1-1-2019

[Accepted Article Manuscript Version (Postprint)] Distributed Data-Gathering and -Processing in Smart Cities: An Information-Centric Approach

Reza Tourani
New Mexico State University

Abderrahmen Mtibaa
New Mexico State University

Satyajayant Misra
New Mexico State University

Research online Publishing www.ronpub.com

Open Access

Open Journal of Internet of Things (OJIOT)
Volume 5, Issue 1, 2019
http://www.ronpub.com/ojiot
ISSN 2364-7108

Distributed Data-Gathering and -Processing in Smart Cities: An Information-Centric Approach

Reza Tourani [A], Abderrahmen Mtibaa [B], Satyajayant Misra [B]

[A] Saint Louis University, 1 N Grand Blvd, St. Louis, MO 63103, USA, reza.tourani@slu.edu
[B] New Mexico State University, 1780 E University Ave, Las Cruces, NM 88003, USA, {amtibaa,misra}@cs.nmsu.edu

Abstract

*The technological advancements along with the proliferation of smart and connected devices (things) motivated the exploration of the creation of smart cities aimed at improving the quality of life, economic growth, and efficient resource utilization. Some recent initiatives defined a smart city network as the interconnection of the existing independent and heterogeneous networks and the infrastructure. However, considering the heterogeneity of the devices, communication technologies, network protocols, and platforms the interoperability of these networks is a challenge requiring more attention. In this paper, we propose the design of a novel Information-Centric Smart City architecture (**iSmart**), focusing on the demand of the future applications, such as efficient machine-to-machine communication, low latency computation offloading, large data communication requirements, and advanced security. In designing iSmart, we use the Named-Data Networking (NDN) architecture as the underlying communication substrate to promote semantics-based communication and achieve seamless compute/data sharing.*

Type of Paper and Keywords

Visionary paper: *data-centric, Information-Centric, heterogeneity, NDN, IoT, smart cities, architecture*

1 Introduction

Recently, smart cities have emerged as an exciting new concept to interconnect multiple networks of sensors, actuators, and other IoT devices in a city, in order to better address real-life and increasingly complex urban needs. While smart cities composition may vary from a city to another, most include networks, such as smart transportation system, smart healthcare, and smart grid as illustrated in Figure 1 – a network with a diverse set of stakeholders. The United Nations has predicted a constant increase of urbanization, projecting that 68% of the worlds population will live in urban areas by 2050 [25]. According to this report, by 2030, the world will have 43 megacities with more than 10 million inhabitants, needing efficient management for sustainable urban growth.

The major challenge in building a smart city will be interconnecting diverse networks, which are heterogeneous in terms of devices, network protocols, policies, and/or platform incompatibility. Current approaches propose the use of a data management layer [20] as a centralized entity in the network which

This paper is accepted at the *International Workshop on Very Large Internet of Things (VLIoT 2019)* in conjunction with the VLDB 2019 conference in Los Angeles, USA. The proceedings of VLIoT@VLDB 2019 are published in the Open Journal of Internet of Things (OJIOT) as special issue.

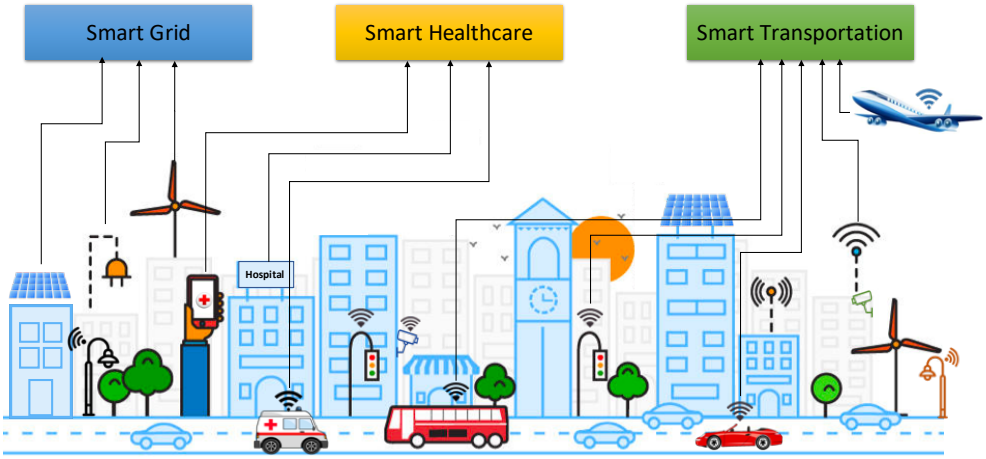

Figure 1: Sample components constituting a smart city (noninclusive list)

gathers, organizes, analyzes, and stores data from different sources for intelligent decision making. Such a data management entity, as a major component (*i.e.,* the brain) of the smart city network, is often slow and inefficient as it deals with heterogeneous data sources, protocols, and systems operating at a very high data rate.

We propose to rethink the current centralized paradigm, by making informed in-network decisions in a fully distributed fashion. We argue that the Named Data networking (NDN) architecture, a realization of the novel Information-Centric Networking (ICN) paradigm, can be leveraged beyond its data centric design to promote seamless and efficient in-network sharing of data and compute resources. We believe that NDN's features, such as data naming, pervasive caching, in-network processing, and built-in security can be used to efficiently address the needs of smart city networks. In particular, NDN allows resources (storage, compute) to be shared using names, which opens opportunities for seamless and easy deployment of inter-connected autonomous network of IoT devices.

