Status-aware and SLA-aware QoS Routing Model For SDN Transport Network

Atefeh Meshinchi, Ranwa Al Mallah, Alejandro Quintero
Presenter:
Atefeh Meshinchi
atefeh.meshinchi@polymtl.ca

Laboratoire de recherche en réseautique et informatique mobile (LARIM)
Ećole Polytechnique de Montréal, Montréal, Canada

October 5, 2020

Bio

Atefeh Meshinchi received the B.E. degree in computer engineering from the Shahid Beheshti University, Tehran, Iran, in 2004, and the M.S. degree in Computer engineering from Montreál University, Ećole de Polytechnique, Montrál, Canada, in 2018. She has almost 15 years of experiences in telecommunication industry, and currently working as the Solution Architect at Bell Canada, Montreál, Canada.

Topics of research interest

- Internet of Things
- Software Defined Networking
- Resource allocation and performance optimization
- Security
- Mathematical modeling and optimization algorithm

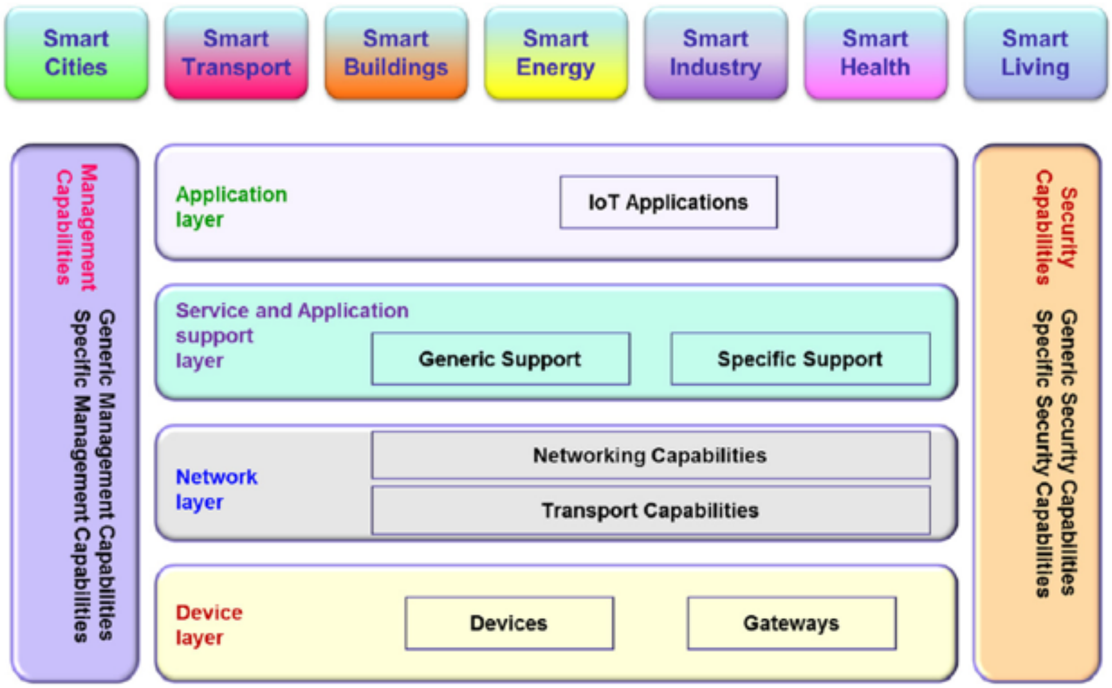
Table of Contents

Table of Contents

Motivation

- Applications are fundamental to the IoT. A large numbers of heterogeneous devices, mostly low-end sensors give the opportunity to develop a wide range of innovative applications and services with stochastic and dynamic traffic, consequently they could have various QoS requirements.
- Internet world-wide availability plus IP stack capability added to smart objects make the Internet the best available choice as transport and communication infrastructure in IoT system. However, the best-effort mechanisms implemented in legacy computer network and the current Internet could impose limitations for efficiency of IoT services. Vendor-dependent platforms and interfaces make the Internet evolution complex and slow to adapt to new business needs.
- IoT system currently lacks the standardized and unified QoS support service framework which considers the inherent heterogeneous, stochastic, and dynamic nature of IoT. It is also challenging to have a standard SLA (Service Level Agreement) for IoT applications.

