Statistical Dissemination Control in Large Machine-to-Machine Communication Networks

Shih-Chun Lin, *Student Member, IEEE*, Lei Gu, *Member, IEEE*, and Kwang-Cheng Chen, *Fellow, IEEE*

***Abstract*—Cloud based machine-to-machine (M2M) communications have emerged to achieve ubiquitous and autonomous data transportation for future daily life in the cyber-physical world. In light of the need of network characterizations, we analyze the connected M2M network in the machine swarm of geometric random graph topology, including degree distribution, network diameter, and average distance (i.e., hops). Without the need of end-to-end information to escape catastrophic complexity, information dissemination appears an effective way in machine swarm. To fully understand practical data transportation, G/G/1 queuing network model is exploited to obtain average end-to-end delay and maximum achievable system throughput. Furthermore, as real applications may require dependable networking performance across the swarm, quality of service (QoS) along with large network diameter creates a new intellectual challenge. We extend the concept of small-world network to form shortcuts among data aggregators as infrastructure-swarm two-tier heterogeneous network architecture, then leverage the statistical concept of network control instead of precise network optimization, to innovatively achieve QoS guarantees. Simulation results further confirm the proposed heterogeneous network architecture to effectively control delay guarantees in a statistical way and to facilitate a new design paradigm in reliable M2M communications.**

***Index Terms*—Machine-to-machine communications, network topology, small-world networks, information dissemination, quality-of-service guarantees, statistical control, Internet of things, *ad hoc* networks.**

I. Introduction

CLOUD based machine-to-machine (M2M) communications [1], [2] have emerged to enable services through interaction between cyber and physical worlds, achieving ubiquitous and autonomous data transportation among objects and the surrounding environment in our daily lives. The wireless network involving tremendous machines that the availability of end-to-end information at each machine is not possible, is referred to the large M2M network, which is getting importance into next-generation wireless systems [3]. While these tremendous machines have short-range communication capabilities, multi-hop networking is a must for information dissemination over machine swarm. The connectivity and low delivery latency in the machine swarm are consequently crucial to achieve reliable communications. However, lacking complete understanding of large network characteristics, effective traffic control for message delivery remains open [4]. As a result, a proper control scheme of routing with quality-of-service (QoS) guarantee regarding end-to-end delay becomes an urgent need to practically facilitate M2M communications. This is even more challenging due to the scalability of multi-hop *ad hoc* networks and energy-efficient and spectral efficient operation for each machine [4]–[6].

To investigate the routing mechanism for large-scale networks, network topology analysis can be scientifically exploited by random network analysis [7]–[21]. Newman [7] provides a comprehensive study in network structure and functions from complex networks perspective. Aiming at social communities mediated by network technologies, [8] reviews the historical research for community analysis and community discovery methods in social media. While Gjoka *et al.* [9] develop an unbiased sampling for users in an online social network by crawling the social graph, they further examine multiple underlying relations for such network in [10] to introduce a random walk sampling. For social networks related research, [11] proposes the information-centric networking as it brings the advantages to the network operator and the end users. Exploring various research challenges in context management, [12] presents a context management architecture that is suitable for social networking systems enhanced with pervasive features. Through a survey of current routing solutions, [13] discuss the trend toward social based routing protocols, which are classified by employed network graph.

In addition, to employ social network analysis in message delivery, Kleinberg [14] remarkably pioneers the methodology to exercise the small-world phenomenon [22] of social networks in navigation, successfully creating transmissions with less delay. Small-world phenomenon plays a crucial role in social networks, which states that each individual in such network links to others by a short chain of acquaintances and has great potential for improving spectral and energy efficiency for shorting the end-to-end delay. Reference [15] also presents

Manuscript received April 6, 2014; revised September 6, 2014; accepted November 17, 2014. Date of publication December 2, 2014; date of current version April 7, 2015. This research was supported in part by the National Science Council, National Taiwan University and Intel Corporation under Grants NSC102-2911-I-002-001 and NTU102R7501, the Ministry of Science and Technology under the grant MOST 103-2221-E-002-022-MY3, the Major Research Plan of the National Natural Science Foundation of China under Grant 91118008, and by Key Projects in the National Science & Technology Pillar Program during the Twelfth Five-year Plan Period under Grant 2014BAC16B01. The associate editor coordinating the review of this paper and approving it for publication was B. Hamdaoui.

S.-C. Lin was with INTEL-NTU Connected Context Computing Center, Taipei 106, Taiwan. He is now with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: slin88@ece.gatech.edu).

L. Gu was with INTEL-NTU Connected Context Computing Center, Taipei 106, Taiwan. He is now with the Shanghai Research Institute of China Telecom, Shanghai 200122, China (e-mail: gulei@sttri.com.cn).

K.-C. Chen is with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei 106, Taiwan (e-mail: chenkc@cc.ee.ntu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2014.2376952

a thorough examination of average message delivery time for small-world networks in the continuum limit. Via random network analysis, [16] studies the properties of giant component in wireless multi-hop networks, while [17] provides a heterogeneous structure for such networks and conducts the throughput and delay analysis. Furthermore, the concepts of rumor and gossip routing algorithms are also widely employed in sensor networks [18] and *ad hoc* networks [19]. As for disconnected delay-tolerant MANETs and generalized complex networks, [20] and [21] respectively provide the social network analysis for information flow and epidemic information dissemination.

In this paper, inspired by small-world phenomenon, we connect data aggregators (DAs) to machine swarm and propose a promising two-tier heterogeneous architecture with DA's small-world network for statistical traffic control in large M2M communication networks. To address efficient dissemination control for routing and QoS such as surveillance applications, we first analytically supply the condition to establish connected M2M networks and explore some essential geometric properties (i.e., degree distribution, network diameter, and average distance) for the networks. Analytic bounds of average distance characterize the average number of hops that machines' packets need to traverse over the swarm, thus dominating the QoS guarantee capability for reliable communications. Furthermore, through G/G/1 (i.e., both inter-arrival time and service time distributions of a traffic queue are arbitrary distributions) queuing network model for traffic modeling, the practical data transportation takes place in connected M2M networks. Both the average end-to-end delay and maximum achievable throughput per machine from information dissemination in machine swarm multi-hop networking are examined. However, to control QoS over such networks still serves great difficulty due to the large network size. Aiming at statistical performance in large M2M networks, we propose a statistical control mechanism for the networks by establishing the heterogeneous network architecture and exploiting statistical QoS guarantee for end-to-end transmissions without the need of feedback control at each link. By forming DA's network with small-world property and linking machines to DAs, this novel heterogeneous architecture significantly improves the performance of end-to-end traffic for tolerable delay and makes dependable communications possible from guaranteing traffic QoS, with extremely simple network operation for each machine.

The contributions of this paper are summarized as follows.

1) To understand geometric properties of large M2M networks and thus benchmark performance, we first analytically examine network connectivity, degree, distribution, network diameter, and average distance under Poisson Point Process (PPP) machine distribution.
2) Introducing queuing network theory into such network analysis for practical data transportation, the average delay and achievable throughput for message delivery in connected M2M networks are analytically obtained as well as the QoS guaranteed throughput in real applications.
3) Standing on hereby established analysis, statistical dissemination control is proposed that incorporates DA's network with machine swarm (or sensor swarm) for favorable heterogeneous network architecture.
4) Due to infeasible end-to-end information exchange and subsequent precise control, we exploit statistical QoS guarantees over two-tier heterogeneous network architecture to exhibit remarkable enhancement of system performance, and to facilitate the merits of small-world phenomenon into engineering reality.

Simulation results show that our proposed control yields the significant throughput gain for delay guarantee performance. Such system performance asymptotically relieves Gupta and Kumar's dilemma [23] in scaling perspective for general wireless networks. Note that this paper is based on our preliminary research in [24] and [25]. However, different from [25] which only considers the average number of hops, this paper deals with the actual packet delay including transmission and queuing latency, which necessitates a new queuing network analysis. Furthermore, while the work in [24] provides the upper bound performance of end-to-end delay dedicated for the proposed routing algorithm, this paper studies the asymptotic performance of several statistical QoS requirements, such as end-to-end delay and maximum throughput as well as the throughput under guaranteed delay, for a general forwarding scheme in M2M network. What is more important, our previous work focuses on obtaining the traffic performance under a specific scenario setting, which can simplify the analysis, while failing to maintain the same level of transmission qualities when the scenario changes, e.g., the network topology or traffic pattern becomes different. In this paper, we solve this challenge through statistical dissemination control by leveraging the heterogeneous network architecture. In particular, the upper layer of DAs' network enables shortcut transmissions to reduce the excess end-to-end delay from the long route transmissions in the lower layer of machine swarm. A comprehensive performance analysis upon such a heterogeneous architecture is also included in this paper. With these accomplishments, we provide an original and significant paradigm to facilitate M2M communications, practically realizing information dissemination control to meet the need of time sensitive applications in next-generation wireless standards.

The rest of this paper is organized as follows. Section II presents related work and system model. Section III and Section IV provide M2M network topology analysis and queuing network model for large M2M networks, respectively. Statistical dissemination control with heterogeneous architecture is proposed in Section V with performance evaluation in Section VI. Section VII gives the conclusion and ends the paper.

II. Background and System Model

M2M communication network consists of tremendous self-organized machines/sensors and enables autonomous connections among different applications for ubiquitous communications upon such large swarm system. To facilitate this scenario into practice, providing the connectivity accompanied with reliable transportation is a must for such large network. In the following, we highlight the relevant research and introduce the M2M network model using geometric random graph (GRG) as its topology and local clustering property are suitable for benchmarking large wireless *ad hoc* sensor networks.

A. Background

Scanning the literature, [26]–[35] explore QoS issues for traffic control over next-generation wireless systems (i.e., 3GPP LTE/LTE-A based cellular systems). [26] offers a reallocation-based assignment that maximizes the spectral efficiency with QoS guarantees in multi-service wireless systems, achieving a good tradeoff between performance and computational complexity. To deal with scheduling policies over multi-hop wireless networks, [36] provides an analytical traffic delay analysis and [37] proposes a low-complexity congestion control with respect to per-flow delay. Aiming to minimize the energy consumption of overall heterogeneous network and preserving the QoS for users, [27] examines the optimal control for wake up mechanisms of femtocells. Studying the delay-throughput tradeoff, [28] utilizes an efficient power allocation scheme with minimized delay and high throughput for real-time services in distributed wireless networks. To reduce latencies and increase fairness in terms of transmitted frames, [38] further provides a distributed and online fair resource management in video surveillance sensor networks. With regard to power efficiency in vehicle-to-roadside infrastructure communication networks, [29] proposes a joint power and sub-carrier assignment policy under delay-aware QoS requirements. Reference [30] studies a tight integration of device-to-device communications into an LTE-A network, desirably exploiting spectrum of the existing radio networks. Reference [31] provides a systematic framework for the power and energy optimal system design in cellular-based M2M communications. Reference [32] proposes an efficient overlay to provide GSM connectivity within an LTE carrier for low data rate M2M customers. Aiming at a great number of applications in M2M communications, [33] proposes a massive access management on the air interface and [34] designs the data collectors to efficiently serve many uplink transmissions based on a random access scheme. To alleviate interference via cognitive radio technology, [39] pursuits capacity maximization under the SINR model in multi-hop cognitive radio networks. Moreover, [35] applies cognitive M2M communications in the smart grid and integrates reliability and timeliness in the QoS study. However, above excellent efforts do not thoroughly explore network characterizations and thus need end-to-end information regarding network topology, while end-to-end information requires non-scalable complexity of overhead. In this paper, instead of acquiring such information upon large networks with catastrophic complexity, we propose a totally different philosophy to proceed statistical dissemination control based on network topology analysis and traffic queuing model in large M2M communication networks.