In this paper, we propose an Information-Centric Smart City (*iSmart*) architecture that uses NDN as its communication substrate to offer a data- and service-centric framework to all smart cities' actors. For instance, *iSmart* allows efficient gathering, processing, and analysis of different sources of information, for instance, regarding an accident on the road (*i.e.,* camera feeds, vehicle reports, pedestrians field of views, etc.). It further offers easy access to such heterogeneous information for various actors, such as a police car, ambulances, or vehicles on the road.

In designing *iSmart* we keep three objectives in mind: *(i)* efficient and seamless intra- and inter-domain stakeholders communications, *(ii)* in-network processing and efficient use of resources for the next generation applications and services, and *(iii)* seamless and easy-to-use security features. To this end, we will elaborate on *iSmart*'s design features that helps achieve these objectives. We will further discuss the open challenges for *iSmart*'s successful deployment and opportunities that it offers.

The remainder of the paper is organized as follows. In Section 2, we provide an NDN primer and present related work on smart cities architectures using TCP/IP and NDN. In Section 3, we introduce our *iSmart* architecture and its main components. We emphasize in the next three sections the benefits of *iSmart*, specifically with data communication (Section 4), compute sharing and re-use (Section 5), and security measurements (Section 6). We present potential challenges and open issues of *iSmart*'s design in Section 7. Section 8 concludes the paper.

2 Background and Related Work

2.1 Named-Data Networking

The Information-Centric Networking (ICN) is a novel paradigm that shifts the existing "host-centric" communication model to a "data-centric" paradigm, in which data is the primary asset. Named-Data Networking (NDN) is the most common ICN realization. In contrast to IP networks that use IP addresses to identify the data source/destination, the fundamental idea of the NDN architecture [12, 6, 29] is unique content naming, pervasive caching, and name-based routing. NDN is designed as a pull-based communication architecture with an inherent flow control; each *Interest* (NDN request) packet elicits a piece of data. This data-centric paradigm aligns network operations with the consumer/producer nature of the IoT applications, which makes it a perfect fit for IoT

networks.

In NDN, routers are equipped with a content store (CS), a pending interest table (PIT), and a forwarding information base (FIB). The FIB (similar to the forwarding table in IP routers) gets populated using a routing algorithm. Any node that receives an Interest packet for a data chunk performs a CS lookup on the content name. If the content is not available in the CS, the router performs a lookup in its PIT to check whether there is an existing entry for the requested content. If the PIT lookup is successful, the router adds the incoming Interest's interface to the PIT entry (Interest aggregation) and drops the interest. Otherwise, the router creates a new PIT entry for the Interest and forwards it using the FIB to an upstream router in the direction of the data source(s).

An Interest can be satisfied either by any intermediate forwarding router which has cached the corresponding content chunk, or the content provider. In both cases, the content takes the interest's reverse-path back to the requester. Upon receipt of a content chunk, a router forwards the chunk along the interfaces on which it had received the corresponding Interest(s). The router may also cache a copy of the content in its CS for subsequent Interests. An important facet of NDN is its strategy layer, which allows a node to leverage cross-layer information to make smart and forwarding decisions with high granularity.

2.2 Smart City Architecture

With large scale data generation, sensors, actuators, consumers, and stakeholders the smart city scenario has been a broad area of research with focus so far mainly on providing a cross-platform, multi-stakeholders platform to enable efficient use and sharing of resources [7, 2]. Research designs [20], frameworks [7], and implementations [14] have been motivated by the projected increase in smart city's population and the number of devices and have predominantly focused on hardware, application, and services advancements. Solutions include: (i) the use of new and different information and communication technologies [9, 21]; (ii) monitoring, controlling, and managing the resources (*e.g.,* electric power [8]); and (iii) the real-world deployment and feasibility studies of new smart cities.

We, in this paper, explore the other dimension by rethinking the underlying networking architecture design and assessing what it can potentially add to existing and future smart city applications and services.

Only a few recent initiatives leveraged NDN for smart cities [15, 5, 10]. At the core, these approaches focused on the communication aspects of smart cities by leveraging NDN's components [5], design and orchestration of smart services in an NDN-enabled platform [15], and secure onboarding and routing [10] for large scale deployment of devices in a smart city network. However, these work fall short in answering an important architectural question–how to build a holistic smart city network by integrating the existing heterogeneous and self-governed networks? Most relevant to our work, NDN has been discussed as a potential architectural solution to IoT networks [19], however this preliminary work does not focus on multiple issues, such as scalability, resiliency, and security onboarding. In this paper, we argue that using NDN as the underlying substrate allows stitching the existing independent networks to build an information-centric smart city network, which will be much more difficult to build using the current IP-based paradigm.

3 An Information-Centric Design for Smart City Architecture

In this section, we elaborate on *iSmart*–our information-centric smart city architecture–which creates a smart city network comprised of independently-governed IoT/CPS networks. In what follows, we assume that each of these networks, such as smart grid, is running independent of others with its services hosted on the cloud or their private hosts. In other words, we consider a service-oriented architecture, where IoT devices interact with a server(s) for delivering their sensory data and receiving command and control messages.

3.1 System Model

In what follows, we describe *iSmart*'s major components.

