<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *IEEE Communications Magazine*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Wang, X., Ning, Z., Hu, X., Ngai, E., Wang, L. et al. (2018)
A City-Wide Real-Time Traffic Management System: Enabling Crowdsensing in Social Internet of Vehicles
IEEE Communications Magazine, 56(9): 19-25
<https://doi.org/10.1109/MCOM.2018.1701065>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-366229>

Intelligent Network Management for Internet-of-Things in Smart Cities

Xiuming Liu

Department of Information Technology
Uppsala University
Uppsala, Sweden
Email: xiuming.liu@it.uu.se

Edith C.H. Ngai

Department of Information Technology
Uppsala University
Uppsala, Sweden
Email: edith.ngai@it.uu.se

Abstract—Building an intelligent, sustainable, and safe city is a global mission for both research communities and general public. Recent advance of technologies, such as Internet-of-Things (IoT), cloud computing, and machine learning, enables new solutions and applications, such as intelligent traffic system, urban environment monitoring, and smart energy systems. However, robust and intelligent network management on such a large scale becomes a great challenge. In this paper, we present a machine learning assisted network management framework for IoT in smart cities. This intelligent framework seamlessly integrates machine learning with sensing and communication, information fusion, and decision making in the city network architecture. We present a case study of vehicular sensing network for urban environment monitoring, in which the services are supported by various machine learning techniques. A novel distributed information fusion scheme that is robust to false data is then introduced. Finally we discuss scalable computation and implementation of machine learning methods in large scale networks.

I. INTRODUCTION

The recent advance of sensing, communication, and computing technologies have enabled great opportunities for further and deeper integration of the information systems and our physical society. It was predicted a decade ago that the machines to machine (M2M) communications, instead of phones or tablets, will take over the majority in the connected devices. This forecast has become the reality and even will be boosted by the internet-of-things (IoT) in the foreseeable future [1]. On the other hand, the largest physical systems human have ever created, cities, are becoming very large intelligent networks which consists of connected buildings, vehicles, and many other small devices with various sensing, communication, and computing capabilities [2].

Smart cities rely on the collection and analysis of big data generated from numerous IoT devices and smart systems (e.g. ICT infrastructures, smart grids, vehicular networks) to communicate and manage resources efficiently. Artificial intelligence could play a key role in numerous applications for smart cities, from air pollution monitoring [3] to smart grids [4], traffic control, and self-driving cars [5]. Machine learning techniques allow us to gain insight from large amount of data to support real-time monitoring, autonomous control, decision-making, and long-term urban planning for smart cities. Intelligent network management also enables efficient

and autonomous control on the smart system itself to improve data quality, resilience, and performance.

The continuing integration of information and physical systems have already presented many critical challenges for researchers, industries, and general public. Focusing on the network management, the challenges emerge from the aspects of service quality, resource efficiency, and robustness towards to cyber-physical attacks [6]. Consider a vehicular sensing network [7] for urban monitoring, its mobile, delay-sensitive and large-scale characteristics require seamless integration of peer-to-peer wireless communications, distributed information fusion and decision making (e.g., path planning). In this context, the quality of the monitoring service is not only affected by the sensing layer, but also the communication and network management layers [8]. In addition, the requirement of resource efficiency arises from the fact that the shared radio spectrum and the on-board computation power are limited by a feasible budget. The decentralized network topology may also make individual agents more vulnerable to false data injection given that global information may not be available in a fully distributed system.

In this paper, we propose an intelligent network management framework for IoT in smart city, assisted by statistical machine learning techniques. We emphasize that the machine learning techniques used in network management shall be interpretable and transparent for human users, and also be scalable for deployment in large networks with “small” (in terms of computation resources) nodes. The framework provides an unifying view of information gain from sensing and communication, and seamlessly integrates with Bayesian information fusion for autonomous decision making that is aware of data quality and robust against false data attacks. On the decision making level, we discuss how machine learning can be applied to solve problems such as resource allocation and path planning. We demonstrate this intelligent network management framework with a case study of vehicular sensing networks in smart cities. Finally we give a few tips on how the presented machine learning algorithms can be implemented in a scalable manner in order to be deployed on large scale networks.

Fig. 1. An illustration of the system architecture of intelligent network management for IoT in smart city. The urban environment and society are divided into three levels: urban, block, and node. IT infrastructures and networks with various capabilities are deployed, ranging from wearable and on-board devices to edge servers and cloud. On each level, there are four interactive functionalities and services: sensing, communication, information fusion, and decision making. Considering various applications, computation capabilities and communication capacities on three levels, different network management strategies with suitable machine learning methods are implemented.

II. INTELLIGENT NETWORK FOR SMART CITIES

An integrated architecture for IoT in smart city is presented in Figure 1. The intelligent network architecture in urban area can be divided three physical layers, where every layer is capable to providing sensing and communication, information fusion, and decision making services. On the lowest level, such as vehicles, building, and personal devices, the sensors are able to collect detailed information of the physical environment, but with very limited computation resources and communication capacity. On the edge level, the increased computation resources enable more complex tasks such as coordination of nearby vehicles, with low communication latency. On the urban level, large scale data fusion and decision support services are available for professional users such as municipalities and public service companies.