B. Network Model

Suggested by [1], [2], [4], a cloud based M2M networks consists of two-tier network architectures. As in Fig. 1(a), the first tier contains a swarm of machines, which equip short-range wireless networking capability. The second tier involves wireless infrastructure and service cloud. The cloud's gateways are smart devices that collect and process data from machines and manage their operation. The service cloud further provides the accesses to M2M service, linking physical world to cyber world. Rather than guided by central controller(s) as in conventional cellular networks, nodes (i.e., machines) in M2M networks distributedly communicate with each other. There is no identity difference among nodes. Furthermore, due to wireless propagation and outage, each node can only communicate with other nodes within certain distance. It is assumed that there are n nodes in a M2M network, while each pair of nodes within distance r can communicate. Meanwhile, instead of exploiting classical Erdös-Rényi random graph [40], the geometric topology and local clustering property of GRG make itself preferable as a mathematical graphical model for wireless *ad hoc* networks [41]. The GRG model is based on a homogeneous PPP that randomly distributes nodes on the unit area to characterize the spatial distribution of nodes. In particular, GRG model with parameter n and r (i.e., $GRG(n, r)$) defines a graph with n nodes following homogeneous PPP and edges that are established when pairs of nodes are closer than radius r. With this understanding, GRG model is well suited for modeling M2M networks, especially for the first tier of machine swarm. Thus we adopt $GRG(n, r)$ for M2M network model, where nodes follow PPP and randomly distribute on the $[0, 1] \times [0, 1]$ flat square.

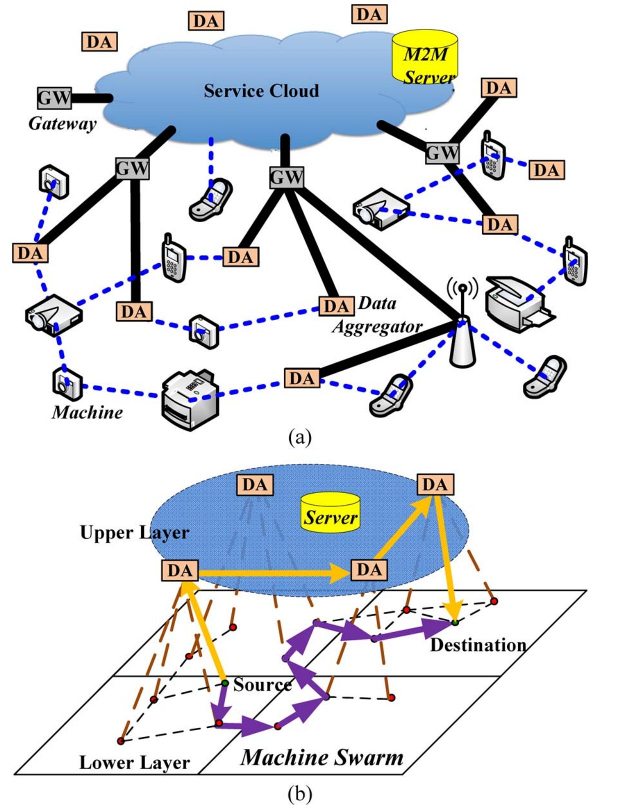

Fig. 1. Network topology of heterogeneous network architecture for statistical control in cloud M2M communication networks. (a) Cloud based scenario for M2M communication networks. (b) Heterogeneous network architecture for proposed statistical control.

C. Connectivity of M2M Networks and Information Dissemination Control

To transmit data packets across large machine swarm, constructing a connected M2M network is the necessity to ensure every packet could be sent to its corresponding destination from the source. Note that, in the rest of paper, we claim that a network has certain properties with high probability (i.e., almost surely) as $n \to \infty$ and $r(n) \to 0$. The connectivity of M2M network will be provided via network topology analysis later in Section III; however, we summarize the idea as follows. While each machine equips with short-range communication capability, multi-hop networking is necessary for end-to-end data transportation [4]. That is, for a single source-destination pair, there exist a source machine, a destination machine, and several relay machines that forward traffic from the source to the destination. Without loss of generality, it is assumed that sequences of packets follow the general arrival process and the general service time, and each transmission link is modeled as a $G/G/1/\infty/FCFS$ (or $G/G/1$ for the conventional abbreviation) queue [42]. In particular, such a queue represents a queuing system with a single server, infinite buffer size, and the scheduling discipline of first-come-first-serve (FCFS), where packet interarrival times have a general (meaning arbitrary) distribution and service times have a (different) general distribution. By connecting each link of $G/G/1$ queue, the entire $G/G/1$ queuing network is established for M2M network. Thus, upon this queuing network model, the analysis of network connectivity and information dissemination (i.e., end-to-end delay and maximum system throughput) are ready to be exploited.

In addition, as some real-time applications require bounded packet delay for end-to-end transmissions, the statistical QoS guarantee is considered. In particular, given the required statistical delay bounds, the QoS guaranteed throughput that satisfies the delay requirement is derived via Markov inequality. Meanwhile, considering source's excessive traffic loads for prodigious incoming data or poor forwarding capability from long multi-hop transportation, we further develop an effective statistical dissemination control with heterogeneous network architecture. This architecture significantly improves the end-to-end traffic performance under tremendous amounts of machines in large M2M network. The idea is as follows. As shown in Fig. 1(b), we aim to establish an "information highway" of ultra-fast forwarding potential for machine swarm by adding few DAs to form a small-world network [14], [22], [43] and connecting them with infrastructure networks. Once source's packets access to this "highway" after link transmissions to DA, these packets only traverse few steps in DA's network to arrive at the area near the destination from small-world phenomenon. The comprehensive examination for such promising architecture with corresponding transmission improvement is presented later in Section V-B. Standing on top of these accomplishments, the proposed statistical control successfully facilitates reliable information disseminations over large M2M communications of tremendous number of machines.

III. M2M Network Topology Analysis

To achieve information dissemination among all nodes in an M2M network, we first investigate the network connectivity by social network analysis [7], [14] for a connected M2M network. In particular, we study some useful geometric properties that facilitate our analysis of information flow later in Section IV.

A. Connected M2M Networks

As mentioned in Sections II-B and II-C, given a $GRG(n,r)$ model for M2M network, we obtain the following lemmas for network connectivity.

Lemma 1: For n nodes following PPP on a unit area (i.e., [0, 1] $\times$ [0, 1] flat square), the partition of flat square into smaller square grids with area $\log n/n$ is applied. Then, there is at least one node in each square grid almost surely, when $n \to \infty$.

Proof: Firstly, under mentioned partition, there are $n/\log n$ square grids in flat square. With PPP, the number of nodes in each square grid follows a Poisson random variable with mean $\log n$ and the probability that a specific square grid has no node is $\exp(-\log n)$. Since the probability of an intersection of events is no larger than any individual among them, the probability that there is a square with no node is $n\exp(-\log n)/\log n$, which approaches to zero for $n \to \infty$. It is noted that *Lemma 1* holds for any partition that gives square grids with area larger than $\log n/n$, especially the one gives square grids with area $(r/\sqrt{5})^2$ and $r \geq \sqrt{5\log n/n}$. Thus, we have *Corollary 1* for the connectivity of M2M networks. ∎

Corollary 1: For a M2M network with $GRG(n,r)$, if $r \geq \sqrt{5\log n/n}$, the network in connected almost surely.

Proof: Followed from *Lemma 1*, each node in a square grid is able to connect to its four neighborhood square grids almost surely and it becomes the lattice structure. Furthermore, since lattice structures are connected, the given GRG model and therefore the M2M network is connected almost surely. ∎

B. Geometric Properties

In the following, aiming at connected M2M networks, we explore geometric properties of networks that are essential for determining the dissemination performance in large networks.

1) Degree Distribution: We model a M2M network (i.e., $GRG(n,r)$) by assuming that all nodes follow PPP and edges are established when pairs of nodes are closer than certain distance r, its degree distribution can be obtained from ER model. Considering a specific vertex in $GRG(n,r)$, an edge connecting to its neighbor is present with the probability πr^2 for total $n-1$ possible neighbors. The degree distribution then follows a binomial distribution. Since we are interested in large M2M networks (i.e., $n \to \infty$), the degree distribution thus can be expressed by a Poisson distribution as

$$p_k = \mathbf{Pr}\{K=k\} = \frac{(n\pi r^2)^k \exp\{-n\pi r^2\}}{k!}. \tag{1}$$

The degree distribution of networks characterizes the number of nodes that potentially connects to other node in networks, representing the incoming traffic load that contributes to a single link for multi-hop communications in M2M networks. This is a crucial factor for reliable information dissemination, since traffic overload might bring deadlocks in link transmissions and deteriorate the QoS of end-to-end traffic.

2) Network Diameter: The network diameter for connected M2M networks can be directly obtained from *Lemma 1* as shown in *Theorem 1*.

Theorem 1: While $r \geq \sqrt{5\log n/n}$ for a connected M2M network with $GRG(n,r)$ model, the network diameter $R(n)$ in terms of hop-count has the upper bound $2\sqrt{5}/r$ and the lower bound $\sqrt{2}/r$ almost surely.

Proof: Please see Appendix A. ∎

3) Average Distance: To obtain the average distance for pairs of nodes in M2M networks, we examine the Euclidean distance over all pairs of vertexes in $GRG(n,r)$ model.

Lemma 2: For Euclidian distance d, let d_x and d_y denote for the projective length of d in unit area for x and y axes, respectively. Then, in connected $GRG(n,r)$, d in terms of hop-count has the upper bound $\sqrt{5}(d_x+d_y)/r$ and the lower bound d/r, given $r \geq \sqrt{5\log n/n}$.

Proof: Please see Appendix B. ∎

While the bounds are $\lceil\sqrt{5}d_x/r\rceil + \lceil\sqrt{5}d_y/r\rceil$ and $\lfloor d/r\rfloor$, we assume $\sqrt{5}(d_x+d_y)/r$ and d/r to be integers in *Lemma 2*. However, when $r \to 0$, the difference can be neglected. Before exploring the average distance of M2M networks, we first have the following proposition for uniform random variable that arises from machines' random positions.

Proposition 1: Let $U(S)$ denote the uniform distribution on S, where S is a connected subinterval of $\mathcal{R}^k$, and $|\cdot|$ and $\|\cdot\|$ denote absolute value for one dimension and Euclidean distance. For a unit square, if two random variables X, $Y \sim U([0,1])$, $\mathbf{E}[|X-Y|] = 1/3$. And if these X, $Y \sim U([0,1]\times[0,1])$, $\mathbf{E}[\|X-Y\|] \geq \sqrt{2}/3$.