Autonomous Network (AN): ANs are the self-governed proprietary networks, such as automated power grids, smart homes, intelligent transport systems, or campus/enterprise networks, which do not share their information with other networks. As shown in Figure 2, we consider three different sizes of ANs: (i) small ANs such as smart homes (shown as the top layer of Figure 2); (ii) mid-size ANs, such as university campuses, and multinational companies, water management system (shown as the middle layer of Figure 2); and (iii) large-size ANs, such as city-wide smart grids and intelligent transport systems (shown as the bottom layer of Figure 2). For the intra-AN communication, each AN is equipped with a local highly available authentication service (AS) that provides basic security services, including identity, key, trust, and access control management. As for the inter-AN communication, which is the missing component of the existing smart city

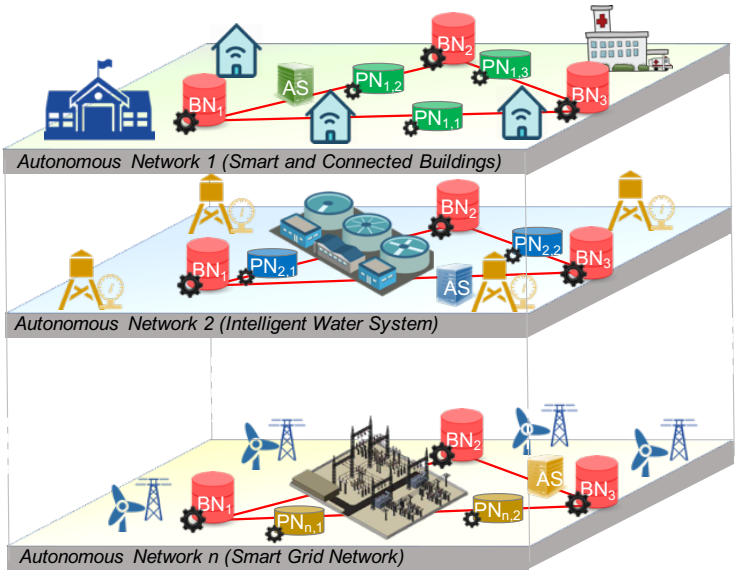

Figure 2: *iSmart* multi-layer architecture with superposed AN placement

proposals, we augment each AN with multiple Peripheral Nodes (PNs) as we describe below.

Peripheral Node (PN): PNs act as ANs' proxies to the outside world. As illustrated in Figure 2, we envision each AN_i to have $k \geq 1$ peripheral nodes, $PN_{i,j}$, where $j = 1 \ldots k$ to allow network resiliency by providing multiple paths between ANs. PNs have three main roles: *(i)* managing the access to the corresponding AN (the IoT networks within the AN) by allowing only authorized requests to enter the AN; *(ii)* caching content as a means for eliminating the single point of failure and reducing the intra-AN traffic; and *(iii)* performing basic edge computing services, such as context extraction from crowdsourcing applications and multi-source video analytic. The PNs can also aggregate data obtained from the IoT nodes in the AN to create aggregated results as well as answer queries.

Backbone Nodes (BN): BNs are dedicated backbone nodes that create a backbone network connecting the ANs. BNs are generally deployed by an authority such as the city's municipality. The placement of these nodes can be planned to create a resilient network that enhances connectivity and can handle nodes and links failures. These backbone nodes operate as NDN nodes as well as edge computing nodes. They can use their caches to reduce content retrieval delay and reduce network traffic volumes.

Edge Computing (EC): EC is a distributed computation model, in which storage and computation resources are deployed in the proximity of the users. In the context of *iSmart*'s architecture, BNs and PNs form the distribute edge network and provide resources for running customized data and compute intensive applications, such as video analytic and knowledge extraction from massive data-sets. They can also provide edge services to end-users' (*e.g.,* smartphones or smart vehicles) applications (*e.g.,* , image annotation, answers to local queries).

3.2 Architecture Overview

One of the major merits of *iSmart*, depicted in Figure 2, is promoting seamless and secure collaboration between several autonomous networks (ANs)–currently, these networks operate independently without sharing information. We argue that the interconnection of ANs and its subsequent advantages, such as information sharing, intelligent decision making, etc., are the key to creating a pervasive smart city network to improve life quality, fuel economic development, gain efficiency in utilizing resources, and introduce positive impacts on environment.

iSmart consists of: *(i)* a diverse set of autonomous networks (ANs) containing IoT devices (subnets) including their independent authentication servers (ASs) for arbitrating access to the subnets and their data (each layer in Figure 2 corresponds to an AN); *(ii)* multiple peripheral nodes (PNs) that are provisioned at the periphery of the ANs, which are equipped with edge computing capabilities; and *(iii)* a set of backbone nodes (BNs), installed in the city infrastructure, scattered around the city to maintain connectivity between ANs and provide a resilient edge computing environment. The superposed placement of BNs in all layers of

Figure 2 represents their geographic locations.

We argue that our proposed ICN-based smart city architecture offers features, which can be leveraged to provide: (*i*) *Semantics-based Communication*: leveraging NDN's semantic naming and pervasive caching for efficient one-to-many, many-to-one, and many-to-many communications in a smart city network; (*ii*) *Edge-based Data and Compute Sharing*: using NDN's in-network processing to promote re-use of data and computation across entities for efficient resource utilization; and (*iii*) Security and Resiliency: utilizing NDN's inherent data integrity and trust assessment, via digital signatures and the NDN trust schema [27], to facilitate private and secure information sharing.