A. Sensing and Communication

Consider a multidimensional dynamic process $\mathbf{x}_t = f(\mathbf{x}_{0:t}) \in \mathcal{X} \subseteq \mathbb{R}^D$. A network of sensors are deployed to obtain measurements $\mathbf{y}_t \in \mathcal{Y} \subseteq \mathbb{R}^M$. In many scenarios, only a small set of sub-dimensions of the dynamic process is monitored, i.e., $M \ll D$. For example, in the application of urban environment monitoring, the measurements are taken in a set of locations which is very small comparing to the entire urban area.

The sensing process at a node m of the network can be model with a parametrized and probabilistic function, $y_t^m \sim h^m(\mathbf{x}_t | \boldsymbol{\theta}^m)$, where the set of parameters $\boldsymbol{\theta}$ are estimated during the calibration process. Given that the IoT deployed in urban area can be heterogeneous in terms of operation & maintenance (O&M), the parameters $\boldsymbol{\theta}^m$ might have different

dispersion across the network [9]. Therefore, each data y_t^m is associated with an uncertainty measure (e.g., variance) in order to be weighted properly in the later process of information fusion.

On the other hand, the wireless communication technologies for IoT in urban environment are often designed to be lower power (short range) and limited capacity, such as 6LoWPAN and NB-IoT [10]. The dynamic network topology in large vehicular networks makes it even more difficult to disseminate data through entire networks via multi-hop communications. Therefore, a node would preferably communicated small size messages with its neighbours.

B. Information Fusion

Given that the global information is limited for a node to access in a fully distributed network, it is important to combine local information in order to refine a node's knowledge about the dynamic process. Information or data fusion can be applied on different levels: when the raw sensor data are exchanged and combined, we usually refer it as sensor fusion; when the local knowledge of individual nodes are exchanged and combined, we refer it as information fusion. Considering the limited communication capacity, it is often preferred to exchange the compressed knowledge (e.g., probability distributions) rather than raw sensor data, especially when the data type is audio, image, or even video. Another advantage of perform information fusion based on the latest local knowledge is that, the local knowledge is a result of all historical data a node and its neighbours have been received. Therefore a node only keeps its latest local knowledge and do not need to exchange and cache historical data from its neighbours.

Considering the uneven sensor quality and increasing risk of cyber-physical attacks, such as the false data injection (FDI) [11], information fusion for IoT in smart city need to be robust and resilient. There are two different modes for a node to handle a message from its neighbour: if the message is from a neighbour with low sensor quality, the information should be weighted properly when use it to updated the local knowledge; if the message is from a corrupted neighbour, then it should be discarded. Nevertheless, it is a non-trivial task to identify which neighbours are under attack. In the later section we will discuss emerging machine learning techniques for detecting FDI.

C. Intelligent Network Management and Applications

For intelligent network management, the general task is to monitor and maintain good service and data quality, support device management and network configuration autonomously. For IoT in smart cities, intelligent network management is essential to support emerging applications and improve their performances. We list a few examples here.

Urban environmental monitoring: The heterogeneity of IoT devices and the variety of operator makes it difficult to control the life-cycles for every connected things. However, from the data and service quality's perspective, it is possible and important to manage the "virtual" life-cycle of IoT devices, defined by the time period when they can provide data or services with satisfying quality.

Consider an air pollution sensor network deployed in urban environment, most of the low-cost sensors are based on electrochemical materials, which have limited time period (e.g., two years) that they can provide measurement with controllable errors (zero drifting and dispersion of coefficients). As time goes by, the electrochemical materials on sensing devices are consumed by the air pollutants, and therefore the measurements become inaccurate. When the meta-data of the IoT devices are registered in the network, the measurements come from low quality devices can be calibrated online and treated with less confidence in the process of information fusion supported by intelligent network management.

Intelligent network management also enables autonomous power-saving on the IoT devices. Based on the data and communication quality, the network can automatically configure itself to turn on and off, or change the sampling rates of the sensors for saving energy. Consider vehicles or drones to be used for collecting urban data, the mobility of sensing devices enables wider coverage but requires intelligent and autonomous management. For instance, optimal paths of the mobile devices are desired for collecting the most informative measurements with minimum cost (e.g., travelling distance). Considering the connectivity constraints and distributed information fusion, a cross-layer intelligent framework is required for such a complicated optimization problem.

Smart energy systems: Smart energy systems are involved in energy generation, transmission, distribution, and consumption processes for power grids, buildings, and transportation systems. Modern smart energy systems promote the integration

of distributed and renewable energy sources, such as wind, solar, nuclear, etc. Due to the unpredictability issues associated with the resource availability, it is challenging to match the energy generation with the consumption. Intelligent network management enables real-time monitoring and control of the energy systems, which makes renewable energy generation and management more efficient. By analyzing the IoT data in real-time, it is possible to predict energy generation and consumption, and give real-time control through the network by combining different energy sources or energy storage to make the supply and demand more balance and cost-effective.