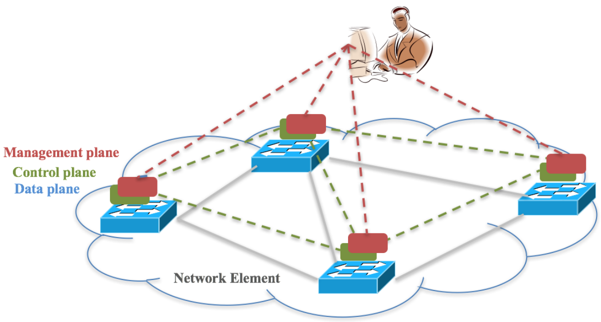
IoT architecture



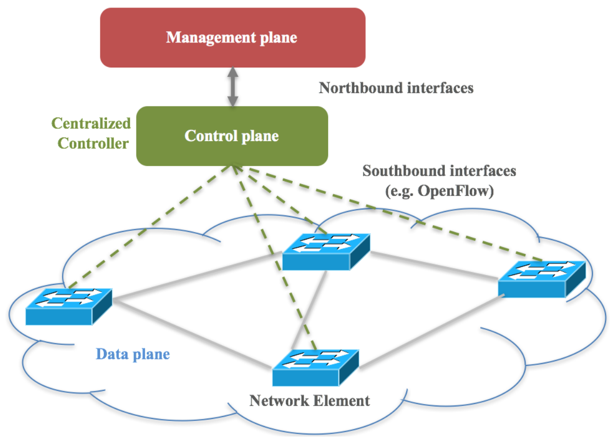
IoT Reference architecture [1]

Software-Defined Network architecture

Traditional Network



Software-Defined Network



- Closed equipment
- Distributed control
- High equipment and operation cost

- Centralized management
- Network programmability
- CAPEX and OPEX efficiency

Figures reproduced from [2]

Table of Contents

Related works

QoS-support design and QoS-based design approaches

- ✓ *QoS-aware scheduling of services-oriented Internet of things* [3]
- ✓ *Middleware to support sensor network applications(MiLAN)* [4]
- ✓ *SDIoT: a software defined based Internet of things framework* [5]

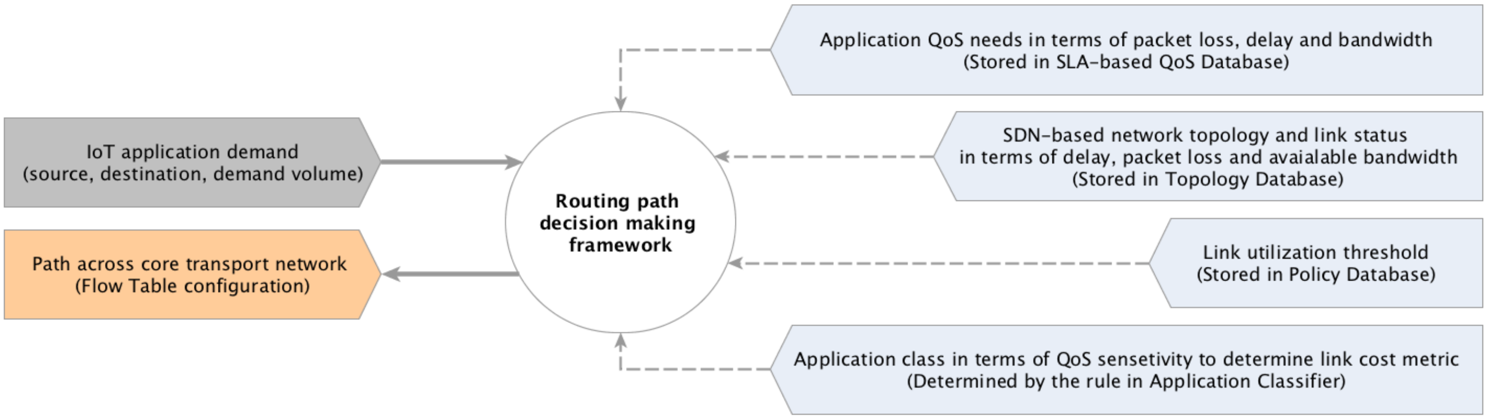
Table of Contents

Proposed Model

IoT framework:

- ✓ SDN technology integrated into IoT architecture
- ✓ Programmable middle-ware layer for IoT system
- ✓ Centralized control over network resources

Decision-making framework for routing calculation across SDN-capable network:



Model definition

SDN-based communication network as an oriented graph $G = (V, E)$:

- List of network elements: $1...v$, $v \in V$;
- List of bidirectional links between network elements: $(i,j) \in E \cup (j,i)$;
- Demands: $1...k$;

