communication of data. Encryption methods with 128-bit and 256-bit key lengths have been tried for the encryption process. The AES library on Arduino was used for encryption. With this study, the data was sent encrypted and the data was reconstructed by decrypting the other side. Different encryption methods and different security mechanisms may be tried in the future..

## REFERENCES

- [1] Satyanarayanan, M. (2017). Edge Computing. *IEEE Computer*, 50(10), 36-38.
- [2] Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A survey on internet of things: Architecture, enabling technologies, security and privacy, and applications. *IEEE Internet of Things Journal*, 4(5), 1125-1142.
- [3] Bull, P., Austin, R., Popov, E., Sharma, M., & Watson, R. (2016, August). Flow based security for IoT devices using an SDN gateway. In *2016 IEEE 4th International Conference on Future Internet of Things and Cloud (FiCloud)* (pp. 157-163). IEEE.
- [4] Meng, W., Li, W., Su, C., Zhou, J., & Lu, R. (2017). Enhancing trust management for wireless intrusion detection via traffic sampling in the era of big data. *Ieee Access*, 6, 7234-7243.
- [5] Wang, T., Zhang, G., Liu, A., Bhuiyan, M. Z. A., & Jin, Q. (2018). A secure IoT service architecture with an efficient balance dynamics based on cloud and edge computing. *IEEE Internet of Things Journal*, 6(3), 4831-4843.
- [6] Akkuş, S. (2016). Gömülü sistem tabanlı, kriptolu TCP/IP veri haberleşmesi uygulaması.
- [7] Cabuk, U. C., Aydın, Ö., & Dalkılıç, G. (2017). A random number generator for lightweight authentication protocols: xorshiftR+. *Turkish Journal of Electrical Engineering & Computer Sciences*, 25(6), 4818-4828.
- [8] Kösemen, C., Dalkılıç, G., & Aydın, Ö. (2018). Genetic programming-based pseudorandom number generator for wireless identification and sensing platform. *Turkish Journal of Electrical Engineering & Computer Sciences*, 26(5), 2500-2511.
- [9] O'Neill ME. PCG: A family of simple fast space-efficient statistically good algorithms for random number generation. Available online at <http://www.pcg-random.org/pdf/toms-oneill-pcg-family-v1.02.pdf>
- [10] Aydın, Ö., & Dalkılıç, G. (2018, July). A hybrid random number generator for lightweight cryptosystems: xorshiftLplus. *The 3rd*

International Conference on Engineering Technology and Applied Sciences (ICETAS).

- [11] Aydın, Ö., & Kösemen, C. (2020). XorshiftUL+: A novel hybrid random number generator for internet of things and wireless sensor network applications. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 26(5), 953-958.
- [12] Armknecht F, Hamann M, Mikhalev V. Lightweight authentication protocols on ultra-constrained RFIDs-myths and facts. In: International Workshop on Radio Frequency Identification: Security and Privacy Issues; 2015. pp.1-18. Springer, Cham.
- [13] Lee JY, Lin WC, Huang YH. A lightweight authentication protocol for internet of things. In: 2014 International Symposium on Next-Generation Electronics (ISNE); New York, USA; 2014. pp. 1-2.
- [14] Aydın, Ö., Dalkılıç, G., & Kösemen, C. (2020). A novel grouping proof authentication protocol for lightweight devices: GPAPXR+. Turkish Journal of Electrical Engineering & Computer Sciences, 28(5), 3036-3051.
- [15] Şeker, Ş.E. (2009). AES ve Rijndael Şifreleme. *Bilgisayar Kavramları*. Retrieved from <http://bilgisayarkavramlari.com/2009/06/03/aes-ve-rijndael-sifreleme/>, Accessed date: 25.11.2020
- [16] Surian, D. (2006). Algoritma Kriptografi AES Rijndael. TESLA Jurnal Teknik Elektro UNTAR, 8(2), 97-101.



# Secure Connection between Google Home and IoT Device<sup>1</sup>

*Ekrem YİĞİT*

---

- ❖ **Affiliation:** Dokuz Eylul University, Faculty of Engineering, Department of Computer Engineering İzmir, Turkey
- ❖ **Email:** ekrem.yigit@ceng.deu.edu.tr

---

<sup>1</sup> This study has been published in the Journal of Emerging Computer Technologies.

## **INTRODUCTION**

Google Home is a device developed by Google that can be customized both with its own functions and with many devices connected via Wi-Fi. It offers an interface that is managed by voice commands and can be customized.

Approximately 6 weeks after the release of Google Home devices, the logs kept on the devices were checked and it was determined that the device recorded the sounds as a log hacker can use these data to penetration attempts or fishing attacks. Researchers have found that smart speakers can be hacked with the help of laser-powered "light commands." Researchers suggest smart speaker makers can fix this vulnerability by adding a light shield around the microphone or using two different microphones on opposite sides to listen to voice commands.

The project is to operate the IoT device with safe commands sent from google home. In the project, the main target is to connect to the cloud from the google home device via IFTTT, then to get the light effect by taking data from the cloud using the NodemCU IoT device and Blynk.

NodemCU is a circuit board with high quality ESP8266 Wi-Fi module that can be programmed from Arduino IDE application and can communicate easily. It is aimed to provide cloud interaction by communicating with this device over the internet by making web transactions thanks to HTTP libraries. IFTTT (if this then that), If this happens, do it. It is an abbreviation of one of the simplest code phrases. This assign device to be triggered and the that assign action that works on device that is triggered. Webhooks service is a IFTTT service that connect to Blynk and IFTTT on Cloud Service. Webhooks is used to send requests to the URL we enter. Blynk server information and the TOKEN information used are entered as URLs and connection is provided and control is provided via IFTTT. In this URL, the pins on the device are entered and the controls of various pins are provided. Finally, my project is to identify vulnerabilities in communication on Blynk and IFTTT connections and to use these two components effectively in the safest possible way.

## **RELATED WORKS**

The interaction of end-users with online IoT devices has been increasing for the last 10 years and with the increasing interactions, the usage of these IoT devices has started to be observed beyond the intended use. IoT

devices need lightweight and new security mechanisms like random number generators [1-3], authentication protocols [4,5] and privacy protected connection methods. IFTTT enables users to use their home assistant devices heterogeneously on multiple platforms with IoT via the internet connection tool. With the widespread use of this environment, users are satisfied with the increase in interaction and they started to load more work on their systems.

In order to meet the increasing demand for products, studies on the field of security in a competitive environment have been quite incomplete. According to the researches carried out today, more than 50% of the 19323-connection established over home assistant devices did not find any security factor.

While the picture taken from the phone is in a private structure, the smart home device translates this picture information into a Public form. In the scenario described, the hacker sees the system vulnerabilities related to IFTTT, which enables the connection of the smart home device to the phone, and the Home Assistant will also make private content public and make confidential data accessible [6].

Security problems related to 45 smart home appliances and IoT devices connecting with them are discussed. It is determined that the most common vulnerability is users. Users endanger their own security by using Home Assistants with different levels of security for the same or similar jobs or IoT devices attached to them. In this research [7], which presents this vulnerability as the simplest and most effective solution, the use of multiple IoT and Home Assistant was shaped and the relational relationships were explained to the users.

It is not very difficult to steal users' IFTTT accounts with various online attack methods. Attackers who will change the authorizations made through this IFTTT account on another device for their own malicious purposes can cause damage as much as cyber-attacks with increasing device interactions [6].

This framework [8], developed to prevent a bad interaction via IFTTT, inserts the MAC address of the device to which the user connects to the IoT device to a PUF at the beginning, and returns the output it receives as OTP, allowing access only from the devices authorized by the user at the beginning.