The equality of *Proposition 1* comes from the derivation in calculus and the inequality comes from $\sqrt{a^2+b^2} \geq (|a|+|b|)/\sqrt{2}$. Finally, we conclude with *Theorem 2*.

Theorem 2: While $r \geq \sqrt{5\log n/n}$ for a connected M2M network with $GRG(n,r)$ model, the average distance $d(n)$ in terms of hop-count is bounded almost surely as

$$\frac{\sqrt{2}}{3r} < d(n) < \frac{2\sqrt{5}}{3r}. \tag{2}$$

Proof: Please see Appendix C. ∎

For connected M2M networks, especially for $r \sim \sqrt{\log n/n}$, *Theorems 1* and *2* suggest that $R(n)$ and $d(n)$ follow the order of $1/r$. The asymptotic notations [44] of $R(n)$ and $d(n)$ are both $\Theta(1/r) = \Theta(\sqrt{n/\log n})$. Furthermore, while $d(n)$ provides average hop number for traversed packets, it highly relates to average delay for end-to-end transmissions as we will see in the following section and thus plays a critical role for information dissemination in M2M networks.

IV. Queuing Network Model

Without the need of end-to-end information to escape catastrophic complexity, information dissemination becomes the only way in machine swarm. As suggested by Section II-C, we exploit an open $G/G/1$ queuing network model for delay and throughput analysis of M2M networks. Furthermore, the diffusion approximation is used to analyze the queuing network. Our analytical methodology to deal with wireless networks have general inter-arrival and service time distributions by providing closed form expressions of end-to-end delay and maximum achievable throughput per node. In the following, to fully understand practical data transportation, we present the traffic model and an equivalent queuing network model in connected M2M networks.

A. The Traffic Model

For an $GRG(n,r)$ of connected M2M network with $r \geq \sqrt{5\log n/n}$, there are n nodes uniformly distributed over an unit area, numbered from 1 to n, and each capable of transmitting at W bits per second. The set of neighbors of node i is denoted by $N(i)$ and each node can be a source or a destination of packets. The external arrival of jobs (i.e., new packets arrive in the network) is a renewal process with rate λ_e. The squared coefficient of variance (SCV) of inter-arrival time of new packets equals to c_A^2. We assume that packets of size L bits are generated by each node according to an independent identical distribution (i.i.d.) Poisson process with rate λ. The mean and SCV of the service time at node i are denoted by $\mathbf{E}[S_i]$ and $c_{Bi}^2 = (\mathbf{E}[S_i^2] - \mathbf{E}[S_i]^2)/\mathbf{E}[S_i]^2$, respectively. As a packet is generated by a node, it transverses over the network by multi-hop relaying until it reaches the destination. The probability that a packet received by its destination is $p(n)$ and is referred as "absorption probability". Alternatively, the probability that a packet received by a node is forwarded to a neighboring node is $(1-p(n))$. If a packet is not absorbed by a node, then all the neighboring nodes are equally likely to be the next hop of the packet. $p(n)$ characterizes the degree of locality of the traffic as: The traffic is highly localized for large $p(n)$, while small $p(n)$ implies unlocalized traffic. Therefore, it is easy to quantify the dependence of delay and capacity on average distance $d(n)$ of the network as presented in Section IV-B. In addition, each node is assumed to have infinite buffers and thus no packets are dropped in the network. The packets are served by nodes in FCFS manner. The queuing network model for connected M2M network is shown in Fig. 2(a). The stations of queuing network represent nodes of M2M network and the forwarding probability p_{ij} equals to the probability that a packet is transmitted from the queue at node i to the queue at node j. Moreover, the queue at a node as a station in queuing network is provided in Fig. 2(b).

B. End-to-End Delay Analysis

The end-to-end delay in a connected M2M network is defined as the sum of the queuing and transmission delays at all intermediate relaying nodes. To evaluate the delay, we first prove *Lemma 3* that relates absorption probability with average distance of the network. After that, by deriving expressions for some parameters of the queuing network, we analytically obtain the end-to-end delay.

Lemma 3: In a M2M network, the number of hops traversed by a packet $d(n)$ equals to $1/p(n)$.

Proof: Please see Appendix D. ∎

$d(n)$ in *Lemma 3* is obtained from (2) and provides $p(n)$ for connected M2M networks.

Lemma 4: Given the number of neighbors of node i as K_i, the probability that a packet is forwarded from node i's queue to node j's queue, denoted by p_{ij}, is $(1-p(n))/K_i$ if $j \in N(i)$; otherwise, it is zero.

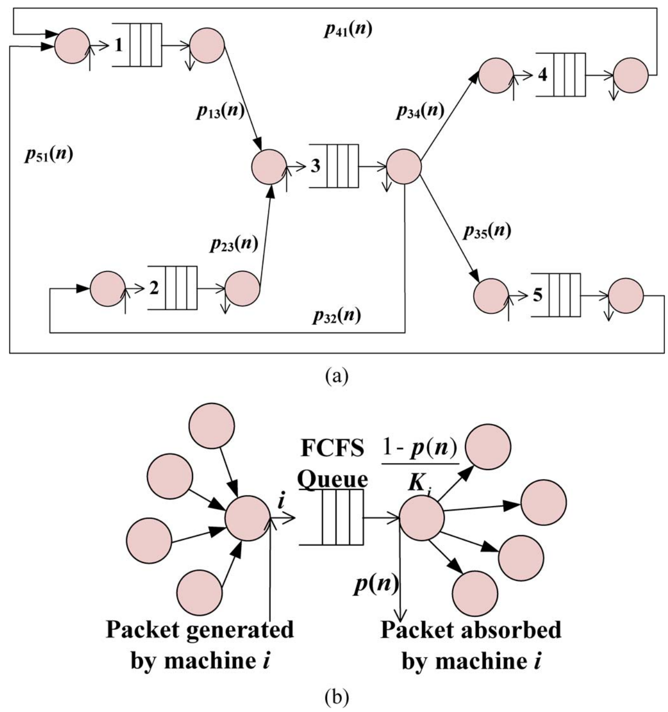

Fig. 2. Queuing network model for a connected M2M network. (a) A connected M2M network as queuing network. (b) A machine in a connected M2M network represented as a station of the queuing network.

While $p_{ij}(n)$ equals to the product of $\mathbf{Pr}\{i$ transmits the packet to j | packet not absorbed$\}$ and $\mathbf{Pr}\{$packet is not absorbed by $i\}$, $\mathbf{Pr}\{i$ transmits the packet to j | packet not absorbed$\}$ equals to $1/K_i$ if $j \in N(i)$; otherwise, it equals to 0. $\mathbf{Pr}\{$packet is not absorbed by $i\} = 1 - p(n)$.

To illustrate how the number of neighbors of a node in a M2M network (i.e., K_i for node i) affects the visit ratio of that node (i.e., the average number of times a packet will visit such node as the packet enters the network) in *Lemma 5*, we start from the following proposition.

Proposition 2: For node i, let K_i denotes the number of its neighbors and $K_i^{(2)}$ denotes the number of its two-hop neighbors. Given a connected M2M network with $GRG(n,r)$ model (i.e., $r \geq \sqrt{5\log n/n}$), we have $1/\delta = \mathbf{E}[\sum_{j=1}^{K_i^{(2)}} 1/K_j] > 1 - \exp\{-n\pi r^2\}$.

From Section III-B1, K_i is a Poisson random variable following the distribution in (1). Via branching process, $K_i^{(2)}$ is the second generation of node i, whose distribution is $q_{k-1} = \mathbf{Pr}\{K^{(2)} = k-1\} = (kp_k)/\chi$ with $\chi = \sum_k kp_k$ for $k \geq 1$. $K_i^{(2)}$ becomes a Poisson random variable with mean $n\pi r^2$ and we have $\frac{1}{\delta} = \mathbf{E}\left[\mathbf{E}\left[\sum_{j=1}^{k} \frac{1}{K_j} | K_i^{(2)} = k\right]\right] = n\pi r^2 \mathbf{E}\left[\frac{1}{K_1}\right]$ and $\mathbf{E}\left[\frac{1}{K_1}\right] = \sum_{k=0}^{\infty} \exp\{-n\pi r^2\} \frac{(n\pi r^2)^k}{k! \times k} > \sum_{k=0}^{\infty} \exp\{-n\pi r^2\} \frac{(n\pi r^2)^k}{k! \times (k+1)} = \frac{1}{n\pi r^2} - \frac{\exp\{-n\pi r^2\}}{n\pi r^2}$.

Lemma 5: With parameter δ obtained for given connected M2M network by *Proposition 2*, the visit ratio of a node i, denoted by e_i, equals to $e_i = \delta/\{n[\delta - (1 - p(n))]\}$.

Proof: Please see Appendix E. ■

Lemma 6: The effective packet arrival rate at a node i, denoted by λ_i, equals to $\lambda\delta/(\delta - 1 + p(n))$.

Since the packet generation process at each node is an i.i.d. Poisson process with rate λ, new packets arrive in the network at rate $\lambda_e = n\lambda$. Furthermore, there are two sources of packet arrivals at a node: The packets that are generated at the node and the packets that are forwarded to the node by other nodes.

The utilization factor of node i, denoted by ρ_i, is given by $\rho_i = \lambda_i \mathbf{E}[S_i]$. Furthermore, regarding the spatial concurrency constraints in link transmissions, nodes close to a receiver must be idle to avoid collisions which results in the loss of packets. With $n\pi r^2$ nodes surrounding node i, the event for successful link transmissions to i follows a Bernoulli process with success probability $1/(n\pi r^2)$. Thus, we have $\mathbf{E}[S_i] = n\pi r^2 L/W$. The SCV of inter-arrival time at node i, denoted by c_{Ai}^2, is approximated using $c_{Ai}^2 = 1 + \sum_{j=0}^{n}(c_{Bj}^2 - 1)(p_{ij}(n))^2 e_j e_i^{-1} = 1 + (c_{Bi}^2 - 1)(1 - p(n))^2 \psi$ and $\psi > [1 - \exp\{-n\pi r^2\}]/(n\pi r^2) - \exp\{-n\pi r^2\}$ where $c_{B0}^2 = c_A^2$ and ψ is obtained from the same manner as for e_i. According to the diffusion approximation, the approximate expression for the probability that the number of packets at node i equals to t, denoted by $\pi_i(t)$, is $1 - \rho_i$ if $t = 0$; otherwise, it is $\rho_i(1 - \hat{\rho}_i)\hat{\rho}_i^{t-1}$ as $t > 0$. where $\hat{\rho}_i = \exp\{-2(1 - \rho_i)/(c_{Ai}^2 \rho_i + c_{Bi}^2)\}$. The mean number of packets at node i, denoted by L_i, is therefore $L_i = \rho_i/(1 - \hat{\rho}_i)$. With above results, we present the end-to-end delay in the following.

