Approximate Reasoning for Data-driven Tasks Management at the Edge

Ms Tahani Aladwani

Supervisors: Dr Christos Anagnostopoulos, Dr Fani Deligianni

University of Glasgow

School *of* Computing Science
Essence: Data Science & Distributed Computing Lab

Introduction

Emerging technologies such as the Internet of Things, social networking, and online games have caused a significant increase in the volume of data being generated at the network edge. Meanwhile, the concept of data-driven tasks has drawn increased attention in the last few years. Data-driven tasks rely heavily on data generated by smart devices (e.g., sensors and smartphones) to build knowledge and make decisions.

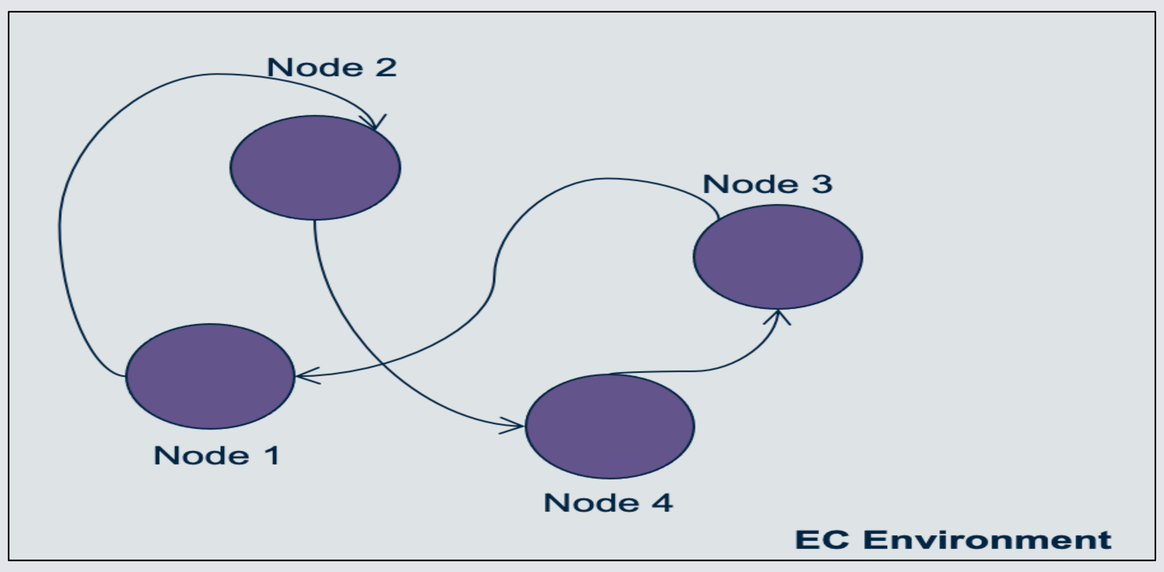
Execution of data-driven tasks can happen on:

- Smart devices
- Cloud Computing or
- Edge computing

Research question : How can data-driven tasks obtain all the required data from various sources without sacrificing the device's computation resources?

Problem Statement

EC system with $\boldsymbol{N} = \{n_1, n_2, n_3, \dots, n_n\}$ Nodes. Data points x= $[x_1, x_2, \dots, x_n]^T \in \mathbb{R}^d$. Each node n_i has a neighbourhood $N_i \subset N$ of directly communicating nodes. Moreover, node n_i communicates with the end-users/applications and a remote cloud server. n_i can execute locally certain data-driven tasks.

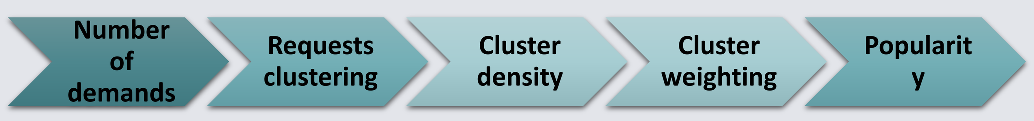


Methods & Results

Task Management Mechanism Factors

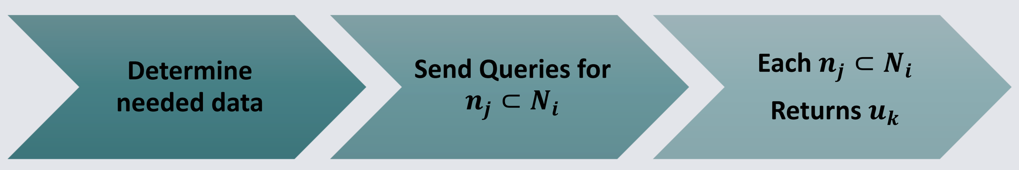
We introduce the basic factors for inferring the right execution decision for each task $\boldsymbol{T_k}$ on each node $\boldsymbol{n_k}$ corresponding to :

1. Task popularity p_k



2. Outliers o_k

3. Data overlapping u_k



Fuzzy Reasoning Approach

All factors are fed to a Fuzzy logic (FL) Inference System to derive the 'probability of offloading r_k '.