4 Semantic-based Communication

In smart cities, various IoT networks are used to provide connectivity among billions of devices for different use-cases and applications, such as intelligent transportation system, smart healthcare, and smart grid (refer to Figure 1). Often, these applications require more sophisticated types of communication, namely: many-to-one, many-to-many, and one-to-many. These types of communication, in today's IoT networks, are achieved by sending multiple unicast packets to each source for polling data (many-to-one), maintaining multicast trees rooted at each source (one-to-many and many-to-many), or custom routing protocols like RPL [26].

Although feasible, these approaches either requires abundant resources at intermediate routers for maintaining numerous multicast trees or incur high communication overhead for routing purposes [4]–a prohibitive approach for large-scale IoT deployments with resource constrained devices. In what follows, we will discuss how NDN addresses the IP limitations in providing one-to-many, many-to-one, and many-to-many communication scenarios (we will refer to them as *X2X* for brevity) in the context of data gathering and device management.

4.1 Data Gathering

In contrast to IP networks, NDN inherently supports *X2X* communication models through naming semantics, pervasive caching, request aggregation, and its stateful forwarding plain. Content naming, as the most fundamental NDN feature, provides the blueprint for consumers' (end-users') to obtain knowledge of the named data and facilitates deployment of the security, provenance, and access control mechanisms by binding data names to their corresponding key names (e.g., name-based access control [28]). The semantics from data names, in conjunction with NDN's stateful forwarding plane, augment the network layer with refined information on application layer logic and the data characteristic, allowing the network to make more informed forwarding decisions.

NDN's pervasive caching (combined with cache replacement) replicates the currently published, popular data across the network, promoting data multi-homing, which provides low latency communication and data resiliency. Furthermore, NDN's request aggregation minimizes the network traffic by eliminating redundant data delivery. All the independent requests for a single data chunk arriving at a forwarding router will be aggregated with only the first one being forwarded. This results in only one copy of the requested data traversing the network upto the aggregating router(s), thus significantly reducing network load. The integration of request aggregation with NDN's stateful forwarding plane helps perform multicast based data distribution without the need for the creation of multicast trees or significant state maintenance at the routers.

For better explanation, let's consider an intelligent transportation system with the network control center managing and monitoring several thousands of autonomous vehicles, traffic lights, speed cameras, CCTVs, etc., which introduces a broad range of tenants across the city. For this use-case, we use a hierarchical naming convention (similar to human-readable URL naming) for devices, leveraging devices' locations and types. For instance, a *traffic light* in *street A* at the intersection of *street B* of *city Z* will be named: "`/city_Z/street_A/intersection(A_B) /traffic_light/node_id`". Note that this convention is illustrative and more compact names can be used. Leveraging this naming convention, one can interact with this traffic light (one-to-one) by sending an Interest including the given name. A more thought-provoking scenario is the many-to-one scenario, in which the current status of all the traffic lights in *street A* of *city Z* is needed. Here, the control center can send an Interest to the network using the name "`/city_Z/street_A /traffic_light/STATUS/*`" name.

This tells each router receiving the Interest to forward it along all of its outgoing interfaces. Thus, the NDN's name-based routing forwards this Interest to all the traffic lights for the given location, resulting in multiple data packet to be delivered to the requester. To reduce data volume, an intermediate router can aggregate the data from multiple traffic lights and send only one packet in response to the interest. This is only possible as the routers can understand the content in the packets they forward. Mechanisms can be deployed on the routers to perform data aggregation while preserving the provenance information corresponding to each lights, so that they can be verified at the control center.

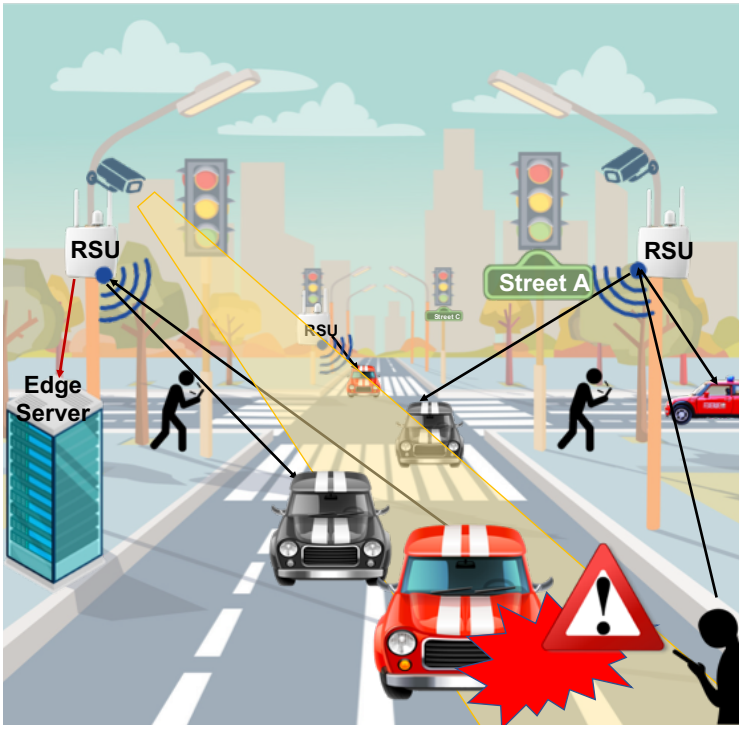

Figure 3: An illustrative smart city use-case: A car accident is automatically reported by explicit messaging from the vehicles involved in the accident, other neighboring vehicles, the CCTV cameras, and potential video feeds from pedestrians. Interested entities, e.g., police cars and ambulance are given role-based controlled access to the aggregated data and the analytics.