Theorem 3: For a connected M2M network with parameter δ described in *Proposition 2*, the average end-to-end delay, denoted by $D(n)$, is

$$D(n) = \frac{\rho\,[(\delta - 1)d(n) + 1]}{\lambda\delta(1 - \hat{\rho})}. \tag{3}$$

Proof: Please see Appendix F. ■

C. Maximum Achievable Throughput

For a connected M2M network, we derive the expression for maximum achievable throughput λ_{max} in the following. λ_{max} is the maximum value of the packet arrival rate λ at the nodes for which the average end-to-end delay remains finite.

Theorem 4: For a connected M2M network with parameter δ described in *Proposition 2*, the maximum achievable throughput is

$$\lambda_{max} = \frac{W\,[(\delta - 1)d(n) + 1]}{\delta d(n) n\pi r^2 L}. \tag{4}$$

Also from (4), $\lambda_{max} = \Theta(1/[d(n)nr^2])$.

To have finite delay, $\lambda_i \mathbf{E}[S_i] = \lambda_i n\pi r^2 L/W < 1$. Furthermore, as $n \to \infty$, we have $\delta \to 1$ from *Proposition 2*.

The result of *Theorem 4* coincides with Gupta and Kumar's classic research [23] for the asymptotic case where $n \to \infty$. It is obvious that λ_{max} increases with decreasing in $d(n)$. From *Theorem 2* in Section III-B3, for a connected M2M network, $r \geq \sqrt{5\log n/n}$ and the number of hops between arbitrary source and destination pair would be $\Theta(\sqrt{n/\log n})$. Thus, we get $\lambda_{max} = \Theta(1/[d(n)nr^2]) = \Theta(1/\sqrt{n\log n})$, which confirms with the asymptotic capacity of multi-hop wireless *ad hoc* networks as from [23].

V. Statistical Dissemination Control

As real applications require dependable networking performance across the swarm, QoS along with large network diameter creates a new challenge. Aiming at statistical performance in large M2M networks [24], [25], we propose a statistical control mechanism for the networks by establishing the heterogeneous network architecture. By constructing the "small-world shortcut" scheme [14] among DAs, the effective traffic control is obtained with great system throughput under desired QoS constraint (i.e., delay bound). In particular, any machine can identify a shortcut through a corresponding DA, such that end-to-end delay can be significantly reduced under such heterogeneous network architecture.

A. Statistical QoS Guarantee

The real-time services generally care the end-to-end delay and demand bounded delays. Rather than providing *deterministic* QoS guarantees (i.e., the probability of delay requirement violation is zero), a more practical and reasonable solution is to provide *statistical* guarantees [45] (i.e., the probability that the packet violates its delay constraint is bounded) for QoS as $\mathbf{Pr}\{Delay \geq D_{max}\} \leq \tau$, where D_{max} is the delay requirement and τ is used to characterize the degree of statistical QoS guarantee. By different inequalities or bounds, $\mathbf{Pr}\{Delay \geq D_{max}\} \leq f(Delay, D_{max})$. Formulating f function by packet delay analysis, we thus get the end-to-end throughput with statistical delay guarantees as the maximum available load from the source that satisfies the constraint: $f(Delay, D_{max}) \leq \tau$.

1) QoS Guaranteed Throughput in M2M Networks: For reliable communications in large connected M2M networks, we obtain the end-to-end throughput with QoS guarantee via Markov inequality as follows.

Proposition 3: For a connected M2M network with parameter δ described in *Proposition 2*, the QoS guaranteed throughput is obtained as $\lambda^{Q}_{max} = \frac{[(\delta-1)d(n)+1](c^2_{Bi}\ln T+2)}{\delta d(n)\mathbf{E}[S_i](2-c^2_{Ai}\ln T)} = \left(\frac{c^2_{Bi}\ln T+2}{2-c^2_{Ai}\ln T}\right)\lambda_{max}$, where $T = 1 - \{d(n)\mathbf{E}[S_i]\}/(\tau D_{max}) = 1 - [d(n)n\pi r^2 L]/(W\tau D_{max})$ and $\lambda^{Q}_{max} = \Theta(1/[d(n)nr^2])$. Via Markov inequality, the statistical delay guarantee is provided as $\mathbf{Pr}\{Delay \geq D_{max}\} \leq \{\rho[(\delta-1)d(n)+1]\}/[\lambda\delta(1-\hat{\rho})D_{max}]$ and λ^{Q}_{max} is the maximum arrival rate holding such inequality. *Proposition 3* shows that even providing statistical QoS guarantee in end user's traffic, we still can maintain the system throughput that closes to the maximum achievable throughput for the asymptotic case (i.e., λ^{Q}_{max} and λ_{max} follow the same order as $n \to \infty$). Note that there exists a tradeoff between the delay requirement and the attainable throughput (i.e., both maximum achievable and QoS guaranteed throughput). In particular, from the formulation of λ^{Q}_{max} and λ_{max}, they both include the delay factor (i.e., $d(n)$) in the denominator and thus increase (decrease) when the corresponding delay decreases (increases).

B. Small-World Shortcut

To leverage small-world feature into machine swarm, we first create a promising two-tier heterogeneous network architecture by adding some DAs to help machines' transmissions. After that, we confirm shorter average distance for pairs of nodes via shortcut in network. We finally present the end-to-end delay reduction with improved system throughput and therefore confirm our proper control for information dissemination in large M2M communication networks.

1) Heterogeneous Network Architecture: The heterogeneous network architecture for connected M2M network, as shown in Fig. 1(b), consists of two layers. The lower layer is the machine swarm modeled by a GRG model with n nodes and radius r, while the upper layer is a 2D-lattice network with m^2 DAs. The size of DA's lattice network is much smaller as compared machine's wireless *ad hoc* network (i.e., $m \ll n$). Such DAs partition the bottom [0, 1] $\times$ [0, 1] unit area into equal $m \times m$ grids with average n/m^2 nodes per grid and each DA can communicate with z random selected nodes in the correspondent grid. To prevent traffic jam and thus deadlock in DA's network, we assume each DA's service rate, denoted by $1/\mathbf{E}[S^D_i]$, is the product of machine's service rate and the number of serving machines. That is, $\mathbf{E}[S^D_i] = (m^2/n) \times L/W$. In the following, through network property and system performance, we exhibit that with few DAs (i.e., m^2 and $m \ll n$) installed in machine swarm, smaller number of hops and thus significant performance improvement are feasible to robust and reliable information dissemination in large M2M networks.

Note that we consider the uniformly deployment of homogeneous DAs here, as machines are uniformly distributed in a given area from the GRG model of M2M network. When the machine's deployment becomes uneven due to issues like mobility, our model can be easily adaptive to the heterogeneous scenario by adjusting two related parameters (i.e., the total number of DAs m^2 and the amount of machines z that each DA can simultaneously communicate with). For example, for the dense machine area, we can increase the number of DAs (i.e., with larger m) or/and enhances the serving capability of DA (i.e., with larger z). Similar concern can be applied to sparse machine area. However, to avoid an unclear delivery of our main contribution, we focus on the homogeneous DAs for heterogeneous network architecture in the rest of paper.

2) Shorter Distance via Shortcut: To obtain the average hop number for node pairs under heterogeneous architecture, we first examine the average hops from nodes (machines) to DAs.

Lemma 7: In connected $GRG(n,r)$ with m^2 DAs forming upper lattice network and each DA serves z nodes, the heterogeneous architecture has the average distance from nodes to DAs $d^L(n)$ in terms of hop-count almost surely when $n, m \to \infty$ and $mr \to 0$ as $\pi/(120zmr) < d^L(n) < 2\sqrt{5}/(3zmr)$.

Proof: Please see Appendix G. ■

With the aid of *Lemma 7*, we provide the average distance via shortcut (i.e., DA's lattice network as an illustration) for heterogeneous architecture in the following.

Theorem 5: While $r \geq \sqrt{5\log n/n}$ for a connected M2M network with $GRG(n,r)$ model and m^2 DAs form lattice network as each DA serves z nodes concurrently, the heterogeneous architecture has the average distance $d^S(n)$ in terms of hop-count almost surely when $n, m \to \infty$ and $mr \to 0$ as $\frac{\pi}{60zmr} + \frac{2m}{3} < d^S(n) < \frac{4\sqrt{5}}{3zmr} + \frac{2m}{3}$.

TABLE I
SIMULATION PARAMETERS AND VALUE SETTINGS FOR PERFORMANCE EVALUATIONS

Simulator Setup	
Parameters	***Values***
Simulator	Matlab
For each evaluated value, run 100 times and average over samples.	
Machine Swarm Module	
Number of machines n	List in each figure
Position of machines	Randomly distribute with PPP in $[0, 1] \times [0, 1]$ flat square
Machine communication range r	$\sqrt{5 \log n/n}$ to ensure connectivity
Heterogeneous Architecture Module	
Length of lattice side m	List in each figure
Number of DAs	m^2 from 2D-lattice network
Position of DAs	Uniformly distribute
Number of machines handled by each DA at one time z	List in each figure
Traffic Module	
Packet arrival rate at each machine	$2(pkt/ms)$
Packet service rate at each machine	$1000(pkt/ms)$
Packet absorption at each machine $p(n)$	$\sqrt{5 \log n/n}$
QoS Requirement	
Maximum delay bound D_{max}	$1000ms$
Delay violation probability τ	0.02

Proof: Please see Appendix H. ∎

With *Theorem 5* and $m = \Theta(1/\sqrt{r})$, the shortcut is exhibited from the average distance as $d^S(n) = \Theta(1/\sqrt{r}) = \Theta(\sqrt[4]{n/\log n})$, which is much smaller than $d(n) = \Theta(1/r)$ for plain machine swarm.

Delay Reduction for Improved Throughput: In this section, aiming at heterogeneous architecture of large M2M networks, we derive average end-to-end delay, maximum achievable throughput, and the achievable throughput under QoS constraint to bring the merits of heterogeneous framework. Since DA's lattice network partitions the machine swarm into $m \times m$ grids, there are approximate n/m^2 nodes in each grid. Similar to *Theorem 5*, $D^S(n)$ is obtained from the aggregation of the delay from a node to DA, denoted by $D^L(n)$, and the delay in upper DA's lattice, denoted by $D^U(n)$.

Theorem 6: While $r \geq \sqrt{5 \log n/n}$ for a connected M2M network with $GRG(n,r)$ model and m^2 DAs form lattice network as each DA serves z nodes concurrently, the average end-to-end delay $D^S(n)$ for the heterogeneous architecture is $D^S(n) = 2D^L(n) + D^U(n) = \frac{2\rho^L[(\delta-1)d^L(n)+1]}{\lambda\delta(1-\hat{\rho^L})} + \frac{\rho^U m^2}{n\lambda(1-\hat{\rho^U})}$, where ρ^L and $\hat{\rho^L}$ are obtained as usual by $p(n) = 1/d^L(n)$, $\rho^U = \{2m\rho^L[(\delta-1)d^L(n)+1]\}/(3n\pi r^2\delta d^L(n))$, and $\hat{\rho^U} = \exp\left\{-2(1-\rho^U)/\left\{\left[1+m(c_{Bi}^2-1)\left(1-\frac{3}{2m}\right)^2\right]\rho^U + c_{Bi}^2\right\}\right\}$.