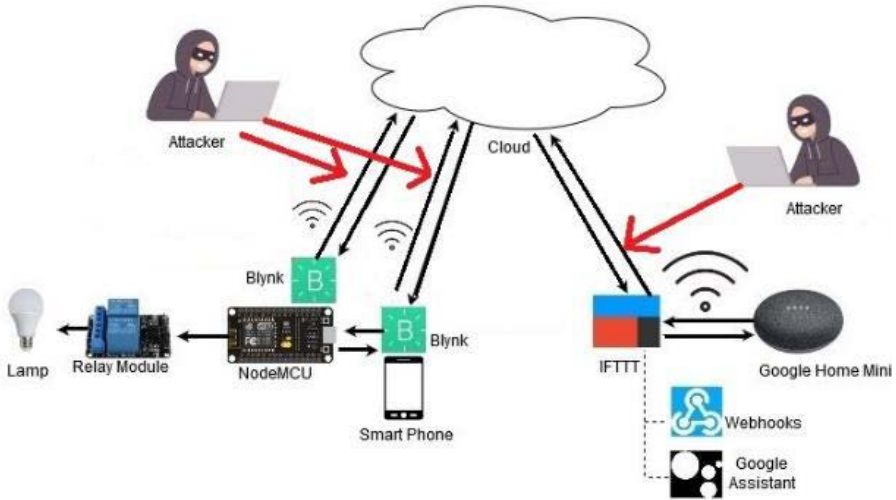
The most important prioritization on the market is always the cost of IoT devices and the platforms actually used. In costing-priority product planning, security is placed in the background. Extra cost or does not require any hardware. The answer that best meets these needs is the AES algorithm, which is also described in another research [9]. Offering both fast and secure service at once, AES also does not cause negative effects on IoT device performance [10].

Due to the lack of encryption in the transmission of data from IFTTT to IoT device of systems that are tried to be created with IFTTT account security and device connection methods seen in various studies, there are also vulnerabilities in leaking many data by logging them, so AES encrypt as an effective and cost effective method. With the method, both the data coming from IFTTT and Google Home Assistant is encrypted, and then the data is decoded without coming to the IoT device and secure transmission of the data [9].

Çepik et. al presented a study in 2020 about security vulnerabilities on the connection between Google Home and IoT device. They tested if this communication is vulnerable to an NTP attack [11].

## **PROPOSED WORK**

The 2 basic devices of the project are the Google Home Mini and NumedCU IoT device. The microphone section on the device must be active. Make AES decryption on NumedCU, AES library has been added to Arduino IDE. The necessary encryption functions can be realized with this library. Password and SSID are necessity to connect NumedCU via Blynk. SSID and password If the information is added directly into the post to provide a link, a security flaw occurs here. While connecting, this two information in the software pose security risks. The attacker can discover certain data on NodemCU by reverse engineering methods via EPPROM.



**Figure 1.** Representation of Possible System Vulnerabilities

All data entered while communicating with Blynk in NodemCU is kept in EEPROM. It is seen in the research that it is possible to access the data in EEPROM physically via a USB [12]. The attacker, who accesses the data via USB, can easily access the internet information on which the device is used, since no hardware measures are taken.

In another study to prevent EEPROM access in terms of hardware, by adding extra parts, the hardware vulnerability of the attackers was prevented but the software vulnerabilities could not be blocked [13].

The attackers' access to the information of the internet connection is a danger to all devices on that connection. To eliminate these risks, SSID and password will be added by encrypting with 128bit AES key in this study and this encrypted data will be provided by decryption while the connection is provided. 128Bit 16ASCII characters will be used for the key.

Reverse engineering of Arduino memory has been handled in the study in which it is possible to access information in the software [14]. The key size of the AES algorithm has a direct effect on performance in various algorithm encounters. When high security is targeted, performance/security threshold has been seen in the examinations; AES is more successful in this rate [15]. After a secure connection is established, the command to turn on the bulb coming from the cloud will be realized.

## TEST RESULTS

The main goal of the project is to add the user-known internet information to the program in a previously encrypted form and to provide an encrypted internet connection. Communication with four devices and a common internet connection has been tested.

```
EkremEv
Encrypted: 2nqyFb+a9uGmJw52SAZtSg==
Decrypted: EkremEv
EkremEv
Encrypted: feRuUTz/lmkPbVaUh0i3Fg==
Decrypted: EkremEv
EkremEv
Encrypted: I6VVDN72z+Erlbtv2jvAyA==
Decrypted: EkremEv
EkremEv
Encrypted: A0NCgvzNjqusq9YR+NddSA==
Decrypted: EkremEv
EkremEv
Encrypted: C+AEieiqt2bgLmuZc6ECvg==
Decrypted: EkremEv
EkremEv
Encrypted: k8x53C+MKWj5Fk2PQpVUQA==
Decrypted: EkremEv
EkremEv
Encrypted: Nku4H7Q6t4aSgNfw79ODkA==
Decrypted: EkremEv
EkremEv
Encrypted: 6HZEKJk4D/zL3uUUfTiBeg==
Decrypted: EkremEv
EkremEv
Encrypted: 7YtI/cP7RrHtxr3SZv26Dw==
Decrypted: EkremEv
EkremEv
Encrypted: 2oPq27mZI6tXlNS+NBMakQ==
```

**Figure 2.** Arduino IDE Serial Port Output

The commands set via IFTTT to the Google home device were given as HTTP request, and the connection with the IoT device was tested and the lamp was successfully controlled. Button on / off communication to Blynk port 0 was transmitted in HTTP request format.

On the security side of the application, Blynk internet connection has been tested to properly encrypt and decrypted data. As seen in the figure below, the software used on the IoT device processes the data properly. When performing the tests, the IV has been re-created while performing each encryption in 16 bytes.

The connection made with the SSID and password of the user, which was previously encrypted, was successfully implemented. These devices, which generally communicate with the cloud, have become a necessity in order to provide sufficient trust. As mentioned in many studies like [16], NodemCU does not have sufficient security, so it is coded into additional security algorithms.

## **CONCLUSION**

As a result, data in the NodemCU memory is not available in a meaningful way. All of important data are kept in encrypted form in memory and communication is provided in encrypted form. Even if the attacker has accessed to NodemCU via USB or software, the attacker will see data encrypted with AES, which he cannot make sense of. This prevents the device from being physically hijacked or the attacker from exploiting vulnerabilities.

.

## REFERENCES

- [1] Cabuk, U. C., Aydın, Ö., & Dalkılıç, G. (2017). A random number generator for lightweight authentication protocols: xorshiftR+. Turkish Journal of Electrical Engineering & Computer Sciences, 25(6), 4818-4828.
- [2] Aydın, Ö., & Dalkılıç, G. (2018, July). A hybrid random number generator for lightweight cryptosystems: xorshiftLplus. The 3rd International Conference on Engineering Technology and Applied Sciences (ICETAS).
- [3] Aydın, Ö., & Kösemen, C. (2020). XorshiftUL+: A novel hybrid random number generator for internet of things and wireless sensor network applications. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 26(5), 953-958.
- [4] Lee JY, Lin WC, Huang YH. A lightweight authentication protocol for internet of things. In: 2014 International Symposium on Next-Generation Electronics (ISNE); New York, USA; 2014. pp. 1-2.
- [5] Aydın, Ö., Dalkılıç, G., & Kösemen, C. (2020). A novel grouping proof authentication protocol for lightweight devices: GPAPXR+. Turkish Journal of Electrical Engineering & Computer Sciences, 28(5), 3036-3051.
- [6] Milijana Surbatovich, Jassim Aljuraidan, Lujo Bauer, Some Recipes Can Do More Than Spoil Your Appetite: Analyzing the Security and Privacy Risks of IFTTT Recipes” WWW '17: Proceedings of the 26th International Conference on World Wide Web April 2017
- [7] Omar Alrawi, Chaz Lever, Manos Antonakakis, Fabian Monrose, Security Evaluation of Home-Based IoT Deployments, 2019 IEEE Symposium on Security and Privacy, 19-23 May 2019
- [8] Jacobs, IS, Bean, CP (1963) “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [9] Bennet Praba1, Shivam Kumar, Ankur Saxena, Sourabh Patel “Securing Assistant Based Home Automation using AES Algorithm” Department of Computer Science and Engineering SRM Institute of Science and Technology, Chennai, India, 2019
- [10] Ross Mcpherson James Irvine, “Using Smartphones to Enable Low-Cost Secure Consumer IoT Devices,” Department of Electronic and Electrical Engineering, Royal College Building, Glasgow G1 1XW, U.K..
- [11] Çepik, H., Aydın, Ö., Dalkılıç, G. (2020) Security Vulnerability Assessment of Google Home Connection with an Internet of Things

- Device. In:7th International Management Information Systems Conference, İzmir, Turkey (9-11 December 2020).
- [12] Kingpin, "Attacks on and Countermeasures for USB Hardware Token Devices," Proceedings of the Fifth Nordic Workshop on Secure IT Systems, 2000
- [13] Grand, Joe. "Practical secure hardware design for embedded systems." Proceedings of the 2004 Embedded Systems Conference, San Francisco, California. 2004.
- [14] Torroja, Yago, et al. "A serial port based debugging tool to improve learning with arduino." 2015 Conference on Design of Circuits and Integrated Systems (DCIS). IEEE, 2015.
- [15] Kumar, M. Anand, and S. Karthikeyan. "Investigating the efficiency of Blowfish and Rejindael (AES) algorithms." International Journal of Computer Network and Information Security 4.2 (2012): 22.
- [16] Utpala, Kuchi NSSSS, N. Suresh Kumar, K. Praneetha, D. Hema Sruthi, and K. Sai Avinash Varma. "Authenticated IoT Based Online Smart Parking System with Cloud." Pramana Research Journal 9, no. 4 (2019).