4.2 Device Management and Provisioning

A smart city network is expected to be large and highly heterogeneous in terms of devices, technologies, applications, software, and stakeholders. To handle such scale and heterogeneity the existing IP-based solutions require multiple complex mappings at the application, transport, and network layers to provide abstractions–this does not scale with the fast growth of devices. To remedy this shortcoming, NDN's naming can be used to provide an abstraction for applications and devices and hide such heterogeneity. The rich semantic of NDN naming can use device's properties including its vendor, functionality, model, and production year for naming to simplify the interactions between the control center and a class of devices with similar features for management, actuation, and over-the-air updates.

In what follows, we elaborate on the benefit of NDN's naming using the well-known Mirai botnet attack and our intelligent transportation scenario as the driving examples. The authors in [1] discovered that the Mirai botnet consisted of $600,000$ compromised IoT devices–with the majority of those being IP cameras, DVRs, and consumer routers sharing similar properties. With the existing IP-based architecture, it is almost impossible to track such a large number of devices by their identities, whether MAC or IP addresses [16]. In contrast, leveraging the NDN naming allows the authorities in each AN to access and update the firmware of the compromised CCTV cameras (produced by *XiongMai* company) using the Interest "`/XiongMai/urgent/firmware_update/CCTV/*`". Given the provenance capability of NDN, such an update can be efficiently globally orchestrated by the company itself. Evidently, such firmware update will require more detailed interactions between the control center (or company) and the devices, which is out of the scope of this paper.

In the intelligent transportation system case (Figure 3) with a large number of deployed IoT devices ranging from in-vehicle sensors to infrastructure-based devices, such as cameras, traffic lights, ultrasonic, and CO_2 emissions sensors. Figure 3 shows the scenario in which an accident has happened on *street A*. Following the accident report, the network control center can request the live feed of all the *CCTV cameras* located in *street A* of *city Z* using "`/city_Z/street_A/LIVE_FEED/camera/*`" name. On receiving the request, the cameras in *street A* return their feed back to the center via the deployed

RSUs. We note that the operation center can interact with particular cameras with the best angle if their names are known, as discussed in the previous subsection.

5 Edge-based Data and Compute Sharing

Edge computing often relies on centralized command and control entities, *e.g.,* a network controller, which gather statistics on network and compute loads and make decisions in task dispatch, load balancing, and service monitoring. However, such a centralized design is not resilient and is prone to failures. In fact, it is well-documented in the software-defined networking literature that if the network controller fails, the operation of the entire edge network might fail or become severely sub-optimal.

In this section, we will discuss how names can be exploited to facilitate in-network data and compute sharing without the need for a centralized controller.

5.1 Data Sharing

In conventional IP architectures, context, locations, and status messages need to be stored and mapped to physical IP devices, which may change due to mobility or dynamic host configurations. This complex mapping and the consequent indirections introduce delays, inefficient use of resources, and increase failure probability which can be fatal for smart city scenarios, such as autonomous driving where content and command delivery can be very time sensitive ($<$ 10 ms latency).

In NDN, Interest and Data packets use semantic naming, which can clearly define the context, can be leveraged to guide forwarding, security, and application requirements without mapping these names to physical machines. Such context-aware naming allows data sharing and collaboration among different stakeholders across multiple domains.

For instance, in the smart city scenario (Figure 3), an accident notification message can be expressed via "`/accident/city_Z/street_A/warning/timestamp`". Other vehicles or pedestrians can also share additional information such as video feeds of the accident, "`/accident/video/city_Z/street_A/timestamp`". The diverse received information are processed by the Road Side Units (RSUs) and edge computing servers. The use of semantic names, in this example, makes it easy for data to be accessed by multiple entities to provide easy access or to request follow-up content without the need to understand the content.

Data from different entities reporting the same event can also be aggregated efficiently due to semantic naming (overlapping names imply overlapping content). Sensed and analyzed data can be easily accessed by other vehicles, insurances, ambulances, etc. Today, due to the non-scalability of IP and the corresponding silo-ed nature of communication sharing of data at the edges is difficult. With mandated device-to-device communications in 5G, such sharing will become prevalent at the edge.

Moreover, context-aware caching strategies can be employed to allow fast and prompt access to information when they are requested (*e.g.,* accident and local weather warnings are proactively cached in the neighboring localities), and thus can reduce overall network backbone traffic.

5.2 Compute Reuse

In smart cities, the scale of data gathering, communication, and task execution at the edge can result in network congestion and/or the overloading of the compute nodes. Often in such compute intensive environment, end users request execution of "similar" tasks, but the way edge computing is setup rather than re-using existing computation results the computations are performed repeatedly with little reuse. This is particularly wasteful with tasks that share a given context, location, or objectives such as annotating a live scene in a stadium or a museum where multiple mobile users would require annotation of similar scenes or request information about similar events.

Leveraging compute re-use in current IP-based systems can be challenging as the metadata which can be used to determine task similarity is only available at the application layer. In NDN, context-aware naming can help with compute reuse and with the elimination of redundant computation. For instance, with efficient namespace design tourists can retrieve pre-annotated touristic scenes without any effective computation. Scenes can have semantic names which include their GPS coordinates, gyroscope data, and camera directions, *e.g.,* "`/city_Z/street_A/X;Y;Z/NE/Camera1/timestamp`" referring to a field-of-view (FoV) from node "`camera1`", at time "`timestamp`", describing a scene in city "`city_Z`", street "`street_A`", where cameras are facing the northeast "`NE`". Edge computing nodes can use the name to filter and classify the content to speed up the search in its pool of precomputed tasks.