Network link status and policy parameters:

$B_{ij} > 0$	Maximum capacity on the link (i,j), $[Mbps]$
$b_{ij} \geq 0$	Available capacity on the link(i,j), $[Mbps]$
$d_{ij} \geq 0$	Delay on the link (i,j), $[Second]$
$pl_{ij} \geq 0$	Packet loss ratio on the link (i,j), $[Percentage]$
$u_{Threshold} > 0$	Link utilization limit on the link, $[Percentage]$

Application Qos and demand parameters:

$S^k \in V$	Source of demand k
$T^k \in V$	Destination of demand k
$F^k \geq 0$	Total demand volume k, $[Mbps]$
$D^k_{SLA} \geq 0$	Maximum acceptable delay for demand k, agreed in SLA
$PL^k_{SLA} \geq 0$	Maximum acceptable packet loss ratio for demand k, agreed in SLA
$B^k_{SLA} \geq 0$	Minimum bandwidth required for demand k, agreed in SLA
P^k	Determined path across the network for demand k

Routing calculation algorithm

Algorithm 1: Proposed QoS support routing algorithm

Input : $G = (V, E)$ as network topology: $V = \{1, 2, ..., v\}$ as nodes and $E = \{(i, j) : i, j \in V, i \neq j\}$ as bidirectional links

Input : (S^k,T^k, F^k) per each demand k in K: $S^k \in V$ as Source of demand, $T^k \in V$ as Destination of demand k, $F^k \geq 0$ as Total demand volume, $[Mbps]$

Output : P^k: Routing path across the network for demand k

Procedure:

for *(i, j) in E* **do**
- Get network link QoS parameters ($b_{ij} \geq 0$, $pl_{ij} \geq 0$, $d_{ij} \geq 0$, $B_{ij} > 0$;
- Calculate the current link utilization rate ;
- Get the link utilization limit $u_{Threshold}$;
- **if** *Link utilization rate* $\geq u_{Threshold}$ **then**
  - Exclude this link from the logical network topology to calculate paths for current active demands;
- **end**

end

for *k in K* **do**
- Get Max acceptable delay D^k_{SLA}, Max acceptable packet loss PL^k_{SLA}, and Min required bandwidth B^k_{SLA} for each demand k in K from SLA-DB ;
- Set the link cost metrics depending on the application class, if any;

end

for *k in K* **do**
- Calculate the best-fit path per each demand k in K based on the proposed mathematical model;

end

Application classification approach

Application classes:

1. Control applications, also called mission-critical applications,such as city-traffic management and emergency management, which need very fast response with as less error as possible,
2. Monitoring applications, such as intelligent security surveillance tasks, which are fed by cameras and need more throughput,
3. Analysis and inquiry applications, such as the inquiry into the transported item state in the intelligent logistics, which are throughput and delay tolerant.

Applications prioritization and queuing approach:

Application Class	**IoT application**	**QoS attributes**	**Priority**	**Type of queue**	**Cisco classification**
Delay-Centric	Mission-critical (event-based application)	$D^k_{max} \leq D_{Threshold}$	1	PQ (Priority Queue)	EF (Expedited Forwarding)
Bandwidth-Centric (Multimedia application)	Real-time monitoring, query-driven application	$D^k_{max} \geq D_{Threshold}$ $BW^k_{min} \geq BW_{Threshold}$	2	Q1	AF (Assured Forwarding)
General	Non-Real time monitoring, analytic application	N/A	3	Q2	BE (Best Effort)

Table of Contents

Objective function:

Multi commodity and constraint based routing allocation: to flow application in the network with minimum cost in respect to the particular cost metrics in multi-service environment:

$$\text{Minimize} \quad \sum_{(i,j)\in E} \sum_{k\in K} C_{ij} X_{ij}^{k}$$

X_{ij}^{k}: Amount of volume associated to the demand k on the link (i,j), $[Mbps]$

$$0 \leq X_{ij}^{k} \leq b_{ij}$$

C_{ij}:Unit cost of the link (i,j)

$$C_{ij} = \alpha \times b_{ij} + \beta \times pl_{ij} + \gamma \times d_{ij} \quad , \forall (i,j) \in E$$
$$\alpha, \beta, \text{ and } \gamma \text{ as scaling factors}$$
$$\alpha + \beta + \gamma = 1 \ , \ 0 \leq \alpha, \beta, \gamma \leq 1$$

Feature Scaling method used for normalization of the range of b_{ij}, pl_{ij}, and d_{ij} into the range [0, 1].