# Practical Evaluation on Serious Games in Education<sup>1</sup>

*Slavica Mileva Eftimova*

---

- ❖ **Affiliation:** High School SOU Jane Sandanski Strumica, North Macedonia
- ❖ **Email:** meslavica@yahoo.com

*Ana Madevska Bogdanova*

---

- ❖ **Affiliation:** Ss. Cyril and Methodius University, Faculty of Computer Science and Engineering, Skopje, North Macedonia
- ❖ **Email:** ana.madevska.bogdanova@finki.ukim.mk

*Vladimir Trajkovik*

---

- ❖ **Affiliation:** Ss. Cyril and Methodius University, Faculty of Computer Science and Engineering, Skopje, North Macedonia
- ❖ **Email:** trvlado@finki.ukim.mk

---

<sup>1</sup> This study has been presented at the 16th International Conference on Informatics and Information Technology (CIIT 2019) and published in the Journal of Emerging Computer Technologies.

## **INTRODUCTION**

Today, the teacher abandons the role of a central figure, becomes a leader who guides the students through the learning process, enabling more learning styles, so that the student can move forward with his\her learning pace [1]. Students, on their part, use technology for communication, searching and finding information, expanding social experiences, and enjoying computer games on a daily basis [2]. The digital games (on their phones or computers) are played by students with a lot of energy and enthusiasm. This commitment is a challenge for the teacher - the learning process can be done through computer games. Students have transferable skills to share online research and access to many digital texts in a number of contexts. If digital literacy is encouraged throughout the teaching program, using positive language is necessary. By making such changes in the language teachers use, with the goal of becoming closer to their students' language, the students themselves will feel closer to the teachers and will understand the learning material more easily [3]. Technology offers a wide range of opportunities for developing learning experiences across a wide range of topics. If digital literacy is promoted through the curriculum, a positive language is required. By changing the language that teachers use, in order to be closer to the students, then the students themselves will feel closer with the teachers and will easily overcome the material.

## **RELATED WORKS**

Confidentiality should be able to ensure that the data is only available to authorized users during the process and that it is not interfered by unauthorized people. Privacy is the most important security item in IoT because many devices can be integrated into IoT. It should be ensured that data received with a measuring device does not provide secure information to neighbouring devices. To ensure this privacy, advanced techniques and others, including key management mechanisms, should be developed and used [2].

It is essential to use wireless data transmission and to encrypt the information transferred between the nodes in order to keep it confidential. The most appropriate encryption algorithms and adequate key management systems are required to secure this data [3].

Wireless sensor networks have a large number of trust-based intrusion detection systems (IDS) that are used to defend against attacks. However, the

effectiveness of IDS decreases in IoT due to the large amount of data produced in a short time. Meng et al. [4] proposed a Bayesian-based trust management method that incorporates traffic sampling into IDS under a hierarchical structure [5].

James King of Lulea University created an IoT network in his work in 2015 within the local network. Arduino used it as a gateway and collected and encrypted data from different devices and sensors. The encryption algorithm used is the Advanced Encryption Standard (AES) with both 128-bit and 256-bit key length [6]. Mahmudur Rahman, Bogdan Carbunar and Umut Topkara from Florida International University collects instant values from devices such as stopwatch, heart rhythm tracker and moisture meter used in fitness studies and transmits the IoT device to a web server remotely. In this study, FirstBeat is used as an encryption algorithm and Arduino Uno device is used as an IoT device. It also belongs to the IoT device from the webserver by taking the coordinate data, we can track the point we are at via GPS [6].

## **BACKGROUND**

With carefully selected concepts and accompanying pictures that create a pleasant and creative atmosphere, children learn more easily through various activities of games and tasks. The games often have a fantastic element that intrigue players and engage them in learning activities [4,5,6].

But in order to apply games in teaching, more conditions need to be met. According to the UNESCO framework, teachers should use educational (serious computer games) in education, preferably in accordance with the application of the teaching experience [3]. Teachers should have the following competencies:

- Using the Internet for online research;
- Using tools for making text and spreadsheets, making presentations;
- Using communication and collaboration tools such as emails, video conferencing and social networks;
- Application of ICT sources for curriculum development;
- Interest in continuing upgrading and improving the teaching content they teach and their teaching skills;

- Knowledge of the subject they teach to be appropriate for the age of the students;
- Have managerial and organizational skills;
- Knowledge of strategies that will help the student to gain in-depth knowledge such as:
  - o Learning Collaboration;
  - o Problem-based learning;
  - o Project-based learning;
  - o Activities based on project development;
  - o Games and simulations;
  - o Research experiments;
  - o Case study;
  - o Exercises;
  - o Mentoring;
  - o Evaluation.

What conditions should the school have in order to introduce educational games in teaching? The answer to this question divided in two parts:

- Provided technical equipment in the classroom and
- Teachers have to have appropriate digital competencies.

The word "competence" means knowledge or expertise in a given area [7]. Accordingly, digital competence is the ability to track, analyse, evaluate, generate and transmit information in digital format. This applies to desktops, laptops, smartphones and similar devices. Regarding the discussion of this term, there are various attempts to define definitions that are in use, as well as a few related names for it, such as information, the Internet or media competence. A person who is digitally competent will have more interlinked skills: knowledge of the basic principles of computer hardware, computer networking skills, the ability to engage in online communities and social networks. By digitizing human knowledge and developing digital technologies (mass production of devices that have access to the Internet). We can conclude that a digitally literate person will have practical knowledge of hardware and software, but also different kind of knowledge that they did not have in the last century. Computer literacy is often considered today as the ability to use the computer programs for some less complex practical tasks or

the ability of individuals to effectively use the computer. Digital competence, or in other words, digital literacy, is considered to be as important today as reading and writing. Digital devices are starting to be used from an increasingly young age, but this does not help much young people to develop the skills they need for further personal and professional upgrading. Digital competence is far more than just accepting new technologies and using social media in order to create some content.

Today, digital literacy is almost equally necessary to attain personal and professional ambitions. It allows seemingly complex tasks to be performed in a much simpler and more efficient way and with better results. It is necessary to focus attention on the way the students use devices in extracting knowledge.

Digitalization and inroads have already led to major changes in our daily lives and our world in terms of information and work. However, these numerous changes have not yet become clearly visible and understandable to us. For this reason, it is particularly important to pay attention to large volumes of new information and to all innovations. To do this, it is necessary to have a sufficiently high level of digital competence.

As far as the technical equipment of the schools is concerned, as a first requirement is that there is an Internet connection and at least one laboratory with a certain number of computers, preferably connected in a local network. Possession of additional tools and equipment can greatly enrich hours with certain activities. The LEGO Mindstorm EV3 Robot [8] and micro:bit tools [9] are used in practical case studies.

## **THE CASE STUDY**

In order to conduct this evaluation, the evaluation framework for integration of digital games into education (MEDGE) was used [10], expanded with two new questions, thus MEDGE+.