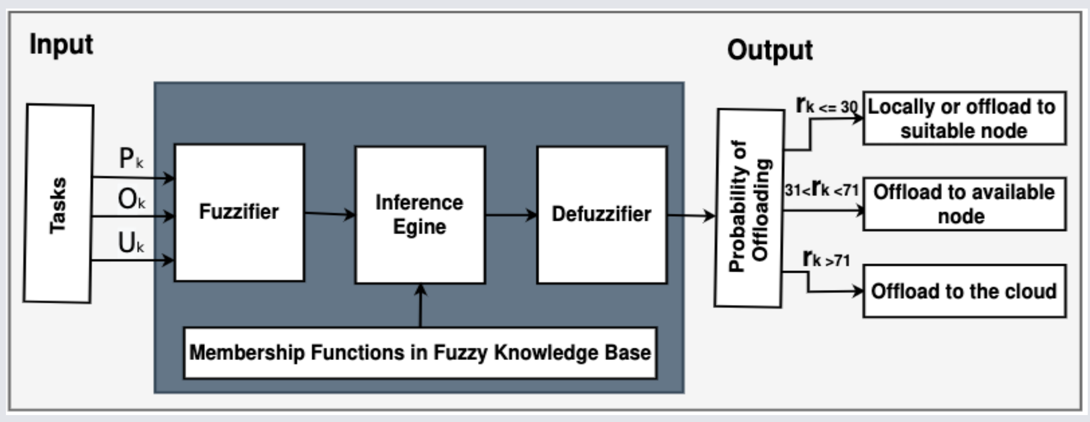


Figure 1. Fuzzy logic inference system.

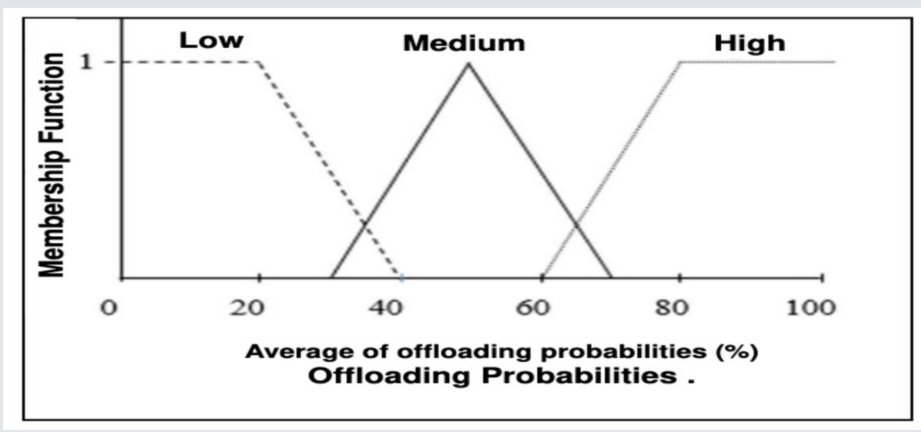


Figure 2. The probability of offloading.

Experiment

We used synthetic dataset to simulate tasks' popularities while the data overlapping experiment has been carried out on real datasets that are collected by four Unmanned Surface Vehicles (USVs) working as nodes n_i to collect data from sensors in a coastal area, dataset link: http://www.dcs.gla.ac.uk/essence/funding.html#GNFUV. In order to obtain the task's data overlapping $\boldsymbol{u_k}$, we have defined for each local dataset $\boldsymbol{D_i}$ the feature boundaries

$$\boldsymbol{D_i} = [x_1^{min}, x_1^{max}, x_2^{min}, x_2^{mix}]$$

The FL engine has been developed in MATLAB considering the popularity p_k of tasks T_k between [1, 40] and outlier o_k either 0 or 1, while the percentages of data overlapping u_k are between [1, 100].

Results:

- **Probability of offloading (r_k) for each task (T_k).**

Table 1. The probability of offloading r_k for each task T_k.

T_k	p_k	u_k	u_k	$M_1(r_k)$	$M_2(r_k)$	$M_3(r_k)$
T_1	Low	Yes	Low	83%	84%	57%
T_2	Med	No	High	30.4%	32%	42%
T_3	Low	No	Low	85%	86%	72.9%
T_4	Med	No	Med	65%	68%	53%
T_5	High	No	High	23%	17.6%	35%
T_6	Low	Yes	Low	85%	86.5%	72.7%
T_7	Med	No	Med	44.7%	37%	47%
T_8	High	Yes	High	14.4%	14.5%	17.7%
T_9	High	Yes	High	15%	15.8%	27.1%
T_{10}	Low	Yes	Low	83%	85%	67.2%
-	-	-	-	9/10	8/10	6/10