Constraint function: Network-Level

- Link Bandwidth Constraint:

$$\sum_{k \in K} X_{ij}^{k} \leq b_{ij}, \ \forall (i,j) \in E$$

- Link Utilization Constraint:

$$\sum_{k \in K} X_{ij}^{k} \leq b_{ij} - (1 - u_{Threshold}) \times B_{ij}, \ \forall (i,j) \in E$$

- Flow Conservation Law Constraint:

$$\sum_{(i,j) \in E} X_{ij}^{k} - \sum_{(j,i) \in E} X_{ji}^{k} = \begin{cases} F^{k}, & i = S^{k} \\ -F^{k}, & i = T^{k} \\ 0, & i \neq S^{k} and \ i \neq T^{k} \end{cases}$$

Constraint function: Application-Level

- Delay Constraint:

$$d_p^k \leq D_{SLA}^k$$

$$d_p^k = \sum_{(i,j)\in E, p^k} d_{ij}$$

$$\textstyle\sum_{(i,j)\in E, P^k} d_{ij} \leq D_{SLA}^k, \ \forall k \in K$$

- Packet Loss Constraint:

$$pl_p^k \leq PL_{SLA}^k$$

$$pl_p^k = \sum_{(i,j)\in E, p^k} pl_{ij}$$

$$\textstyle\sum_{(i,j)\in E, P^k} pl_{ij} \leq PL_{SLA}^k, \ \forall k \in K$$

POLYTECHNIQUE MONTRÉAL
LE GÉNIE EN PREMIÈRE CLASSE

IARIA

Table of Contents

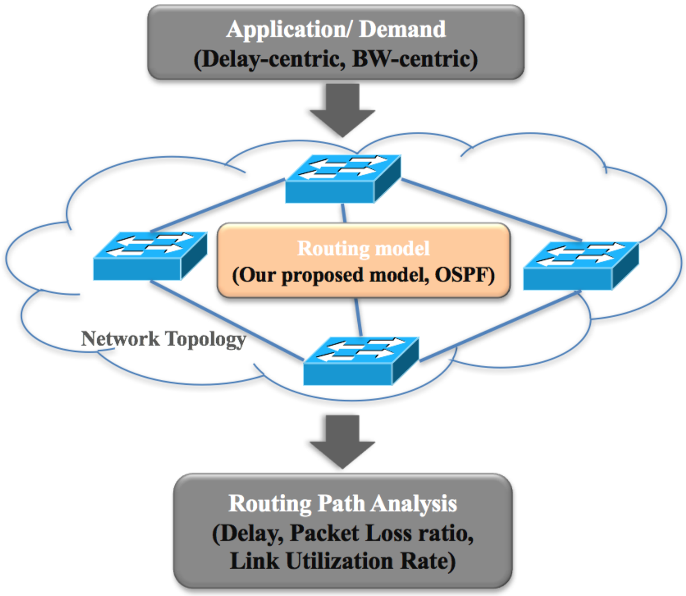
Experimental Methodology

Methodology:
Mixed Integer Linear Programming

Application cost metrics:

Application class	Cost metric
Delay-centric	Delay, Packet Loss Ratio
BW-centric	BW, Packet Loss Ratio

Link utilization rate limit: 75%



Experiment approach

Experiment scenario

Single-demand and multi-demand tests:

	Single-demand	
Topology A	Delay-centric	3
	BW-centric	4
Topology B	Delay-centric	3
	BW-centric	4
Topology C	Delay-centric	4
	BW-centric	4

	Multi-demand	
Topology A	Test 1	1 Delay-centric, 1 BW-centric
	Test 2	2 Delay-centric, 1 BW-centric
	Test 3	1 Delay-centric, 2 BW-centric
Topology B	Test 1	2 Delay-centric, 1 BW-centric
	Test 2	2 Delay-centric, 2 BW-centric
	Test 3	2 Delay-centric, 3 BW-centric
Topology D	Test 1	2 Delay-centric, 1 BW-centric
	Test 2	3 Delay-centric, 2 BW-centric
	Test 3	3 Delay-centric, 3 BW-centric