The following questions were answered by several professors at the “Jane Sandanski” High School:

- Is the game easy to use? (EASY)
- What is the educational goal of the game? (VAL)
- Does the game adapt to educational goals? (ADT)
- Pleasure / acceptance of the game by students? (QoE)

- What is the teacher's subjective opinion about the game? (SUB)
- What is the motivation of students to adopt the material? (MOT)

In order to achieve better motivation for the students when applying serious games, two questions were added, where the students were asked the following questions (MEDGE+):

- Through the game, I will easily overcome the material (EL);
- Through the game, my motivation for material adoption (MS) will increase.

The following scale of responses was offered:

I totally agree (5). I agree (4), I am neutral (3), I disagree (2), I completely disagree (1).

## **Case Study 1 - LEGO**

A class in high school using a memory game (with the help of the LEGO Mindstorm EV3 Robot) [8].

Teaching unit: One-dimensional arrays;

*Type of lesson: lecturing;*

In order to get the students interested as much as possible on this thematic unit, a memory game in Python was used through simulation of robotic games. Preparing for this guide: LEGO Mindstorm EV3 Robot is the tool used for this game. A robot with specific parts is built in the instructions for building the LEGO Mindstorm Education Core Set. The goal and the game is to build the main body for the robot (base unit) and the color sensor.

*Effects*

Computer Science (Python) - This lesson will help students understand the use of an array, from an abstract concept to a point where they actually understand how the color storage works in random order. Students will be introduced to the random functions used to generate colors.

*Exercise*

Create a program that will put random colors in sequence, and then the robot will repeat (express) the order of colors. The student must remember the

colors and show colors on the card to the robot sensor in the same order as given. In the end, the robot announces whether the study of time wins or not.

*Reflection*

Students learned how to store colors in arrays and to check if the generated colors are the same as shown in front of the sensor. They can also make different versions of this program by counting a score and saving other type of data. Based on a poll conducted after playing this game, the following results were obtained, given in the Fig1:

|           | EAS<br>Y | VA<br>L | AD<br>T | Qo<br>E | SU<br>B | MO<br>T | E<br>L  | M<br>S  |
|-----------|----------|---------|---------|---------|---------|---------|---------|---------|
| rob<br>ot | 2,5      | 3,4     | 5       | 3,8     | 4,4     | 4,6     | 4,<br>4 | 4,<br>1 |

**Fig. 1.** Review of robot responses, MEDGE+

**Case Study 2 - KAHOOT**

*Teaching unit:* Basics in programming with C++  
*Type of lesson:* Kahoot quiz to check the acquired knowledge on the topic Introduction to programming in C ++.

Kahoot [11] is a formative learning tool that uses quizzing technology, discussions and surveys. The principle of work is basically a game in which the whole class participates in real time. For the preparation of this class, a quiz with 10 questions was developed, which examines the gained initial knowledge in programming in C ++, which is necessary to start programming the simplest tasks in C ++.

*Exercise*

Students get the link and join the game.

*Reflection*

The class with Kahoot was filled with excitement and euphoria like no previous one. The competitive spirit was at the highest level. After the quiz was conducted, questions that were incorrectly answered were discussed. The students asked after each teaching unit (or at least after completing a theme) to have a time dedicated to competing with the Kahoot Quiz. Based on the

conducted survey after playing this game, the following results were obtained, given in the Fig2:

|        | EAS<br>Y | VA<br>L | AD<br>T | Qo<br>E | SU<br>B | MO<br>T | E<br>L | M<br>S |
|--------|----------|---------|---------|---------|---------|---------|--------|--------|
| Kahoot | 4,6      | 5       | 5       | 5       | 4,7     | 5       | 4,8    | 4,7    |

Fig 2. Review of Kahoot responses, MEDGE+

**Case Study 3 - MICRO:BIT**

A class in high school using a micro:bit [9] to verify acquired knowledge of algorithms with a branched structure.

*Teaching unit:*

Algorithms and programming;

*Type of lesson:*

Algorithms and their representation. In order to hold this lesson, the OOU Sando Masev Strumica was visited, where, with the help of the British Council, the school owns 30 micro:bit devices [12]. At this class, the application of the algorithm with a branched structure was presented, realized practically with the micro:bit device.

*Exercise* The acquired experiences were used for introduction into the branch structure. Other approaches were introduced to explain algorithms with a branch structure. The students were able to solve other examples.

*Reflection*

By using the micro:bit, the programming becomes clearer, closer and more attentive to the students. The survey by MEDGE+ has produced the table in Fig 3.

|           | EAS<br>Y | VA<br>L | AD<br>T | Qo<br>E | SU<br>B | MO<br>T | E<br>L | M<br>S |
|-----------|----------|---------|---------|---------|---------|---------|--------|--------|
| micro:bit | 4,3      | 5       | 5       | 4,5     | 4,2     | 5       | 4,8    | 4,5    |

Fig.3. Review of micro: bit responses, MEDGE+

The evaluation methodology MEDGE+ produces the net presented in Fig.4. It gives graphical representation of the game/tool acceptance in the three case studies.

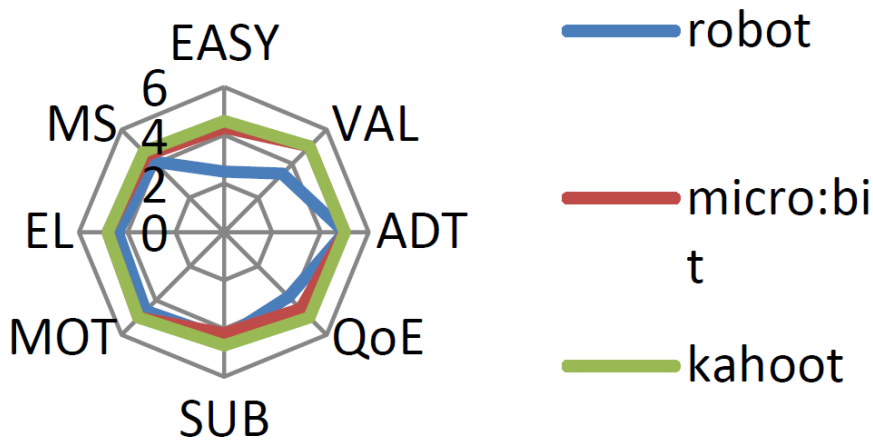


Fig 4. Evaluation of the explored games with MEDGE+

### DISCUSSION

In this part a discussion of the three case-studies is given. At the lesson of one-dimensional strings, the good sides were: greater interest and curiosity towards something new, resulting in increased interest in learning and interest in research through additional sources for the new concepts. The negative sides were that the use of robots in teaching requires additional costs, one robot is not enough for a group of 24-30 students. Another point is the programming language - in order to use the robot in the teaching process, another programming language was needed, and not the one that is studied in the regular classes. The good sides in the realization of classes using the Kahoot quiz for repeating the material for the basics of programming were: initiating a competitive spirit, raising awareness of teamwork, getting quick results for the correct answers and showing greater interest in understanding the reasons of choosing the wrong answers. Also, all students answered the same questions at the same time and there was no fear of consequences if wrong answer was submitted.

In the third case study, the algorithms for a branched structure with micro: bit was introduced. The positive side was this new approach to learning algorithms. It was more interesting to the students because of the visualization and the ease of use. One disadvantage of using micro:bit in high schools is

that following the curriculum, micro:bit can be applied only at the beginning of the programming courses because the latter material is more complex.

## **CONCLUSION**

The introduction of games in IT teaching is a very positive experience [13,14,15], but of course, the realization of each of these serious games and tools has positive and negative aspects.

In this paper, following the evaluation methodology [10], MEDGE+ was introduced that can give even more insight when choosing an appropriate educational game in the teaching process. We have explored three case studies, and used the MEDGE+ methodology to measure the acceptance of the given games and tools.

The performed analysis was done over three different games: robot simulation, Kahoot and micro:bit coding. It showed that all three activities positively influenced the process of adopting new knowledge and knowledge testing, while the greatest satisfaction and motivation according MEDGE+ achieved the quiz Kahoot. The extension of the methodology [10] considered the inclusion of the student's opinion, contributing to the teacher's decision which games/tools should be chosen in the educational process.

The use of games in teaching obviously has many advantages, and as a future work, more examples of games and tools will be evaluated using the information provided by MEDGE+.