Proof: Please see Appendix I. ∎

When $n \to \infty$, $D^L(n)$ dominates end-to-end delay as $m = \Theta(1/\sqrt{r})$. Furthermore, since the utilization factor proportions to the average distance (e.g., $\rho^L \propto d^L(n)$), $D^S(n) = \Theta(1/\sqrt{r}) = \Theta(\sqrt[4]{n/\log n})$ and $D(n) = \Theta(1/r) = \Theta(\sqrt{n/\log n})$. Therefore, through the assistance from shortcut of DA's network, the heterogeneous architecture for M2M networks enjoys significant delay reduction. In the following, we further provide maximum achievable and QoS guaranteed throughput for this architecture.

Theorem 7: While $r \geq \sqrt{5 \log n/n}$ for a connected M2M network with $GRG(n,r)$ model and m^2 DAs form lattice network as each DA serves z nodes concurrently, the maximum achievable throughput λ_{max}^S for the heterogeneous architecture is

$$\lambda_{max}^S = \min\left(\frac{W\left[(\delta-1)d^L(n)+1\right]}{\delta d^L(n)n\pi r^2 L}, \frac{3W}{2mL}\right). \tag{5}$$

Also from (5), $\lambda_{max}^S = \Theta(1/[d^L(n)nr^2])$. Furthermore, since $\hat{\rho^L}$ and $\hat{\rho^U}$ are functions of λ (i.e., arrival rate per machine), the QoS guaranteed throughput, denoted by λ_{max}^{SQ}, satisfies $\nu^L/(1-\hat{\rho^L})+\nu^U/(1-\hat{\rho^U}) = \tau D_{max}/(2\mathbf{E}[S_i])$ where $\nu^L = d^L(n)$ and $\nu^U = m^3/3n^2\pi r^2$ and $\lambda_{max}^{SQ} = \Theta(1/[d^L(n)nr^2])$, too.

Proof: Please see Appendix J. ∎

Theorem 7 shows that λ_{max}^S increases with decreasing $d^L(n)$, and $D^L(n)$ dominates $D^S(n)$ for the asymptotic case (i.e., $n \to \infty$). Both λ_{max}^S and λ_{max}^{SQ} follow $\Theta(1/\sqrt[4]{n(\log n)^3})$. Consequently, we successfully provide a heterogeneous architecture that embraces significant throughput gain, as benchmark to Gupta and Kumar's results [23], for large M2M communications.

VI. PERFORMANCE EVALUATION

We compare the performance of the proposed heterogeneous network architecture with plain machine swarm. Simulation results confirm that heterogeneous architecture achieves remarkable delay reduction as well as high throughput gain with only few DAs installed, favored by practical implementation in large M2M networks. All simulation parameters and value settings are listed in Table I. In particular, to ensure every packet could be sent to its corresponding destination from the source,a connected M2M network is first established via the proposed analysis (i.e., selecting the appropriate machine communication range r with respect to the total machine number n). When a source machine generates a packet, it routes the packet to a specific destination, uniformly selected among other machines. Moreover, for plain machine swarm, source simply hops forward based on the sensing and relaying; for heterogeneous architecture, it employs dissemination without selecting a particular DA. In the following, we first evaluate average distance to DAs and end-to-end distance for plain machine swarm and heterogeneous architecture. Next, end-to-end packet delay, maximum system throughput, and throughput under guaranteed delay are thoroughly examined for such different architecture and compared with Gupta and Kumar's milestone results [23]. Finally, the simulation validation in the Metropolis is established to facilitate our design into an even more practical stage.

A. Average Distance for Plain Machine Swarm and Heterogeneous Architecture

To examine the improvement of average distance under heterogeneous architecture, we first study the average distance from machines to DAs. Regarding 2000 machines in swarm and averaging over 100 samples, Fig. 3(a) shows that such distance decreases with respect to increased amounts of DAs m and their connection links to nodes z. m and z jointly decide the required number of hops for machines' data to leave for DAs' small-world shortcut, which equivalently provides an ultra-fast information "highway" among machines. Based on these results, Fig. 3(b) and (c) show the asymptotic behaviors of average end-to-end distance for information dissemination within various settings of heterogeneous architecture as compared within plain machine swarm. As $z = 2$, Fig. 3(b) shows that such distance reductions become conspicuous with respect to increasing size of machine swarm in scaling perspective. Furthermore, establishing DAs' network with $m = 3$, Fig. 3(c) demonstrates such distance improvement regarding DAs' connection links. Both figures demonstrate consistency with average distance scales with the order of $\Theta(\sqrt{n/\log n})$ for plain machine swarm and with the order of $\Theta(\sqrt[4]{n/\log n})$ for heterogeneous architecture as suggested in *Theorem 2* and *Theorem 5*. As a summary, all these results confirm asymptotically greater performance for information dissemination (i.e., less end-to-end delay and more system throughput) in our proposed network architecture over plain machine swarm.

B. End-to-End Delay Reduction via Small-World Shortcut

Fig. 4 provides the asymptotic behaviors of average end-to-end delay for M2M networks under plain machine swarm and proposed heterogeneous architecture. While Fig. 4(a) shows the delay reduction via small-world shortcut of proposed schemes regarding $z = 3$ (i.e., each DA links to three machines)and various m (i.e., there are m^2 DAs in heterogeneous architecture), Fig. 4(b) shows the similar reduction regarding $m = 3$ and various z. Such improvement on delay comes from the innovative design of DAs' structure to proceed the small-world phenomenon (i.e., the upper lattice network). Moreover, above results also serve as the benchmark for the theoretical lower bound of achievable end-to-end delay and indicate the practicability of heterogeneous architecture. To sum up, by adding few DAs and tailoring their network structure, heterogeneous architecture brings much less delay and thus facilitates reliable disseminations in large next-generation networks such as M2M networks.

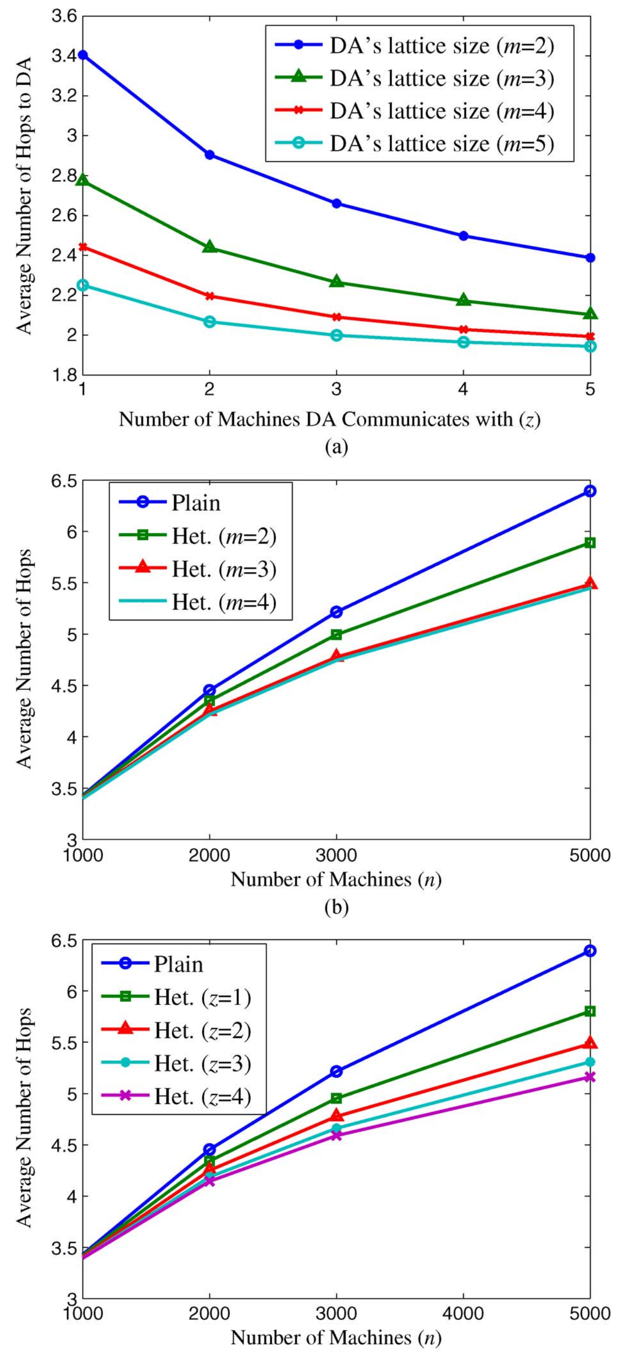

Fig. 3. Average distance in scaling perspective for M2M networks under plain machine swarm (Plain) and proposed heterogeneous architecture (Het.), where m^2 is the amount of DAs and z is the number of machines that each DA communicates with. (a) Average distance to DA over 2000 machines for various heterogeneous network architectures with respect to various m and z. (b) Asymptotic behaviors of average end-to-end distance for plain machine swarm and various heterogeneous network architectures with $z = 2$. (c) Asymptotic behaviors of average end-to-end distance for plain machine swarm and various heterogeneous network architectures with $m = 3$.

C. Maximum Throughput and Throughput Under Guaranteed Delay

%While the performance over plain machine swarm is predicted by Gupta and Kumar's analysis [23], Fig. 5(a) and (b) depict the maximum achievable system throughput for

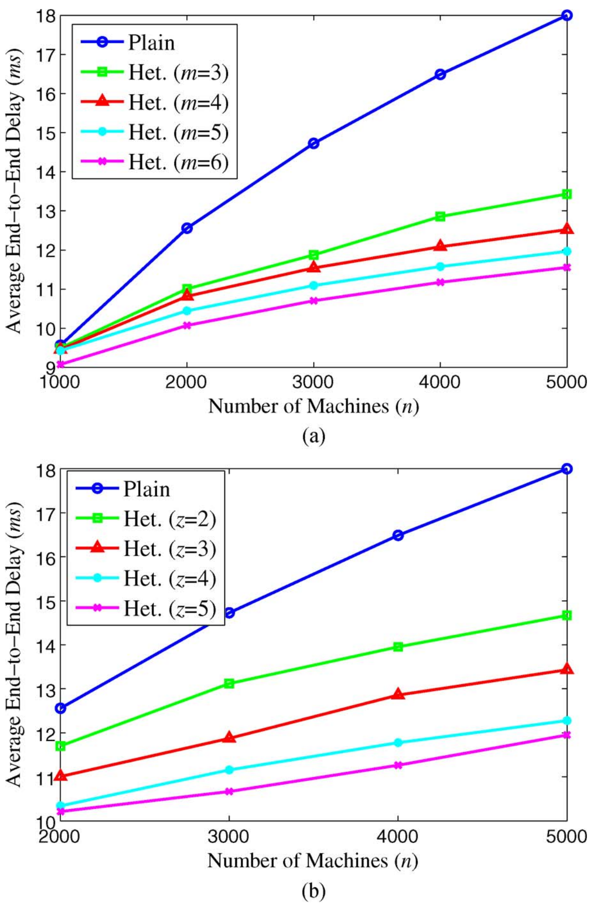

Fig. 4. Asymptotic behaviors of average end-to-end delay for M2M networks under plain machine swarm (Plain) and proposed heterogeneous architecture (Het.), where m^2 is the amount of DAs and z is the number of machines that each DA can communicate with. (a) Average end-to-end delay for plain machine swarm and various heterogeneous network architectures with $z = 3$. (b) Average end-to-end delay for plain machine swarm and various heterogeneous network architectures with $m = 3$.

