Abstract

Motivated by increased mobile communication traffic, a vast amount of data processing and associated energy consumption, as well as strict latency requirements, the author of this thesis presents his research on energy consumption minimization in fog computing networks that distribute communication and computation services along the cloud-to-end-devices continuum. Task execution latency constraints are also considered.

The thesis of this dissertation is the following: *There exist optimal solutions to computational task offloading problems in fog networks, minimizing energy consumption while maintaining required latency levels.* The main goal of the thesis is to propose such solutions.

After the state-of-the-art research on energy-aware fog computing networks is analyzed, the author's original contributions to solving the problem of communication and computing task allocation in fog networks are presented. First, the author focuses on modeling the delay and energy consumption within the fog and cloud tiers of the network. Models are parameterized using values representing real-world equipment for communication and computing resources and diverse user requests. Results presenting the impact of different core network parameters on energy consumption and delay in fog computing networks are shown for various parameter setups.

Then, the author formulates an optimization problem to find an assignment of offloaded tasks to nodes in the fog and cloud tiers that minimize energy consumption while keeping their delay requirements. The objective function includes energy costs related to transmission and computing, while the optimization space includes choice of Fog Nodes (FNs) and Cloud Nodes (CNs) executing the tasks as well as their Central Processing Unit (CPU) frequencies. Two solutions, called Energy-EFFicient Resource Allocation (EEFFRA) and Low Complexity (LC)-EEFFRA, to this non-convex optimization problem are proposed. The simulation results for various input parameters are provided and compared against the benchmark algorithms.

Next, the fog-network model is expanded by including the wireless transmission between Mobile Devices (MDs) and FNs. It has an impact on the objective function and constraints, and adds a new set of decision variables. Despite this, an analytical solution to the optimization problem is found. The results are examined for various input parameters and compared against those achieved by the baseline solutions.

Finally, the author explores the offloading of tasks modeled as sequential graphs. These tasks consist of smaller subtasks, each of which can be processed at a different node. The optimization problem is still about minimizing energy consumption and maintaining the required delay while the proposed solution involves clustering similar nodes to significantly reduce the size of the search space.

The major conclusion of this dissertation is that the author's original solutions can significantly reduce energy consumption in the fog network with latency constraints compared with standard cloud-delegation practices. Key parameters, such as arithmetic intensity, are identified and their impact on the efficiency of offloading solutions is shown through the results of multiple simulations.