## REFERENCES

- [1] L. Phipps, V. Alvarez, S. de Freitas, K. Wong, M. Baker, and J. Pettit, "Conserv-AR: A Virtual and Augmented Reality Mobile Game to Enhance Students' Awareness of Wildlife Conservation in Western Australia", Proceedings of the 15th World Conference on Mobile and Contextual Learning (mLearn 2016), Sydney, Australia, vol. 1, pp. 214-217, 2016
- [2] H.A. Spire, "21st century skills and serious games: Preparing the N generation," in L.A. Annetta, Serious educational games. Rotterdam, The Netherlands: Sense Publishing, 2008
- [3] United Nations Educational, Scientific and Cultural Organization, unesco ict competency framework for teachers, 2011;
- [4] P. Fotaris, T. Mastoras, R. Leinfellner, and Y. Rosunally, "Climbing Up the Leaderboard; An Empirical Study of Applying Gamification Techniques to a Computer Programming Class," Electronic Journal of e-Learning, vol. 14, no. 2, pp. 94-110, 2016.
- [5] T.M. Connolly, E.A. Boyle, E. MacArthur, T. Hainey, and J.M. Boyle, "A systematic literature review of empirical evidence on computer games and serious games," Computers & Education, vol. 59, pp. 661–686, 2012.
- [6] J.C. Burguillo, "Using game theory and Competition-based learning to stimulate student motivation and performance," Computers & Education, vol. 55, no. 2, pp. 566–575, 2010
- [7] Digital competence  
<https://europass.cedefop.europa.eu/resources/digital-competences> , 15.12.2018;
- [8] Lego education, <https://education.lego.com>, 25.02.2019;
- [9] Micro:bit, <https://microbit.org/code/>, 25.02.2019;
- [10] Maja Videnovik , Ana Madevska Bogdanova , Vladimir Trajkovik, "Serious games evaluation methodology", Proceedings of ICERI2018 Conference 12th-14th November 2018, Seville, Spain;
- [11] M. Videnovik, L. Kionig, T. Vold, and V. Trajkovik, "Testing framework for investigating learning outcome from quiz game: A Study From Macedonia and Norway," in 17th International

- Conference on Information Technology Based Higher Education and Training (ITHET), pp. 1-5, IEEE, 2018
- [12] Materials by teacher Biljana Nikolova, OO Sando Masev Strumica, 15.12.2018;
- [13] Gamestorming, <https://gamestorming.com/category/games-for-problem-solving/>, 12.02.2019;
- [14] M. Popescu, S. Arnab, R. Berta, J. Earp, S. De Freitas, M. Romero, I. Stanescu, and M. Usart, "Serious games in formal education: discussing some critical aspects," in Proceedings of 5th European Conference on Games-Based Learning, Athens, Greece, pp. 486–493, 2011
- [15] Immersing, addaptive learning <https://www.ixl.com>, 12.02.2019

# Digital Twin Based Disaster Management System Proposal: DT-DMS <sup>1</sup>

*Özgür Doğan*

---

- ❖ **Affiliation:** Department of Computer Engineering, Muğla Sıtkı Koçman University, Muğla, Turkey
- ❖ **Email:** ozgurdogan3@posta.mu.edu.tr

*Oğuzhan Şahin*

---

- ❖ **Affiliation:** Department of Computer Engineering, Muğla Sıtkı Koçman University, Muğla, Turkey
- ❖ **Email:** oguzhansahin3@posta.mu.edu.tr

*Enis Karaarslan*

---

- ❖ **Affiliation:** Department of Computer Engineering, Muğla Sıtkı Koçman University, Muğla, Turkey
- ❖ **Email:** enis.karaarslan@mu.edu.tr

---

<sup>1</sup> This study has been published in the Journal of Emerging Computer Technologies.

## **INTRODUCTION**

Metropolitan cities are becoming more crowded and their transportation, telecommunication, electricity networks, and natural gas networks are becoming more complex day by day. Natural disasters are a reality and these big cities have to be ready for pre and post-disaster situations. Turkey is no exception.

The number of natural disasters such as earthquakes and their destructiveness are expected to increase in the future. The most recent earthquake in Elazığ showed that disaster management practices and the tools to train staff for pre and post-disaster situation are not sufficient. The fundamental elements of post-disaster interventions were late and communication infrastructures collapsed.

Digital twin technology can be used as a solution. Digital twin can be defined as the bridge between the physical and virtual world. The data received from physical assets are transferred to the digital twin. Digital twin technology can be used to build smarter buildings and infrastructures like transportation, telecommunication, electricity, and natural gas networks. Such a solution should be developed before any crisis to allow emergency staff to be trained to meet the surmountable challenges of a natural disasters, especially an earthquake.

In the next section, fundamentals of digital twin solution are explained. Related works are provided in the third section. The proposal of the system is explained in the fourth section followed by the implementation in section five. The results and conclusions are included in final section.

## **FUNDAMENTALS**

### **Disaster Management**

Today's technology is not yet sufficient in predicting when disasters will take place [1]. However, today's technology is sufficient to supply means which enables us to take measures before and after the disaster to reduce catastrophic effects of any incident on human lives and property. Disaster management can be used for the following:

- Prevent and reduce the impact of the disasters,
- Respond to the events in a timely, fast and effective manner.

- Create a new and safer environment for the victims.

Disaster management can be examined as pre and post disasters. In pre-disaster, the following actions can be taken to reduce risk and harm.

- Determining regions that exist on the fault line,
- Informing and raising awareness of the society about disaster hazard and risk,
- Improving disaster management capacity and capability,
- Installation of IoT devices in the buildings and in the critical infrastructures such as transportation, telecommunication, electricity and natural gas networks.

In a post-disaster situation, timely response by allocating resources to the most affected areas is one of the most critical aspects to reduce human suffering and number of casualties. In this context, gathering current information becomes the most crucial thing. Information/data about the affected areas can be gathered by using the IoT devices that are installed during the pre-disaster period. This data can be transferred to the rescue teams and could be used to allocate scarce resources to the most affected and areas to facilitate efficient relief operations. Drones, satellites, and similar devices could also be used to collect information on conditions of the disaster region. The data that is collected from these devices could be used in making decisions about the establishment of the following:

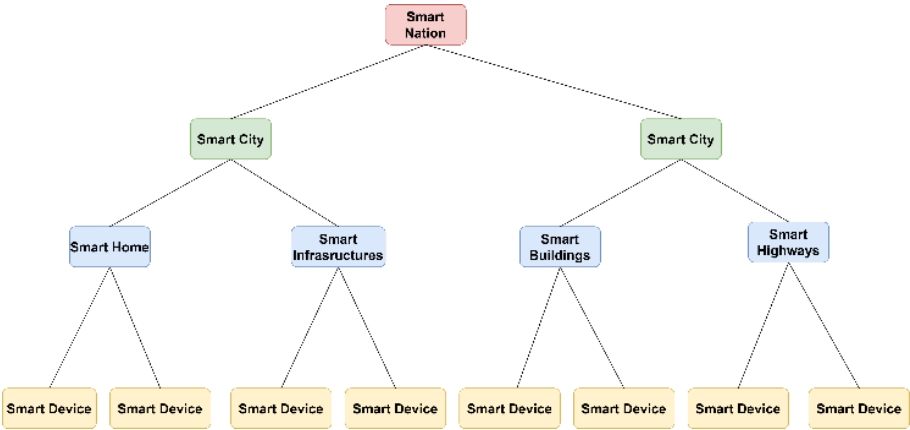
- number and nature of rescue teams,
- kind of equipment,
- mode of transportation,
- location of the relief camps.

There is a need for a decision support system that combines all the collected data and helps in producing an insight into the disaster.

## **Smart City and Smart Nation**

The migration of people to metropolitan cities for certain reasons makes management difficult in those regions. The population in cities in Europe was 50 percent in 1950, now it is over 75 percent [2]. The urban population will be approximately 6.7 billion in 2050 [3]. As a result of this urbanization, systems generate big data. The smart city [4] concept is about collecting these data and using them in decision support systems. Technologies such as IoT,

AI (artificial intelligence), cloud computing [5], and blockchain [6] can be used. These all can be used to form a smart nation as shown in Fig. 1.