Result: Delay and packet loss rate for delay-centric application

		Single-Demand			
		Our model		OSPF	
		Delay (ms)	Loss rate (%)	Delay (ms)	Loss rate (%)
Topology-A	Delay-centric 1	0.06	3	0.07	4
	Delay-centric 2	0.03	2	0.06	3
	Delay-centric 3	0.05	4	0.08	5
Topology-B	Delay-centric 1	0.04	5	0.06	5
	Delay-centric 2	0.02	2	0.04	3
	Delay-centric 2	0.06	5	0.08	7
Topology-C	Delay-centric 1	0.03	2	0.1	6
	Delay-centric 2	0.05	4	0.11	4
	Delay-centric 3	0.03	1	0.07	3
	Delay-centric 4	0.03	3	0.08	5

POLYTECHNIQUE MONTRÉAL
LE GÉNIE EN PREMIÈRE CLASSE

IARIA

Result: Delay and packet loss rate for delay-centric application

			Multiple-Demand			
			Our model		OSPF	
			Delay (ms)	Loss rate (%)	Delay (ms)	Loss rate (%)
Topology A	Test1	Delay-centric 1	0.03	2	0.06	3
	Test2	Delay-centric 1	0.03	2	0.06	3
		Delay-centric 2	0.05	4	0.08	5
	Test3	Delay-centric 1	0.03	2	0.06	3
Topology B	Test 1	Delay-centric 1	0.04	5	0.06	6
		Delay-centric 1	0.03	2	0.05	3
	Test2	Delay-centric 1	0.04	5	0.06	6
		Delay-centric 2	0.03	2	0.05	3
	Test3	Delay-centric 1	0.04	5	0.06	6
		Delay-centric 2	0.03	2	0.05	3
Topology D	Test1	Delay-centric 1	0.04	4	0.08	8
		Delay-centric 2	0.04	4	0.06	6
	Test2	Delay-centric 1	0.04	4	0.08	8
		Delay-centric 2	0.04	4	0.06	6
		Delay-centric 3	0.02	2	0.04	4
	Test3	Delay-centric 1	0.04	4	0.08	8
		Delay-centric 2	0.04	4	0.07	7
		Delay-centric 3	0.02	2	0.04	4

POLYTECHNIQUE MONTRÉAL
LE GÉNIE EN PREMIÈRE CLASSE

IARIA

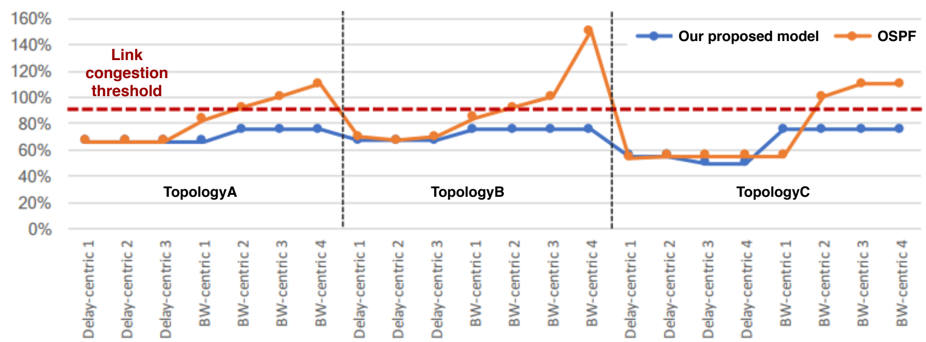
Result: delay and packet loss rate for bandwidth-centric application

		Single-Demand			
		Our model		OSPF	
		Delay	Packet lost	Delay	Packet loss
Topology A	BW-centric1	0.05	4%	0.08	5%
	BW-centric2	0.5	4%	0.08	5%
	BW-centric3	0.07	4%	0.08	5%
	BW-centric4	0.07	4%	0.08	5%
Topology B	BW-centric1	0.04	3%	0.06	4%
	BW-centric2	0.06	3%	0.07	3%
	BW-centric3	0.06	3.5%	0.07	3%
	BW-centric4	0.06	3.5%	0.06	6%
Topology C	BW-centric1	0.06	3	0.06	3%
	BW-centric2	0.03	2.5	0.1	6%
	BW-centric3	0.08	2.8	0.1	6%
	BW-centric4	0.08	3.5	0.1	6%

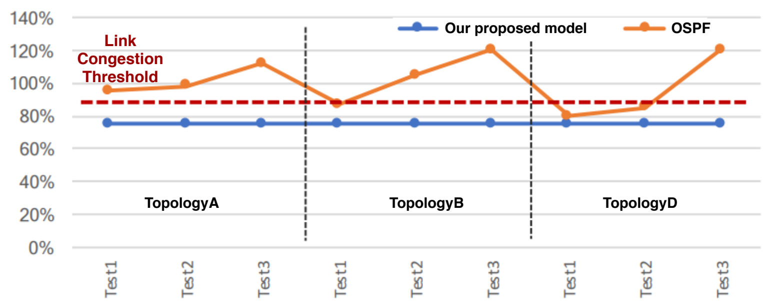
Result: Delay and packet loss rate for bandwidth-centric application