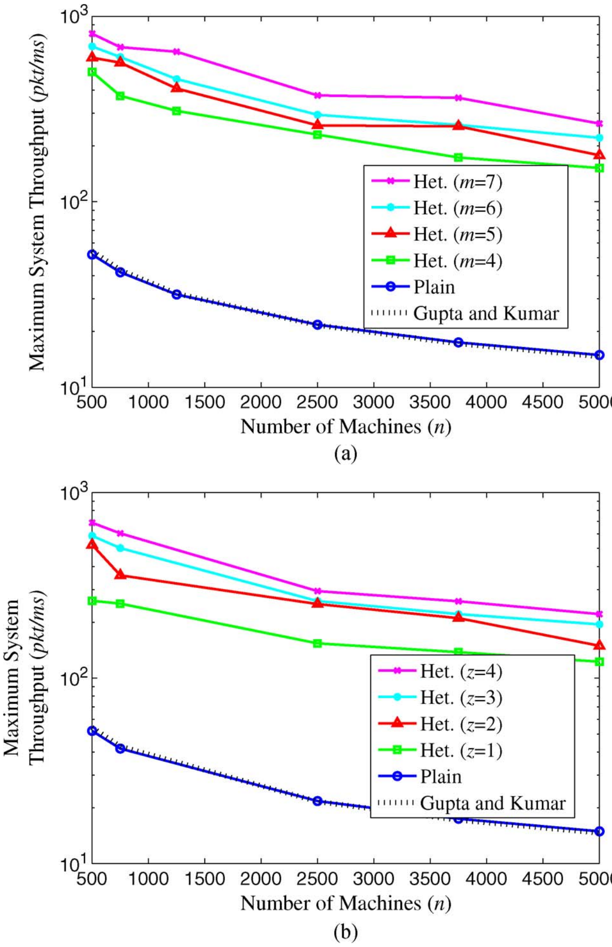

Fig. 5. Asymptotic behaviors of maximum system throughput with regard to QoS guarantee for M2M networks, where m^2 is the amount of DAs and z is the number of machines that each DA can communicate with. (a) Maximum system throughput for Gupta's results [23] (i.e., $\Theta(1/\sqrt{n\log n})$), plain machine swarm (Plain), and various heterogeneous network architectures (Het.) with $z = 4$. (b) Maximum system throughput for Gupta's results [23] (i.e., $\Theta(1/\sqrt{n\log n})$), plain machine swarm (Plain), and various heterogeneous network architectures (Het.) with $m = 6$.

different heterogeneous architectures. While the results of [23] accurately predict the networking performance of machine swarm, especially for great amounts of machines, proposed controls provide remarkable throughput improvements due to tolerable end-to-end delay. Such enhancement becomes significant as the small-world phenomenon of DAs' network takes charge of end-to-end data transportation (i.e., more DAs as m increased and better accessibility to DAs as z increased). Specifically, by decreasing the time duration in machines' *ad hoc* network or increasing the time sojourned in DAs' network, machines' packets experience much less delay and thus bring greater system throughput for M2M communication networks.

Furthermore, the corresponding circumstance also happens for QoS guaranteed throughput as shown by Fig. 6(a) and (b). Given the delay bound $D_{max} = 1000$ ms and the violation probability $\tau = 0.02$, Fig. 6(a) and (b) provides the throughput under QoS guarantees with regard to maximum throughput in Fig. 5(a) and (b). While both maximum and QoS guaranteed throughput of plain machine swarm asymptotically follow $\Theta(1/\sqrt{n\log n})$ as expected from *Theorem 4*, more throughput is obtained for proposed schemes due to heterogeneous architecture as it asymptotically follows $\Theta(1/\sqrt[4]{n(\log n)^3})$ shown in *Theorem 7*.

In addition, with respect to different QoS requirements, Fig. 7 exhibits the effectiveness of heterogeneous network architecture for greater achievable throughput as compared with plain swarm under the same QoS constraint. Loose τ gives more throughput for both plain and heterogeneous schemes as suggested by the previous discussion of existing tradeoff, but heterogeneous schemes are able to provide promising guaranteed throughput even under strict QoS demand for tight τ. Moreover, Fig. 8 further provides the exhaustive throughput comparison among different scenarios to complete our evaluation. While QoS guaranteed throughput is upper bounded by maximum achievable throughput, the great throughput improvement is provided by heterogeneous architecture as compared with plain machine swarm. This suggests that even when there are tremendous amounts of machines as in the next-generation networks, our methodology still accommodates desire QoS guarantees and system throughput by establishing ultra-fast (in terms of routing) "highway" through heterogeneous architecture.

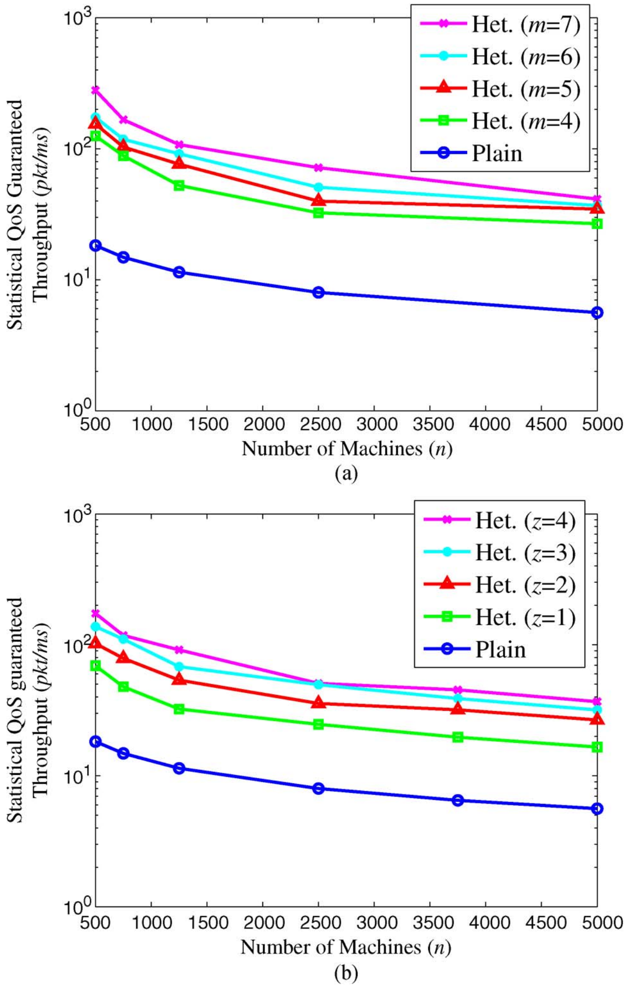

Fig. 6. Asymptotic behaviors of guaranteed system throughput with regard to QoS guarantee for M2M networks. (a) Statistical QoS guaranteed throughput with $D_{max} = 1000$ ms and $\tau = 0.02$ for plain machine swarm (Plain) and various heterogeneous network architectures (Het.) with $z = 4$. (b) Statistical QoS guaranteed throughput with $D_{max} = 1000$ ms and $\tau = 0.02$ for plain machine swarm (Plain) and various heterogeneous network architectures (Het.) with $m = 6$.

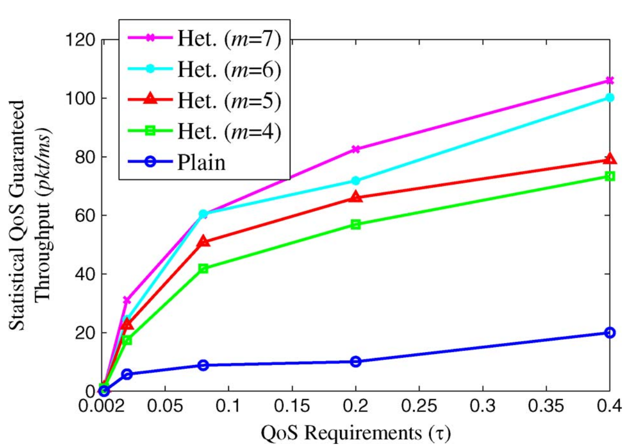

Fig. 7. QoS guaranteed throughput with respect to τ, $D_{max} = 1000$ ms and 5000 machines for plain machine swarm (Plain) and heterogeneous network architectures (Het.) with $z = 2$, where m^2 is the amount of DAs and z is the number of machines that each DA communicates with.

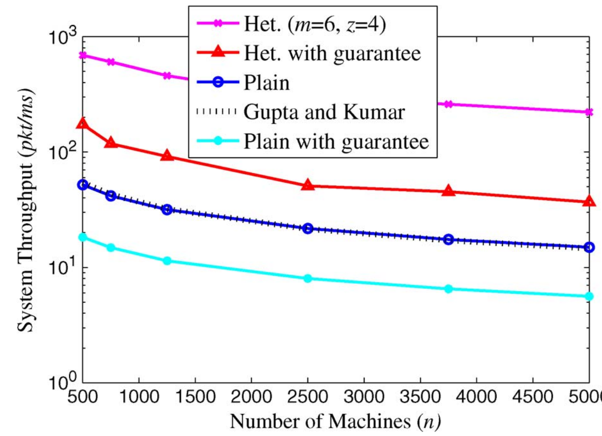

Fig. 8. Comprehensive throughput comparisons for Gupta's results [23], plain machine swarm (Plain), and heterogeneous network architecture (Het.).

TABLE II
METROPOLIS SIMULATION PARAMETERS AND VALUES

Communication Type	Transmission Technology	Average Data Rate
M2M	Zigbee (802.15.4)	$100kbps$
Machine-to-DA	Wifi ($802.11a/b/g/n/ah$)	$1Mbps$
DA-to-DA	$3.5/4G$: LTE (802.16)	$10Mbps$

D. Simulation Validation in the Metropolis

A metropolis is an extremely large city normally within the area of hundreds of km^2 that sets up with several blocks for varies purposes, e.g., business, industries, and residence. Each block consists of millions machines to support human's daily life. To simulate upon this real-world scenario with the proposed heterogeneous architecture in Fig. 1(b), each grid in the figure can serve as a block in Metropolis and each end-to-end data transportation includes three types of communications. In particular, those are the M2M communication with low data rate and energy cost, the machine-to-DA communication with medium data rate, and the DA-to-DA communication with high data rate. We adopt the related values from [3] as shown in Table II and set up the experiment as follows. The 1 Mb data is sent from the source machine to the destination machine in both plain machine swarm and heterogeneous architecture separately. Moreover, DAs' communication capabilities are characterized as the number of machines z that can be served simultaneously by each single DA.