6 Security and Resiliency

In this section, we elaborate on NDN's features that help improve the security and communication resiliency of smart city networks.

6.1 Information and Network Security

In this subsection, we review the four major security requirements of smart city: integrity and provenance, access control enforcement, Distributed Denial of Service (DDoS) resiliency, and trust management.

Integrity and Provenance: NDN's built-in security mechanisms provide data-oriented security, in which data authenticity and integrity propagate with the data itself, rather than being a feature of the communication channels (*e.g.,* OpenSSL in today's applications). In NDN, providers are mandated to sign each chunk of the their generated information before publication. This data-oriented security model is in contrast with the IP networks security, in which trusting data validity and authenticity needs realtime interaction with a server. NDN's built-in security is the key enabler of its pervasive caching–allowing the closest data replica to be retrieved without violating the expected trust. In fact, integrating security in data rather than hosts, promotes the underlying network from a primitive delivery substrate to an intelligent participant.

Access Control: Serving the data from in-network caches, however, raises a security issue–the content owners lose control of their published content. To address this concern and achieve confidentiality, approaches based-on broadcast encryption, proxy re-encryption, and attribute-based encryption have been proposed, allowing efficient cache utilization of encrypted content [11, 23]. Among all, we envision the access control frameworks with authentication and authorization delegation to be more viable due to their distributed nature. Thus, we adopt TACTIC [24], to promote access control delegation to the semi-trusted edge infrastructure. TACTIC's edge deployment effectively enforces the data access policy for the cached content and has been shown to scale with the network's size with low communication and computation overhead.

The multi-stakeholders nature of communication requires engagement of multiple domains in delegating their access enforcement to multiple stakeholders. To accommodate such a need and to protect users security and privacy, we envision the deployment and orchestration of access control as a virtualized network function (VNF) at the edge–a case for access control-as-a-service.

DDoS Resiliency: Another advantage of the NDN architecture, compared to the IP architecture, is its flexibility in handling DDoS attacks. In the IP networks, arrays of compromised IoT devices are being used to orchestrate massive DDoS attacks, such as Mirai and Lizard Botnet attacks on *Dyn DNS* and the *Rio Olympic*. The contemporary mitigation techniques redirect the suspicious traffic across the Internet to scrubbing centers for malicious traffic filtering. The scrubbing centers then return the clean traffic to the network for destination delivery–a costly practice, which results in network congestion and significant overhead in terms of computation and communication.

In contrast, the NDN's fundamentals including its pull-based model and customizable strategy layer allow the distribution of the attack detection and mitigation load to the edge servers, closer to the attack sources. Deploying the defense mechanisms closer to the attack source, prevents the malicious traffic from entering the core network and reduces the communication overhead. In this regard, approaches similar to TACTIC, which authenticate the network ingress traffic at the edge can be leveraged to achieve DDoS resiliency.

Distributed Trust Management: Considering the distributed nature of our design– multiple autonomous ANs with their own trust roots–we envision a distributed trust management system based on reputations (for the services and nodes) by the integration of emerging technologies such as distributed ledger. Such a system should provide unimpeachable and reliable services even in the presence of compromised infrastructure. Thus, the trust related information of the network, services, and devices, such as the certificate revocation lists, users' digital certificates, and trust roots will be recorded in the ledger. For the intra-AN identification and trusted communication, all PNs and IoT devices use their existing certificates. To advance scalability, we envision an AN's devices to be loaded with the required credentials to avoid interaction with the ledger. For building the outward facing trust (inter-AN communication), the trust information of the entities should be retrieved from the ledger, allowing the parties to validate each other in a distributed manner.

6.2 Communication Resiliency

In the emerging cyber-physical system deployments, nodes are equipped with multiple radio access technologies (multi-RAT), such as WiFi, cellular, and ZigBee radios for better communication resiliency and quality of service (QoS). The current IP-based use of multi-radio access technologies (multi-RAT) aim at using cross-layer information to stripe data traffic across multiple wireless interfaces, thus optimizing concurrent radio technologies and paths for improved application quality. However, such solutions use application level proxies, which result in additional delay for decision-making and sub-optimal performance [17, 18].

In contrast, the connectionless nature of NDN enables the concurrent use of multiple interfaces at the network layer, promoting seamless multicast communication. Such an inherent multicasting

Figure 4: Open challenges and opportunities of *iSmart* design

capability is, in part, due to NDN's strategy layer, allowing customized packet forwarding logics to be implemented in the forwarding plane. Recent work in NDN-based multi-RAT [22, 13] has shown a significant advantage (over 30-40%) in communication latency and resiliency via (i) establishing resilient paths between source and destination pairs [13] and (ii) traffic load balancing [22, 13]. Nonetheless, more sophisticated forwarding strategies can be designed for optimizing the communication latency and establishing resilient communication paths based on the link and network statistics.

7 Challenges and Opportunities

In this section, we draw a general picture of the main challenges that need to be addressed for *iSmart*'s successful deployment in future smart cities, and also discuss the main opportunities enabled by *iSmart*. Figure 4 list few challenges and opportunities, which we will discuss below.