**Fig.1:** Smart Nation Concept

**Digital Twin**

Digital twin technology is based on creating a virtual copy of an object, process, or human being in a virtual world. This virtual copy is kept current with continuous incoming data.

The concept of digital twin is not new. This technology was used for the first time in the 1960s by NASA (National Aeronautics and Space Administration). NASA has created virtual duplicate systems for a space mission to test its equipment and estimate potential costs. The “digital twin” term was introduced in 2002 by NASA’s advisor Dr. Michael Grieves. It was also developed in the US Air Force and applied to the aircraft body models in 2013. Digital twin technology has become one of the most promising technology trends with the widespread usage of IoT (Internet of things).

Essential simulations and examinations to test a system, mechanism or a tool can be done easily with this technology reducing potential costs and manpower as well as avoiding harms to human life. The digital twin consists of three main concepts; physical assets, sensors, and virtual copy. The virtual copy displays the data from the IoT devices as 2D/3D or non-visual. Decision-making process of humans or machines can become easier using this display.

As open-source software has become more popular in the computer industry [7], it is also sensible to use a variety of these open-source protocols,

frameworks, and tools for digital twins and IoT. Eclipse Ditto, AllJoyn, and IoTivity frameworks can be used. Communication protocols that can be used are; MQTT (Message Queue Telemetry), AMQP (Advanced Message Queueing Protocol), COAP (Constrained Application Protocol), XMPP (Extensible Messaging and Present Protocol) [8] and RabbitMQ.

### Digital Twin Communication

One of the most important features of humankind is the ability to communicate with each other. Life proceeds in a certain order with the help of this communication. Roads, buildings and infrastructures in the smart city which were created exist individually in the system. They can act by each other's feedback.

The connectivity mostly occurs between IoT devices. The number of IoT connected devices increased by 285% since 2015 and the number of these devices is around 38.5 billion [9]. The sectors such as manufacturing, logistics, health, agriculture, and automotive use IIoT (Industrial Internet of Things) often. Disruptions, delays, and stoppages of jobs can cause substantial losses. Therefore, communication is an important component for the IIoT. In smart cities, the communication must not be interrupted between digital twins for the sustainability of management and life.

The fourth period of the industry has begun (Industry 4.0) which is based on CPS (cyber-physical systems). CPS is a hybrid system that has a physical and virtual environment in communication with each other. In CPS everything is interconnected wirelessly. Smart factories, autonomous cars, quadcopters, etc. are substantial examples of CPS.

**Table 1.** IoT Connectivity Types

| Type      | Radio Frequency    | Range       | Data Transfer Speed |
|-----------|--------------------|-------------|---------------------|
| Bluetooth | 2.4 GHz-2.5 GHz    | 0,5 m-100 m | $\leq 24Mbps$       |
| Wi-Fi     | 2.4 GHz-5 GHz      | 45 m-91 m   | $\leq 250Mbps$      |
| Zigbee    | 860 Mhz-2.4 GHz    | 10 m-100 m  | $< 250Kbps$         |
| GPRS      | Diff. in each reg. | 8 km-40 km  | 56-114 Kbps         |
| 3G        | Diff. in each reg. | 8 km-40 km  | 0,8-2 Mbps          |
| 4G        | Diff. in each reg. | 8 km-40 km  | 0.2-1 Gbps          |
| 5G        | Diff. in each reg. | 8 km-40 km  | $\geq 1Gbps$        |

These devices are in constant communication with each other and also connected to the internet. There are lots of different ways to make these

connections such as RFID, ZigBee, WPAN, WSN, DSL, UMTS, GPRS, WiFi, WiMax, LAN, WAN, 3G, etc [10]. These types are given in Table 1.

## **RELATED WORKS**

There are mainly touristic implementations in the smart city concept. Newcastle, Dubai, Qatar and Singapore started to form their digital twins. Qatar and Singapore have the aim to use digital twins in a touristic sense. In the three-dimensional virtual copy of the city, it also aimed for people to go on the streets and see touristic places [11].

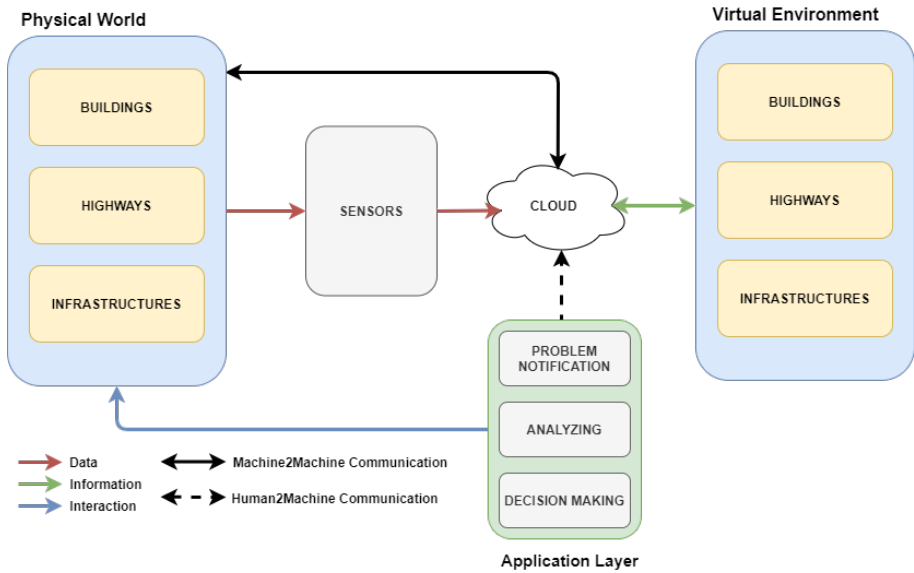
However, using digital twins for disaster management is rare. Digital twin technology is used for so many different projects and studies in the literature. [12-15]. For example, it is used for controlling floods in Newcastle by monitoring the amount of water on the roads [16].

However, there is a lack of academic studies that can be used for better solutions and standardization. Only a recent study [17] addressed this idea and proposed a “disaster city” concept. Our study is an attempt to develop a proof of concept how to implement Digital Twin on a disaster management setting where further studies can be built upon. This prototype will address a variety of disaster (earthquake, storm, fire etc.) types and collect data from a variety of sources such as electricity, natural gas and communication infrastructures of the buildings are planned to be controlled.

## **SYSTEM PROPOSAL**

The proposed system aims to develop essential measures for pre- and post-disaster environment by performing various simulations on a virtual copy. These measures include information on as buildings, transportation, telecommunication, natural gas and electricity networks. This system could also be used real-time to collect data following a disaster to decide how rescue operations are managed by understanding of the occupancy rate of critical facilities such as hospitals.

Digital twin architecture is shown in Fig. 2. Sensors collect data from the physical world and sent it to the virtual environment which is stored in the cloud. Continuous flow of data is the most essential and crucial requirement of the system. This data can be collected from many sources such as IoT sensors, social media, data resources from government and municipality. This data will be stored in database systems on the cloud.