			Multiple-Demand			
			Our model		OSPF	
			Delay (ms)	Loss rate (%)	Delay (ms)	Loss rate (%)
Topology A	Test1	BW-centric 1	0.05	4	0.08	5
	Test2	BW-centric 1	0.05	4	0.08	5
		BW-centric 2	0.06	3	0.07	4
	Test3	BW-centric 1	0.06	3	0.07	4
Topology B	Test 1	BW-centric 1	0.03	3	0.07	4
	Test2	BW-centric 1	0.03	3	0.07	4
		BW-centric 2	0.07	6	0.08	8
	Test3	BW-centric 1	0.03	3	0.07	4
		BW-centric 2	0.07	6	0.08	8
		BW-centric 3	0.02	2	0.02	2
Topology D	Test1	BW-centric 1	0.06	5	0.07	7
	Test2	BW-centric 1	0.06	5	0.07	7
		BW-centric 2	0.08	4	0.08	8
	Test3	BW-centric 1	0.06	5	0.07	7
		BW-centric 2	0.08	4	0.08	8
		BW-centric3	0.04	5	0.02	3

Result: Network throughput



Maximum link utilization across network links in single-demand scenario



Maximum link utilization across network links in multi-demand scenario

Table of Contents

Result analysis

- ✓ Proposed model provides optimized routing paths in terms of the delay and packet loss for the delay-centric demand compared to the OSPF.
- ✓ Proposed model enhances the performance and efficiency of mission-critical and delay intolerant IoT applications
- ✓ Proposed model offers the acceptable level of delay expected by BW-centric application depending on the network status at any given time
- ✓ it can be seen that our model keeps the link utilization rate stable with the maximum 75% while respecting to the other constraints. On the contrary, we could see that the maximum link utilization rate across the network with the same configuration exceeds 100% in OSPF leading to congestion. Demands passing through the congested link could not be served as desired and they may suffer from more delay and loss.
- ✓ In our model takes multi-path approach when multiple BW-intensive demand requests for data transfer service. Also the model seeks to direct the flow toward the links with the higher available bandwidth which cost less.

Table of Contents

Work summary and contribution:

- ✓ This work models a flexible and adaptive common QoS support framework for IoT applications which enforces the application QoS preferences in resource allocation across IoT networking segments.
- ✓ SDN-based middleware in the networking layer makes our model dynamically aware of the network status and of SLA-related application QoS needs. This information is directly applied to the resource allocation process which leads to improving application experiences.
- ✓ Model gives flexibility to the application to define their sensitivity to delay or bandwidth as well as customized their QoS needs and this input is considered in defining link cost metrics per each application.
- ✓ Results shows that this new model provide less delay paths for mission-critical applications enforcing their effectiveness, although enhancing network throughput and minimizing the congestion probability. This could provide more suitable system for mission-critical IoT applications in which the congestion costs a lot.

Future work:

- ✓ We aim to test the performance of proposed algorithm in real IoT environment and enhance our model considering different routing algorithms in different network situations and application classes.

References

[1] A. Haidine, S. El Hassani, A. Aqqal, and A. El Hannani, "The role of communication technologies in building future smart cities," in *Smart Cities Technologies*. InTech, 2016.

[2] ONF, "Openflow switch specification (version 1.5.1)," 2015. [Online]. Available: https://www.opennetworking.org/wp-content/uploads/2014/10/openflow-switch-v1.5.1.pdf

[3] L. Li, S. Li, and S. Zhao, "Qos-aware scheduling of services-oriented internet of things," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1497–1505, 2014.

[4] W. B. Heinzelman, A. L. Murphy, H. S. Carvalho, and M. A. Perillo, "Middleware to support sensor network applications," *IEEE network*, vol. 18, no. 1, pp. 6–14, 2004.

[5] Y. Jararweh, M. Al-Ayyoub, E. Benkhelifa, M. Vouk, A. Rindos *et al.*, "Sdiot: a software defined based internet of things framework," *Journal of Ambient Intelligence and Humanized Computing*, vol. 6, no. 4, pp. 453–461, 2015.