7.1 Challenges

Privacy: As we discussed in Subsection 5.1, NDN's content naming and pervasive caching facilitate data sharing–a common requirement in a range of emerging applications. However, such data sharing can result in privacy violation–a significant concern that is growing with the ever-increasing number of data breach incidents. The *General Data Protection Regulation* (GDPR) act by the European Union is one of the first initiatives aimed at protecting the personal data and privacy of the European citizens. More will follow soon and applications and networks have to account for the corresponding stipulated user privacy requirements.

Security: The proliferation of IoT devices and their adaptation introduces new threats and expands the attack surface. In the context of smart cities, cyber attacks can have devastating consequences: hackers can break into the network and shut down the city's electricity supply, similar to the attack on Ukraine's power grid in December 2015. Thus, more resources should be invested into the security of smart cities to avoid such events.

Regulations: Smart cities may handle data exchange for transnational systems, such as logistics for airlines and commercial transportation. Different states (or nation) may have their own laws and strategies making coordination between states (or countries) a major necessities. Moreover, regulations and standardizations are also needed to identify the responsibilities of each entity (*e.g.,* stakeholders) in case of disputes and assessment of ownership of generated data.

7.2 Opportunities

Connectivity: We argue that *iSmart* is inherently resilient to link and system failures. In fact, *iSmart* provides intelligent stateful routing and multi-interface/path data retrieval, which aids easy and seamless recovery from network and system failures. In the context of disaster recovery (*e.g.,* flood, hurricanes, and earthquakes), most smart city designs require a centralized failure management systems to define

recovery strategies [20]–prone to failure. In such scenarios, *iSmart*'s fully distributed and resilience design offers inherent data and network fault tolerance.

Economics: Deployment of smart cities can bring forth numerous socioeconomic benefits. For instance, automation is often coupled with cost reduction and fewer errors. Investments in smart city programs is considered as one of the most profitable businesses due to the potential economic growth associated with urbanization and population needs [3].

Big Data: *iSmart*'s in-network processing capabilities, enables new areas of research for distributed data analytic. Big data management is not limited to data gathering, but efficient sharing, combining/superposing, summarizing, and analyzing data, which summed together are challenging. *iSmart* can provide an effective network layer that makes such big data applications effective and secure and improves user privacy.

8 Summary and Conclusions

In this paper, we proposed *iSmart*, an information-centric smart city network architecture. *iSmart* employs the emerging NDN architecture as the underlying substrate, which promotes pervasive caching, name-based routing, in-network processing, and built-in security. We elaborate the *iSmart*'s semantic-based communication, which facilitates large data gathering and network management, and its capability in promoting in-network processing, data sharing across multiple domains, and data-centric security, which are essential for scaling to very large IoT networks. Finally, we discuss the generic smart city network challenges that need to be addressed for secure and large scale deployments an opportunities we see with our architecture and with the use of the information-centric paradigm in general.

Acknowledgements

Research supported by NSF awards #1800088; #1719342; #1345232, EPSCoR Cooperative agreement OIA-1757207; ARO grant #W911NF- 07-2-0027, and Intel grant #34627535. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the federal government and other funding agencies.

References

[1] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, and A. Halderman, "Understanding the mirai botnet," in *26th USENIX Security Symposium (USENIX Security 17)*, 2017, pp. 1093–1110.

[2] N. Z. Bawany and J. A. Shamsi, "Smart city architecture: Vision and challenges," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 11, pp. 246–255, 2015.

[3] J. Bélissent, "Getting clever about smart cities: New opportunities require new business models," Cambridge, Massachusetts, USA, Tech. Rep., 2010.

[4] M. Farooq and T. Kunz, "Iot-rf: A routing framework for the internet of things," in *Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2017, pp. 1–7.

[5] S. B. Hussain, H. Ahmed, D. Kim, and H. Song, "Named-data-networking-based its for smart cities," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 105–111, 2017.

[6] V. Jacobson, D. K. Smetters, J. D. Thornton, M. F. Plass, N. H. Briggs, and R. L. Braynard, "Networking named content," in *Proceedings of the 5th international conference on Emerging networking experiments and technologies*. ACM, 2009, pp. 1–12.

[7] J. Jin, J. Gubbi, S. Marusic, and M. Palaniswami, "An information framework for creating a smart city through internet of things," *IEEE Internet of Things journal*, vol. 1, no. 2, pp. 112–121, 2014.

[8] N. Khatavkar, A. Naik, and B. Kadam, "Energy efficient street light controller for smart cities," in *2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS)*. IEEE, 2017, pp. 1–6.

[9] D. Magrin, M. Centenaro, and L. Vangelista, "Performance evaluation of lora networks in a smart city scenario," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–7.

[10] T. Mick, R. Tourani, and S. Misra, "Laser: Lightweight authentication and secured routing for ndn iot in smart cities," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 755–764, 2018.

[11] S. Misra, R. Tourani, F. Natividad, T. Mick, N. Majd, and H. Huang, "Accconf: An access control framework for leveraging in-network cached data in the icn-enabled wireless edge," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 1, pp. 5–17, 2017.

[12] NDN, "Named data networking community," Last accessed 30th May 2019, http://named-data.net/.

[13] G. Panwar, R. Tourani, T. Mick, A. Mtibaa, and S. Misra, "DICE: Dynamic Multi-RAT Selection in the ICN-enabled Wireless Edge," in *Proceedings of the SIGCOMM Workshop on Mobility in the Evolving Internet Architecture (MobiArch)*. ACM, 2017, pp. 31–36.