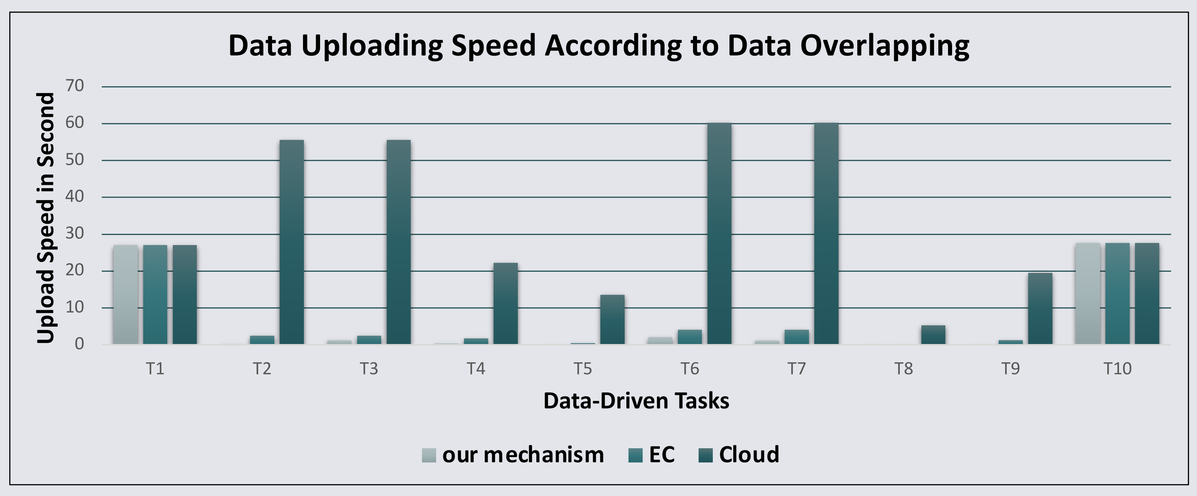
- **Data uploading speed and Execution time**

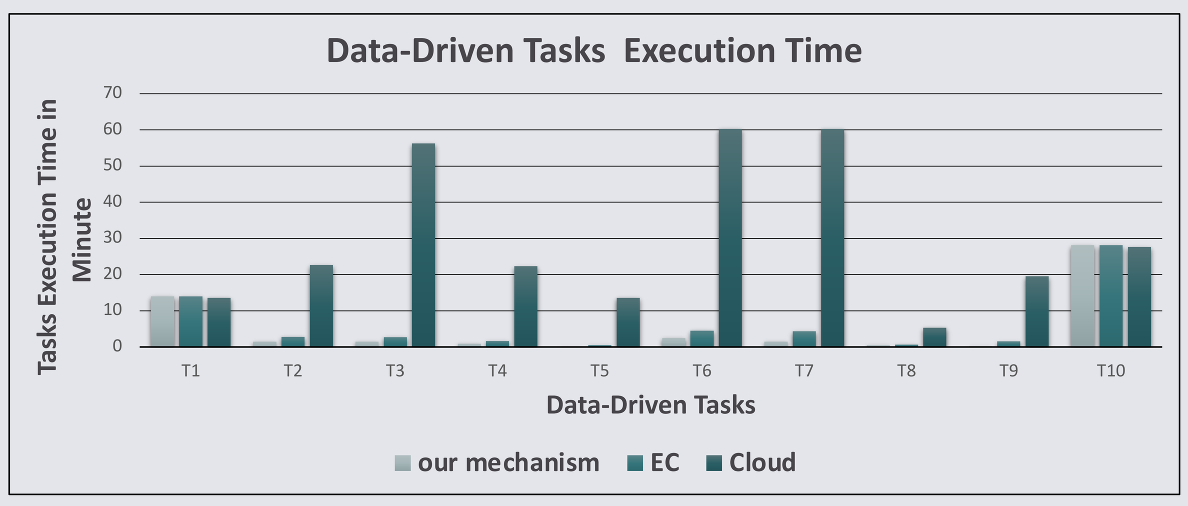
Simulator:
CloudSim Plus [*]

Models:
EC, Ours, Cloud

Performance Metrics:

- Data Uploading Speed
- Task Execution Time





Conclusions

Our mechanism performance has been evaluated according to the probability of offloading data-driven analytics tasks to the correct nodes according to the optimal solution and against two other mechanisms.

As evidenced by the results, our mechanism significantly outperforms the benchmark mechanisms in terms of decision-making accuracy. Furthermore, this mechanism can reduce the probability of a task being offloaded to an unsuitable node by up to 90 %. In addition, our method has been evaluated in terms of resource utilization, showing that it provides higher data uploading speeds compared to EC-based and cloud-based methods.

References

[1] F. Samea, F. Azam, M. Rashid, M. W. Anwar, W. Haider Butt, and A. W. Muzaffar, "A model-driven framework for data-driven applications in serverless cloud computing," Plos one, vol. 15, no. 8, p. e0237317,2020.

[2] K. Kolomvatsos and C. Anagnostopoulos, "A deep learning model for demand-driven, proactive tasks management in pervasive computing," IoT, vol. 1, no. 2, pp. 240–258, 2020.

[3] J. Cheng and D. Guan, "Research on task-offloading decision mechanism in mobile edge computing-based internet of vehicle," J. Wireless Communications and Networking, vol. 2021, no. 1, pp. 1–14, 2021.

[4] K. Cui, B. Lin, W. Sun, and W. Sun, "Learning-based task offloading for marine fog-cloud computing networks of usv cluster," Electronics, vol. 8, no. 11, p. 1287, 2019