Fig. 9 shows the optimal required number of DAs for heterogeneous architecture with respect to the number of machines. As the DA's capability linearly increases, the required number of DAs drops exponentially. It suggests that few powerful DAs are preferable than bunch of DAs with limited capability. Furthermore, Fig. 10 shows the average end-to-end delay with respect to different area sizes of Metropolis. As the area size increases (so does the number of machines in each block), the heterogeneous architecture supports much less traffic delay than the plain machine swarm. For example, with the area size 60 km^2 and 10^8 machines, the delay from heterogeneous architecture is 115 s as compared to 2,500 s from the swarm. Moreover, the linear curves in the log scale of Fig. 10(b) confirms our asymptotic results, and suggest that the heterogeneous architecture outperforms the plain machine swarm

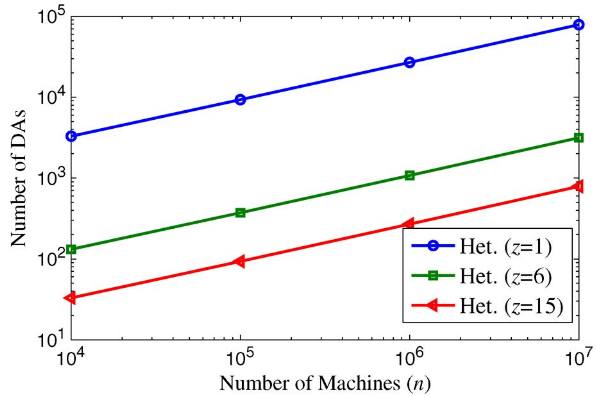

Fig. 9. The optimal number of DAs with respect to the number of machines and DAs' communication capabilities z.

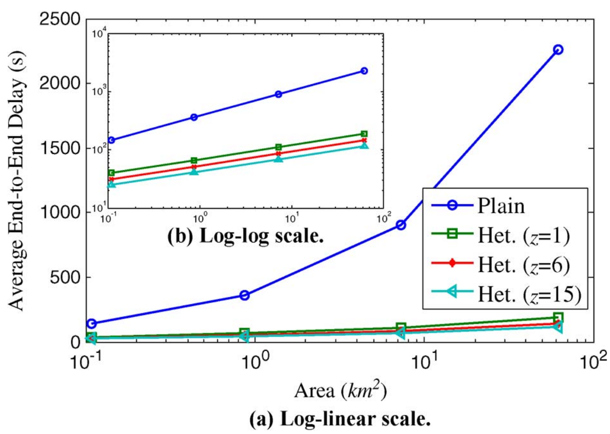

Fig. 10. Average end-to-end delay of plain machine swarm (Plain) and various heterogeneous architecture (Het.) with optimal number of DAs and different communication capabilities z.

with about 95% delay reduction for 10 billion machines. To conclude, by efficiently connecting few DAs to construct small-world shortcuts, proposed statistical control accompanied with heterogeneous architecture resolves the undependable end-to-end transmissions via asymptotically $\Theta(\sqrt[4]{n/\log n})$ delay and $\Theta(1/\sqrt[4]{n(\log n)^3})$ throughput, thus fulfilling the reliable information dissemination in large M2M communication networks.

VII. Conclusion

In this paper, we resolve the most critical challenge on providing statistical control for reliable information dissemination over large M2M communication networks. Examining network topology of M2M networks, the geometric properties of such large networks are well studied to analytically characterize message delivery over connected M2M networks. Moreover, by leveraging queuing network model, the practical data transportation is employed and both the average end-to-end delay and maximum achievable throughput for these connected networks are accessible. Based on above explorations, the promising statistical control with sophisticated small-world network of data aggregators and thus the heterogeneous architecture are proposed to establish shortcut paths among machine communications. Performance evaluation verifies that instead of exploiting long concatenation of multi-hop transmissions in the machine swarm, our heterogeneous network architecture enables machines to communicate through overlaid ultra-fast "highway", like shortcut in small-world networks, with desired throughput. It is particularly crucial for next-generation networks of tremendous amounts of machines. Therefore, we successfully achieve reliable communications via our proposed methodology and facilitate novel traffic control in M2M communication networks.

Appendix A
The Proof of Theorem 1

Partitioning the unit area into square grids with area $(r/\sqrt{5})^2$ as from *Lemma 1* and *Corollary 1*, each node in a square grid is connected to its neighbor nodes within four adjacent square grids. The center of bottom left square grid is at $(r/2\sqrt{5}, r/2\sqrt{5})$ and the center of top right is at $(1 - r/2\sqrt{5}, 1 - r/2\sqrt{5})$. Hence, the Euclidean distance has lower bound $\sqrt{2}(1-2r/\sqrt{5})$. With maximum step length r for each hop, the lower bound of $R(n)$ is given as $\sqrt{2}/r - 2\sqrt{2}/\sqrt{5} \approx \sqrt{2}/r$. On the other hand, the upper bound of $R(n)$ can be obtained as follows. While each node in the square grid is able to connect with nodes in its four adjacent square grids and there is at least one node in each square grid, it exists a path from bottom left square grid, passing bottom right square grid, to top right square grid. As the side of square grids $r/\sqrt{5}$, this path length is $2\sqrt{5}/r$ and gives the upper bound.

Appendix B
The Proof of Lemma 2

Partitioning the unit area into square grids with area $(r/\sqrt{5})^2$ as usual and substituting each square grid with a single nodes, these nodes are connected and form a lattice structure from *Corollary 1*. For an arbitrary path with length d in the lattice graph, there must be a corresponding path, satisfying the upper bound $\sqrt{5}(d_x + d_y)/r$, in $GRG(n,r)$. On the other hand, while the maximum step length is r for each hop, d follows the lower bound and therefore we end the proof.

Appendix C
The Proof of Theorem 2

For $GRG(n,r)$, let y be a random chosen node with position X_n and $x_1, x_2, \ldots, x_{n-1}$ be the rest of nodes with positions $X_1, X_2, \ldots, X_{n-1}$, we have $X_1, X_2, \ldots, X_n \sim U([0,1] \times [0,1])$. We further let the projections of these n nodes on x axis be $Y_1, Y_2, \ldots, Y_n \sim U([0,1])$. For $d(n)$ of M2M network, we have the lower bound $(\|X_n - X_1\| + \|X_n - X_2\| + \ldots + \|X_n - X_{n-1}\|)/[r(n-1)]$ and the upper bound $[2\sqrt{5}(|Y_n - Y_1| + |Y_n - Y_2| + \ldots + |Y_n - Y_{n-1}|)]/[r(n-1)]$ from *Lemma 2*. By law of large number, $(\|X_n - X_1\| + \ldots + \|X_n - X_{n-1}\|)/(n-1) \to \mathbf{E}[\|X_n - X_1\|]$ and $(|Y_n - Y_1| + \ldots + |Y_n - Y_{n-1}|)/(n-1) \to \mathbf{E}[|Y_n - Y_1|]$. From *Proposition 1*, two bounds are obtained to end the proof.

APPENDIX D
THE PROOF OF LEMMA 3

Let s denote the number relay nodes that forward a packet before reaching the destination. $\mathbf{Pr}\{s=l\}=(1-p(n))^{l-1}p(n)$ for $l\geq 1$ and $d(n)=\mathbf{E}[s]=\sum_{l=1}^{\infty} l(1-p(n))^{l-1}p(n)=1/p(n)$.

APPENDIX E
THE PROOF OF LEMMA 5

From diffusion approximation method [42], the visit ratio of a node in a queuing network is defined as the average number of times a packet is forwarded by (i.e. visit) the node. For node i, $e_i=p_{0i}(n)+\sum_{j=1}^{n} p_{ji}(n)e_j$ where p_{0i} denotes the probability that a packet enters the queuing network from node i. While the packets arrive at each node according i.i.d. Poisson process, $p_{0i}(n)$ (i.e. also the probability that node i generates the new packet) equals to $1/n$. Applying this and $p_{ji}(n)$ from *Lemma 4*, we have $e_i=1/n+\sum_{j=1}^{K_i^{(2)}} e_j(1-p(n))/K_i=1/n+e_i(1-p(n))\sum_{j=1}^{K_i^{(2)}} 1/K_i$ where the last equality comes from the assumption of symmetry [46] (i.e. $e_i=e_j,\ \forall 1\leq i;j\leq n$). Furthermore, with the aid of *Proposition 2*, we conclude with $e_i=1/n+[e_i(1-p(n))]/\delta$.

APPENDIX F
THE PROOF OF THEOREM 3

Let D_i denote the average packet delay at node i. According to Little's Formula, $D_i=L_i/\lambda_i=[\rho_i(\delta-1+p(n))]/[\lambda\delta(1-\hat{\rho}_i)]$. By symmetry, the average packet delay at all nodes is the same. Thus, the average end-to-end delay equals to D_i times the average distance (in terms of hop-count) between the source and destination nodes, i.e. $D(n)=d(n)D_i$, which leads to (3).

APPENDIX G
THE PROOF OF LEMMA 7

The upper bound is a direct result from re-scaling *Theorem 2*. For the lower bound, we assume the link to the lattice is at the boundary of each grid, while $z=1$. The average Euclidean distance to that link then has the lower bound as $m^4\int_0^{1/m}\int_0^{v}\int_0^{u}\int_0^{2\pi} r^2 d\theta dr du dv=\pi/\{120\mathrm{m}\}\leq A/B$ where $A=\int_0^{1/m}\int_0^{1/m}\int_0^{1/m}\int_0^{1/m}\min\{u,v,1/m-u,1/m-v,\sqrt{(x-u)^2+(y-v)^2}\}dudvdxdy$ and $B=\int_0^{1/m}\int_0^{1/m}\int_0^{1/m}\int_0^{1/m} dudvdxdy$. With another re-scaling again, the lower bound is obtained for general z (i.e. $z\geq 1$) and ends the proof.

APPENDIX H
THE PROOF OF THEOREM 5

As usual, we examine $z=1$ case and then obtain results for general z from a re-scaling. Considering the lower bound of $d^S(n)$, we first prove that the minimal distance of almost all pairs of nodes comes from the assistance of upper DA's lattice as follows. While the maximum distance of node pairs through upper lattice is $4\sqrt{5}/mr+2m$ from a similar approach as in *Lemma 7*, the coverage area with such distance for machine swarm is $\pi(2mr+4\sqrt{5}/m)^2$ and tends to zero when $m\to\infty$ and $mr\to 0$. Thus, for $(1-\varepsilon)n(n-1)/2$ node pairs with any small $\varepsilon>0$, the distance of these pairs can be provided by the minimal distance coming from the assistance of upper DA's lattice. From *Lemma 7*, both bounds of $d^S(n)$ are thus obtained from the summation of the average distance to DA (from the closet uplink for lower bound) and the average distance in DA's lattice network.

APPENDIX I
THE PROOF OF THEOREM 6

Firstly, $D^L(n)$ is directly obtained from *Theorem 3* as absorption probability $p(n)$ equals to $1/d^L(n)$. On the other hand, let D_i^U denote the average packet delay at DA i. From Little's Formula, $D_i^U=3\rho_i^U m/[2n\lambda(1-\hat{\rho}_i^U)]$. By symmetry, the average packet delay at all DAs is the same. $D^U(n)$ equals to D_i^U times the average distance in DA's network (i.e. $2m/3$) and thus ends the proof.