[14] G. Pasolini, C. Buratti, L. Feltrin, F. Zabini, C. De Castro, R. Verdone, and O. Andrisano, "Smart city pilot projects using lora and ieee802. 15.4 technologies," *Sensors*, vol. 18, no. 4, p. 1118, 2018.

[15] G. Piro, I. Cianci, L. A. Grieco, G. Boggia, and P. Camarda, "Information centric services in smart cities," *Journal of Systems and Software*, vol. 88, pp. 169–188, 2014.

[16] H. Rahbari, J. Liu, and J. Park, "Securematch: Scalable authentication and key relegation for iot using physical-layer techniques," in *Conference on Communications and Network Security (CNS)*. IEEE, 2018, pp. 1–9.

[17] G. Rossini and D. Rossi, "Evaluating ccn multi-path interest forwarding strategies," *Computer Communications*, vol. 36, no. 7, pp. 771–778, 2013.

[18] K. Schneider and U. Krieger, "Beyond network selection: Exploiting access network heterogeneity with named data networking," in *Proceedings of the 2nd ACM Conference on Information-Centric Networking*. ACM, 2015, pp. 137–146.

[19] W. Shang, A. Bannis, T. Liang, Z. Wang, Y. Yu, A. Afanasyev, J. Thompson, J. Burke, B. Zhang, and L. Zhang, "Named data networking of things," in *2016 IEEE First International Conference on Internet-of-Things Design and Implementation (IoTDI)*. IEEE, 2016, pp. 117–128.

[20] B. N. Silva, M. Khan, and K. Han, "Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities," *Sustainable Cities and Society*, vol. 38, pp. 697–713, 2018.

[21] P. Sotres, J. R. Santana, L. Sánchez, J. Lanza, and L. Muñoz, "Practical lessons from the deployment and management of a smart city internet-of-things infrastructure: The smartsantander testbed case," *IEEE Access*, vol. 5, pp. 14 309–14 322, 2017.

[22] R. Tourani, S. Misra, and T. Mick, "Ic-mcn: An architecture for an information-centric mobile converged network," *IEEE Communications Magazine*, vol. 54, no. 9, pp. 43–49, 2016.

[23] R. Tourani, S. Misra, T. Mick, and G. Panwar, "Security, privacy, and access control in information-centric networking: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 566–600, 2017.

[24] R. Tourani, R. Stubbs, and S. Misra, "Tactic: Tag-based access control framework for the information-centric wireless edge networks," in *International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2018, pp. 456–466.

[25] United Nations, "2018 revision of world urbanization prospects," https://population.un.org/wup/, 2018.

[26] T. Winter, P. Thubert, A. Brandt, J. Hui, R. Kelsey, P. Levis, K. Pister, R. Struik, J. Vasseur, and R. Alexander, "Rpl: Ipv6 routing protocol for low-power and lossy networks," 2012. [Online]. Available: https://rfc-editor.org/rfc/rfc6550.txt

[27] Y. Yu, A. Afanasyev, D. Clark, V. Jacobson, and L. Zhang, "Schematizing trust in named data networking," in *Proceedings of the ACM Conference on Information-Centric Networking*, 2015, pp. 177–186.

[28] Y. Yu, A. Afanasyev, and L. Zhang, "Name-based access control," Named Data Networking, Tech. Rep., 2015.

[29] L. Zhang, A. Afanasyev, J. Burke, V. Jacobson, P. Crowley, C. Papadopoulos, L. Wang, B. Zhang *et al.*, "Named data networking," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 66–73, 2014.

AUTHOR BIOGRAPHIES

Reza Tourani, Ph.D., is an assistant professor in Computer Science at Saint Louis University. He completed his M.Sc. and P.hD. in Computer Science from New Mexico State University, Las Cruces, NM, USA, in 2012 and 2018, respectively. His research interests include security and privacy, Internet of Things and CPS, future Internet architecture, Information-Centric Networking, and smart grid architectures and protocols. He has authored more than 20 peer-reviewed IEEE/ACM journal articles and conference proceedings.

Abderrahmen Mtibaa, Ph.D., is an assistant professor at the department of Mathematics and Computer Science in the University of Missouri-St. Louis. Prior to that, he was visiting professor, postdoc at New Mexico State University, Texas A&M University and Carnegie Mellon University. He earned his Computer Science PhD. degree from the Sorbonne University (Paris, France). He received his MS and BS degrees in Computer Science from the University of Manouba (ENSI, Tunisia). His main research interests include, but not limited to, network security, mobile crowd-sensing, ubiquitous computing, mobile networking, and social networking. He published over 35 fundamental peer-reviewed research papers in prestigious venues such IEEE Transactions, IEEE Communication Magazines, IEEE Infocom, and ACM CoNext.

Satyajayant Misra, Ph.D., is an associate professor in Computer Science at New Mexico State University. He completed his M. Sc. in Physics and Information Systems from BITS, Pilani, India in 2003 and his Ph.D. in Computer Science from Arizona State University, Tempe, USA in 2009. His research interests include security, privacy, and resilience in wireless networks, the Internet, IoT/CPS, and supercomputing. He has served on several IEEE journal editorial boards and IEEE/ACM conference executive committees. He has authored more than 80 peer-reviewed IEEE/ACM journal articles and conference proceedings, which have received over 4500 citations. More information can be obtained at www.cs.nmsu.edu/∼misra.