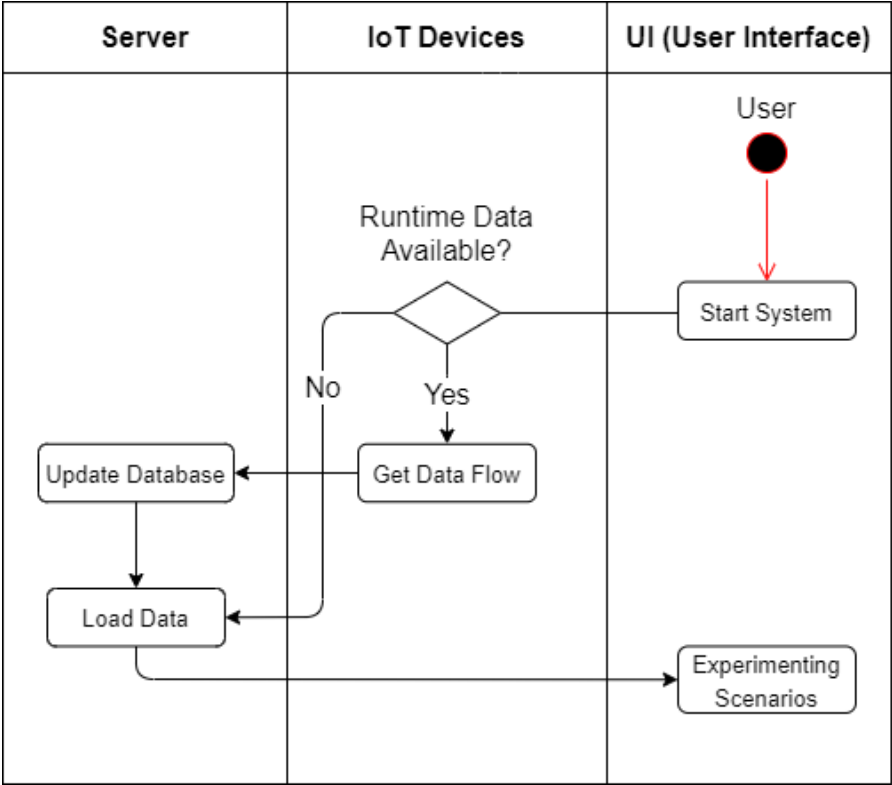


**Fig. 2: Digital Twin Architecture**

There are two main actors of the system that are the staff at the disaster management centers and the rescue teams. Organizations can use this system from the rescue centers and the teams can use these data during the intervention. Also in the post-disaster scenario; buildings and rescue center information of the relevant streets can be checked by clicking on the icon of buildings and the rescue center. In the infrastructures section, electricity, water, and telecommunication facilities can be inspected. A timeline button is used in order to analyze the effects of the disaster on time. The system can be run in two modes:

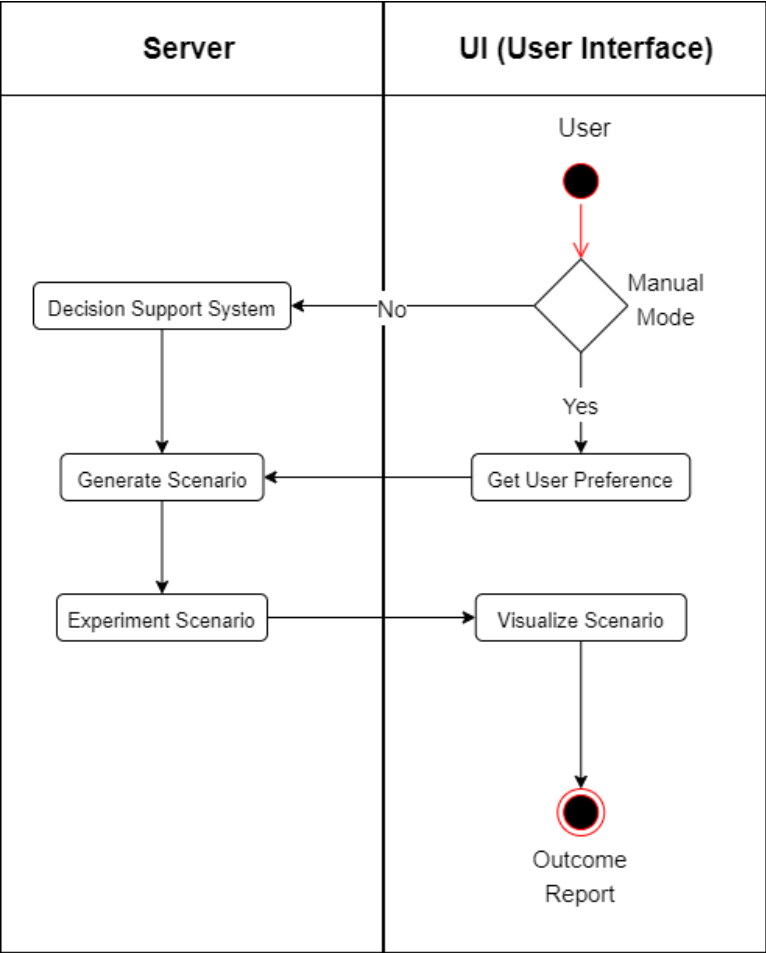
- **Education mode:** Based on the decision of the management staff, the simulation is made and possible outcomes are reported.
- **Estimating mode:** Decision support system runs in several different choices and tries to find the best decision for the best outcome.

The action diagram of the system is in Fig 3. Diagram consists of three components; server, IoT Devices, and UI (User Interface). When the user starts the system, the data can be received from realtime data or stored data. The system may have dynamic live data obtained instantly from the IoT devices. If not, previously-stored simulation data is loaded and relevant scenarios are run on this data.



**Fig. 3:** Action Diagram of the System

Experimenting scenarios are given in detail in Fig. 4. This can be done in two ways. In manual (education) mode, choices are made by the trainee. Users' preferences are taken and according to these preferences, appropriate scenarios are created. Otherwise, the decision support system creates different scenarios by itself and decides which result is most effective in these different scenarios. Also, the decision support system can make changes to the real IoT devices. At the end of the scenario, a report is presented which includes how many people survived, state of buildings, and infrastructures. Depending on these, the effectiveness of the scenario is analyzed.



**Fig. 4:** Action Diagram of Experimenting Scenarios

**Table 2.** GUI Types

|    | Advantages                      | Disadvantages                       |
|----|---------------------------------|-------------------------------------|
| 3D | More accurate modelling         | Difficult to model very large areas |
|    | Suitable for buildings, bridges | Need more computational power       |
| 2D | Large areas easily to build     | Hard to model complex structures    |
|    | Low computational power         | Less detail, low accuracy           |

The graphical user interface is also one of the issues to be considered. It can be examined under two main options: 3D or 2D. All of these approaches have important advantages and also limitations as well. Some of the strengths and weaknesses of each system are highlighted on Table 2. 3D modelling requires too much computational power and time. Since our sources are

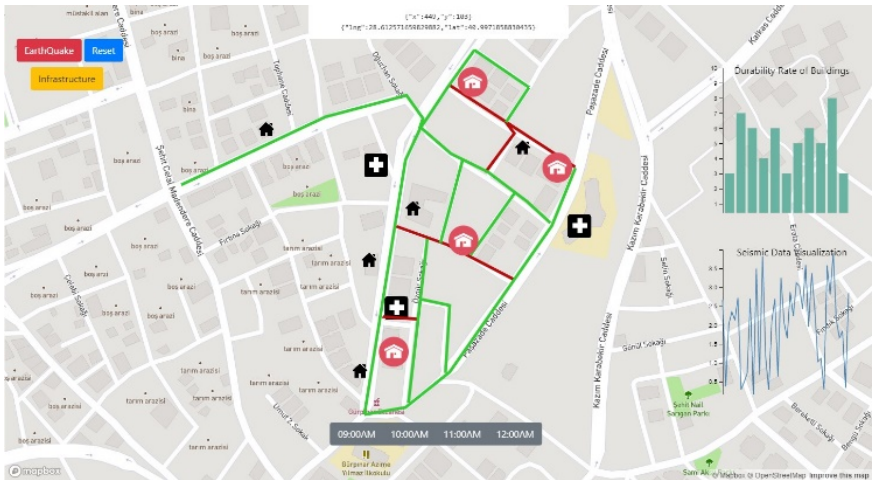
limited, it also was not considered appropriate. Given all these situations, 2D model was chosen for the system.

## IMPLEMENTATION

The implementation aims to educate the emergency centre staff. The system prototype for the education mode is made where an earthquake scenario was implemented. The event of a disaster will be simulated; the situation from the beginning of the earthquake in a timeline can be analyzed by users. Post-earthquake infrastructure data can also be examined. Possible outcomes of the decisions will be reported accordingly.

A prototype simulation is implemented as a proof of concept on an ordinary laptop with 8 GB RAM, 256 GB SSD, 6 GB GTX 1050 Ti and an Intel i5-8300H 2.3 GHz CPU.

The incoming and stored data must be visualized to make it more meaningful. MapBox is selected to create the map service [18]. Conditions of the buildings and the roads are marked by visualization elements.



**Fig. 5: Post-Disaster Prototype Interface**

The user enters the map service and reaches the necessary live pre-disaster visualizations and analysis about buildings, rescue centres, and roads. The prototype interface is shown in Fig. 5. The user can switch to the map style containing the infrastructure information and analyzes the water, electricity, and communication infrastructures. This prototype interface is shown in Fig. 6. The environment is simulated with instant data.