APPENDIX J
THE PROOF OF THEOREM 7

Similar to *Theorem 4* and *Proposition 3*, λ_{max}^S is obtained from restricting finite delay in nodes and DAs, while λ_{max}^{SQ} is acquired by employing Markov inequality with $D^S(n)$.

REFERENCES

[1] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. Johnson, "M2M: From mobile to embedded Internet," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 36–43, Apr. 2011.

[2] S. Lien, K. C. Chen, and Y. Lin, "Toward ubiquotous massive access in 3GPP machine-to-machine communications in 3GPP," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 66–74, Apr. 2011.

[3] 3rd Gen. Partnership Project (3GPP). Study on Facilitating Machine to Machine, Sophia Antipolis Cedex, France, TR 22.868. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/22_series/22.868/22868-800.zip

[4] K. C. Chen and S. Y. Lien, "Machine-to-machine communications: Technologies and challenges," *Ad Hoc Netw.*, vol. 18, pp. 3–23, Jul. 2014.

[5] A. Zemlianov and G. de Veciana, "Capacity of *ad hoc* wireless networks with infrastructure support," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 3, pp. 657–667, Mar. 2005.

[6] U. Niesen, D. Shah, and G. W. Wornell, "Caching in wireless networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6524–6540, Oct. 2012.

[7] M. E. J. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, 2003.

[8] H. Sundaram, Y. R. Lin, M. D. Choudhury, and A. Kelliher, "Understanding community dynamics in online social networks: A multidisciplinary review," *IEEE Signal Process. Mag.*, vol. 29, no. 2, pp. 33–40, Mar. 2012.

[9] M. Gjoka, M. Kurant, C. T. Butts, and A. Markopoulou, "Practical recommendations on crawling online social networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 9, pp. 1872–1892, Oct. 2011.

[10] M. Gjoka, C. T. Butts, M. Kurant, and A. Markopoulou, "Multigraph sampling of online social networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 9, pp. 1893–1905, Oct. 2011.

[11] B. Mathieu, P. Truong, W. You, and J. F. Peltier, "Information-centric networking: A natural design for social network applications," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 44–51, Jul. 2012.

[12] I. Roussaki *et al.*, "Context-awareness in wireless and mobile computing revisited to embrace social networking," *IEEE Commun. Mag.*, vol. 50, no. 6, pp. 74–81, Jun. 2012.

[13] M. R. Schurgot, C. Comaniciu, and K. J. Runser, "Beyond traditional dtn routing: Social networks for opportunistic communication," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 155–162, Jul. 2012.

[14] J. M. Kleinberg, "Navigation in the small world," *Nature*, vol. 406, no. 6798, p. 845, Aug. 2000.

[15] K. C. Chen, M. Chiang, and H. V. Poor, "From technological networks to social networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 548–572, Sep. 2013.

[16] X. Ta, G. Mao, and B. D. O. Anderson, "On the properties of giant component in wireless multi-hop networks," in *Proc. IEEE INFOCOM*, Apr. 2009, pp. 2556–2560.

[17] D. M. Shila, Y. Cheng, and T. Anjali, "Throughput and delay analysis of hybrid wireless networks with multi-hop uplinks," in *Proc. IEEE INFOCOM*, Apr. 2011, pp. 1476–1484.

[18] D. Braginsky and D. Estrin, "Rumor routing algorithm for sensor networks," in *Proc. ACM WSNA*, Sep. 2002, pp. 22–31.

[19] J. Luo, P. T. Eugster, and J.-P. Hubaux, "Route driven gossip: Probabilistic reliable multicast in *ad hoc* networks," in *Proc. IEEE INFOCOM*, Apr. 2003, pp. 2229–2239.

[20] E. M. Daly and M. Haahr, "Social network analysis for information flow in disconnected delay-tolerant manet," *IEEE Trans. Mobile Comput.*, vol. 8, no. 5, pp. 606–621, May 2009.

[21] P. Y. Chen and K. C. Chen, "Information epidemics in complex networks with opportunistic links and dynamic topology," in *Proc. IEEE GLOBECOM*, Dec. 2010, pp. 1–6.

[22] S. Milgram, "The small world problems," *Psychology Today*, vol. 1, no. 1, pp. 61–67, May 1967.

[23] P. Gupta and P. R. Kumar, "Capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[24] S. C. Lin, L. Gu, and K. C. Chen, "Providing statistical qos guarantees in large cognitive machine-to-machine networks," in *Proc. IEEE GLOBECOM*, Dec. 2012, pp. 1–6.

[25] L. Gu, S. C. Lin, and K. C. Chen, "Small-world networks empowered large machine-to-machine communications," in *Proc. IEEE WCNC*, 2013, pp. 1–6.

[26] F. M. Lima, T. F. Maciel, W. C. Freitas, and F. R. P. Cavalcanti, "Resource assignment for rate maximization with qos guarantees in multiservice wireless systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 3, pp. 1318–1332, Mar. 2012.

[27] L. Saker, S. E. Elayoubi, R. Combes, and T. Chahed, "Optimal control of wake up mechanisms of femtocells in heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 664–672, Apr. 2012.

[28] J. Abouei, A. Bayesteh, and A. K. Khandani, "On the delay-throughput tradeoff in distributed wireless networks," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2159–2174, Apr. 2012.

[29] H. Zhang, Y. Ma, D. Yuan, and H. H. Chen, "Quality-of-service driven power and sub-carrier allocation policy for vehicular communication networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 1, pp. 197–206, Jan. 2011.

[30] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an undrlay to LTE-Advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[31] H. S. Dhillon, H. C. Huang, H. Viswanathan, and R. A. Valenzuela, "Power-efficient system design for cellular-based machine-to-machine communications," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5740–5753, Nov. 2013.

[32] X. Lin and H. Viswanathan, "Dynamic spectrum refarming with overlay for legacy devices," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5282–5293, Oct. 2013.

[33] S. Y. Lien and K. C. Chen, "Massive access management for qos guarantees in 3gpp machine-to-machine communications," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 311–313, Mar. 2011.

[34] T. Kwon and J. M. Cioffi, "Random deployment of data collectors for serving randomly-located sensors," *IEEE Trans. Wireless Commun.*, vol. 12, no. 6, pp. 2556–2565, Jun. 2013.

[35] Y. Zhang *et al.*, "Cognitive machine-to-machine communications: Visions and potentials for the smart grid," *IEEE Netw.*, vol. 26, no. 3, pp. 6–13, May/Jun. 2012.

[36] G. R. Gupta and N. B. Shroff, "Delay analysis and optimality of scheduling policies for multi-hop wireless networks," *IEEE/ACM Trans. Netw.*, vol. 19, no. 1, pp. 129–141, Feb. 2011.

[37] P. K. Huang, X. Lin, and C. C. Wang, "A low-complexity congestion control and scheduling algorithm for multi-hop wireless networks with order-optimal per-flow delay," *IEEE/ACM Trans. Netw.*, vol. 21, no. 2, pp. 495–508, Apr. 2013.

[38] Y. Durmus, A. Ozgovde, and C. Ersoy, "Distributed and online fair resource management in video surveillance sensor networks," *IEEE Trans. Mobile Comput.*, vol. 11, no. 5, pp. 835–848, May 2012.

[39] Y. Shi, Y. T. Hou, S. Kompella, and H. D. Sherali, "Maximizing capacity in multi-hop cognitive radio networks under the sinr model," *IEEE Trans. Mobile Comput.*, vol. 10, no. 7, pp. 954–967, Jul. 2011.

[40] P. Erdös and A. Rényi, "On the evolution of random graphs," *Pub. Math. Inst. Hungarian Academy Sci.*, vol. 5, no. 1, pp. 17–61, Sep. 1960.

[41] M. D. Penrose, *Random Geometric Graphs*. London, U.K.: Oxford Univ. Press, 2003, ser. Oxford Studies in Probability.

[42] G. Bolch, S. Greiner, H. D. Meer, and K. S. Trivedi, *Queueing Networks and Markov Chains*. New York, NY, USA: Wiley, 1998.

[43] D. J. Watts and S. H. Strogatz, "Collective dynamics of small world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998.

[44] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*., 2nd ed. New York, NY, USA: MIT Press, 2001.

[45] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 630–643, Jul. 2003.

[46] D. Q. Qian and X. D. Zhang (2011, Jul.). Potential distribution on random electrical networks. *Acta Mathematicae Applicatae Sinica (English Series)* [Online]. *27(3)*, pp. 549–559. Available: http://dx.doi.org/10.1007/s10255-011-0091-9

Shih-Chun Lin (S'06) received the B.S. degree in electrical engineering and the M.S. degree in communication engineering from National Taiwan University in 2008 and 2010, respectively. He is currently with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. His research interests include large machine-to-machine communication, wireless underground sensor networks, software-defined networking, and statistical scheduling in wireless systems.

Lei Gu (M'13) received the B.S. degree in mathematics and the Ph.D. degree in applied mathematics both from Shanghai JiaoTong University, Shanghai, China, in 2006 and 2011, respectively. From 2011 to 2012, he was a Postdoctoral Research Associate with Intel-NTU Connected Context Computing Center, Taipei, Taiwan, ROC. From 2011 to 2013, he was a Postdoctoral Research Associate with Fudan Univerysity, Shanghai, China. Since 2013, he has been with China Telecom Cooperation Shanghai Research Institute, Shanghai, China, as a Senior Technical Staff Member. His research interests include analysis of complex networks, networking and wireless communications, machine-to-machine communication and smart cities.

Kwang-Cheng Chen (M'89–SM'94–F'07) received the B.S. degree from the National Taiwan University in 1983, and the M.S. and Ph.D degrees from the University of Maryland, College Park, MD, USA, in 1987 and 1989, all in electrical engineering. From 1987 to 1998, he worked with SSE, COMSAT, IBM Thomas J. Watson Research Center, and National Tsing Hua University, in mobile communications and networks. Since 1998, he has been with National Taiwan University, Taipei, Taiwan, ROC, and is the Distinguished Professor and Associate Dean for academic affairs at the College of Electrical Engineering and Computer Science, National Taiwan University. His recent research interests include wireless communications, network science, and data analytics. He has been actively involved in the organization of various IEEE conferences as General/TPC chair/co-chair, and has served in editorships with a few IEEE journals. He also actively participates in and has contributed essential technology to various IEEE 802, Bluetooth, and LTE and LTE-A wireless standards. He has authored and co-authored over 200 IEEE papers and near 30 granted US patents. He co-edited (with R. DeMarca) the book *Mobile WiMAX* (Wiley, 2008) and authored the book *Principles of Communications* (River, 2009), and co-authored (with R. Prasad) another book *Cognitive Radio Networks* (Wiley, 2009). He is an IEEE Fellow and has received a number of awards including the 2011 IEEE COMSOC WTC Recognition Award, and has co-authored a few award-winning papers published in IEEE journals and conferences, including the 2014 IEEE Jack Neubauer Memorial Award and the 2014 IEEE COMSOC AP Outstanding Paper Award.