**Fig. 6:** Post-Disaster Prototype Interface for Infrastructures

The simulation environment is used to show the status of the buildings, infrastructures (such as communication, power, electricity, water and natural gas) and the roads in a timeline after the disaster. The most appropriate directions/paths for the rescue teams are given by the decision support system. The system recommends the following:

- Number and type of rescue teams according to the civilization density and the status of the buildings.
- The best live-saving scenario(s) with the most appropriate paths from the best available rescue centre to the wreck.

Simulation provides the users several options with potential success rates. When a decision is chosen, the outcome is also shown in the simulation. Then a rescue team vehicle departs from the rescue centre and goes from the shortest and safest way to save the people. The outcomes of this decision is presented to the user.

A system that uses artificial intelligence algorithms such as Breadth-First Search (BFS) [19] and Uniform-Cost Search (UCS) [20] is being developed. This will help in recommending the best options which will help rescue teams to identify better alternative to save lives.

Social media such as Twitter is an important communication channel in times of emergency. The iniquitousness of smartphones enables people to announce an emergency in real-time. Because of this feature, more agencies

(i.e. disaster relief organizations and news agencies) are interested in programmatically monitoring Twitter. It's not always clear whether a person's words are announcing a disaster. An NLP (Natural Language Processing) model [21] which detects whether the tweets are fake or not was developed. Transfer learning, particularly models like Allen AI's ELMO, OpenAI's OpenGPT, and Google's BERT [22] is being used to smash multiple benchmarks with minimal task-specific fine-tuning. These are also providing the rest of the NLP community with pre-trained models that could easily (with less data and less compute time) be fine-tuned and implemented to produce state of the art results. Hence, in this project, it is decided to use Google's pre-trained BERT model.

Google's pre-trained BERT model is preferred in this project. The dataset used for training the model consists of two components. The first is the text of the tweet. The other is the estimation of whether the tweet is about a real disaster (1) or not (0). This dataset was produced using a CSV file obtained from Kaggle [23]. This file contains information; a unique identifier for each tweet, the location the tweet was sent from, and a particular keyword from the tweet. This dataset contains 7613 tweet texts. The size of the dataset is relatively small compared to modern deep learning models. Hence, the data was split into an approximately 80:10:10 ratio for the training set, development set, and test set respectively. This breaks up of 6091 training examples, 1522 validation examples, 1522 test examples. Binary cross-entropy is chosen as a loss function of this model. AdamW is chosen for optimizer of the model, because it optimizes the learning rate used in gradient descent.

## **RESULTS**

The results of the simulation as part of our first experiments are promising. The simulation covered a small area, therefore the amount of data was limited. Building and infrastructure data was entered manually to the database. The "Earthquake Activity in Istanbul in the Last Year" dataset [24] from IBB open data was used. However, for a digital twin; the data must be provided dynamically from social media and IoT sensors. The system will be better when such an environment is provided. Using live data is to be implemented in a future study. Moreover, also the NLP model result were promising. The dataset was pretty simple and small which is not good for BERT model. It is trained just 4 epochs. After training section, accuracy on

test data is about 67% that is not considered as. Increasing the size of dataset and making necessary optimizations for better performance of model is considered as a future study.

## **CONCLUSION**

Natural and manmade disasters are an inevitable part of our daily lives and governments, institutions, local authorities and individuals must be prepared for such an event. The necessary measures before the disasters are key to reduce the damage on property and the number casualties. “Earthquakes don't kill people, buildings do” expression should not be forgotten.

Disaster management is an important issue to be addressed and digital twin technology can make it easier for us. The concept is still new for many countries, but it is expected to become mainstream within the next five to ten years. Simulations in the virtual copy can help us to learn and prepare the necessary precautions. After the disaster, even seconds are important for saving lives. It is pretty crucial to intervene immediately. Digital twins can also help us with this issue.

Test results are promising and simulation scope is planned to be extended. The simulation results are to be used to have the idea of taking necessary precautions on housing, transportation, telecommunication, natural gas, and electricity networks. The real-time data after a disaster is planned to be used on deciding how rescue operations to be carried out, figuring out the fullness/sturdiness ratio of the mission-critical buildings like hospitals.

Future works will include running this whole system in the cloud, as the amount of data from sensors and social media will increase gradually. These sensors will be placed on the buildings and infrastructures like electricity, transportation, telecommunication, and natural gas. This study will be extended to the city level and a central intelligence in which all regions facing the risk of earthquakes will be considered.

## REFERENCES

- [1] Coppola, D. P. (2006). Introduction to international disaster management. Elsevier.
- [2] Caragliu, A., Del Bo, C., & Nijkamp, P. (2011). Smart cities in Europe. *Journal of urban technology*, 18(2), 65-82.
- [3] United Nations, "World Urbanization Prospects: The 2018 Revision
- [4] Su, Kehua, Jie Li, and Hongbo Fu. "Smart city and the applications." 2011 international conference on electronics, communications and control (ICECC). IEEE, 2011.
- [5] Brian, H., Brunschwiler, T., Dill, H., Christ, H., Falsafi, B., Fischer, M., & Kaiserswerth, M. (2008). Cloud computing. *Communications of the ACM*, 51(7), 9-11.
- [6] Nofer, Michael, et al. "Blockchain." *Business & Information Systems Engineering* 59.3 (2017): 183-187.
- [7] Tan, L., & Wang, N. (2010, August). Future internet: The internet of things. In 2010 3rd international conference on advanced computer theory and engineering (ICACTE) (Vol. 5, pp. V5-376). IEEE.
- [8] Madalli, D. P., Barve, S., & Amin, S. (2012). Digital preservation in open-source digital library software. *The Journal of Academic Librarianship*, 38(3), 161-164.
- [9] Sun, H., Wang, C., & Ahmad, B. I. (2017). *From Internet of Things to Smart Cities: Enabling Technologies*. Chapman and Hall/CRC.
- [10] S. Smith, IoT Connected Devices to Triple to Over 38Bn Units, Juniper Research, 2020. [Online]. Available:<https://www.juniperresearch.com/press/press-releases/iot-connected-devices-to-triple-to-38-bn-by-2020>
- [11] Farsi, Maryam, et al., eds. *Digital Twin Technologies and Smart Cities*. Springer, 2020.
- [12] Aydın, Ö., & Karaarslan, E. (2020, June). Covid-19 Belirtilerinin Tespiti İçin Dijital İkiz Tabanlı Bir Sağlık Bilgi Sistemi. Online International Conference of COVID-19 (CONCOVID).
- [13] Yukcu, S., & Aydın, Ö. (2020). Maliyet Düşürme Yöntemi Olarak Dijital İkiz. *Muhasebe Bilim Dünyası Dergisi*, 22(3), 563-579
- [14] Fan, C., Zhang, C., Yahja, A., & Mostafavi, A. (2021). Disaster City Digital Twin: A vision for integrating artificial and human intelligence for disaster management. *International Journal of Information Management*, 56, 102049.

- [15] Yükçü, S., Aydın, Ö., & Koçakoğlu, Ü. Ö. (2020). Bağımsız Denetimde Dijital İkiz Uygulaması. Mali Cozum Dergisi/Financial Analysis, 30(161).
- [16] Ham, Youngjib, and Jaeyoon Kim. "Participatory Sensing and Digital Twin City: Updating Virtual City Models for Enhanced Risk-Informed Decision-Making." *Journal of Management in Engineering* 36.3 (2020): 04020005.
- [17] Damjanovic-Behrendt, V., & Behrendt, W. (2019). An open source approach to the design and implementation of Digital Twins for Smart Manufacturing. *International Journal of Computer Integrated Manufacturing*, 32(4-5), 366-384.
- [18] CADENAS, Cecilia. Geovisualization: Integration and Visualization of Multiple Datasets Using Mapbox. 2014.
- [19] Beamer, Scott, Krste Asanovic, and David Patterson. "Direction-optimizing breadth-first search." *SC'12: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*. IEEE, 2012.
- [20] VERWER, Ben J.. H.. ; VERBEEK, Piet W.. ; DEKKER, Simon T.. . An efficient uniform cost algorithm applied to distance transforms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, 11.4: 425-429.