Autonomous vehicles: Autonomous vehicles are considered as common long-term vision for smart cities. Intelligent network management with IoT have already made smart parking, smart lighting, and adaptive signal control possible. For the future, vehicles will be able to communicate with each other as well as with the road infrastructures. All the traffic infrastructures and entities will be connected and managed by intelligent network, which can coordinate traffic for the best of road efficiency and safety. For example, traffic congestion can be mitigated by diverging traffic through coordinating the vehicles and imposing different traffic policies adaptively. Intelligent network can also support real-time communication and coordination of vehicles to avoid traffic accidents and give early warnings to potential threats.

In terms of communications, mobility such as connected vehicles or mobile devices carried by pedestrians pose new challenge for intelligent network management. For example, when a mobile device moves from one cell to another, the handover process is required to ensure seamless connection. Machine learning techniques can be applied to the network for predicting user mobility pattern and therefore leading to less service quality degradation caused by the handover process.

III. CASE STUDY: VEHICULAR SENSING NETWORK FOR URBAN ENVIRONMENT MONITORING

Machine learning and artificial intelligence is not a new concept for communication society. In fact, Claude Shannon's pioneering work about information theory founds the basis for many modern statistical machine learning algorithms [12]. In this section, we discuss how different problems in sensing, communication, and network management layers can be unified from a information theoretic view in order to improve the service quality for IoT in smart city. We take the environment monitoring with vehicular sensing networks as an exemplary problem. In this problem, the task is to monitor a spatial-temporal dynamic process over the urban area with a distributed and mobile sensor network.

A. Sensing, Communication, and Information Gain

In the vehicular sensing network, the decision making need to be real-time and most of current vehicle-to-vehicle (V2V) communication technologies are implemented in a peer-to-peer manner to ensure low latency. Therefore, we consider a decision support service implemented on a node level. The information about the dynamic process comes from two

Fig. 2. An example of the KLD based hierarchical clustering for detecting false data injection attacks.

sources: the local sensor measurement and the neighbour's knowledge about the dynamic process. By taking a local measurement or receiving a message from its neighbours, a vehicular node obtains a piece of information regarding the current states of the dynamic process.

The information gain of sensing and communication can therefore be measured by the mutual information between the sensor data and the dynamic process, $I(y_t^m, x_t)$, and the Kullback-Leibler divergence (KLD) [12] between a local knowledge (a probability distribution $p^m(x_t)$ of the dynamic process) and the neighbour's knowledge $p^n(x_t)$, $D_{KL}(p^m, p^n)$. The KLD is a non-negative measure of difference between two probability distributions, the large KLD indicates more information gain. This is also known as the maximum entropy principle in machine learning, as the relative entropy is another name for the KLD. When the two distribution are identical, the KLD is minimized to zero, indicating no information gain.

Therefore, for a vehicular network operated on distributed wireless communication with limited ranges, a trade off between collecting more informative measurements (moving away from the rest of the network) and receiving more information from neighbours (moving towards to the rest of the network). We will revisit this trade off in later sections.

B. Bayesian Information Fusion

For a node in the vehicular network, it constantly updates its knowledge of the dynamic process by combining the latest local sensor data and the knowledges from its neighbours. This task can be implemented as the Bayesian information fusion.

Bayes' theorem gives the posterior distribution based on both data and the prior distribution. From the belief updating's perspective [13], the posterior distribution $p^m(x_t|y_{0:t}^m)$ is a result derived by minimizing the average negative log-likelihood and the KLD to the prior distribution:

$$\arg \min_q \int_{\mathcal{X}} -\ln \ell(y_{0:t}^m, x_t) q(x_t) dx_t + D_{KL}(q, \pi^m), \quad (1)$$

where $\ell(y_{0:t}^m, x_t)$ is the likelihood, $\pi^m(x_t)$ is the prior distribution which can be obtained by predicting based on the dynamic process and the knowledge about the previous states.

The next step is to combine the knowledge $p^n(x_t|y_{0:t}^n)$ from neighbours $n \in \mathcal{N}_t^m$, where $\mathcal{N}_t^m$ is the set of neighbours for

node m at time t . Note that the topology is dynamic in the vehicular network, therefore the neighbour set is changing according to time. Following the same belief updating framework above, the refined knowledge about x_t at node m is

$$\arg \min_q w^m D_{KL}(q, p^m) + \sum_{n \in \mathcal{N}_t^m} w^n D_{KL}(q, p^n), \quad (2)$$

where $\{w^m, w^n\}_{n \in \mathcal{N}_t^m}$ are the weights for the local knowledge and the knowledge from neighbours. This information fusion procedure is also known as the KLD averaged consensus filtering [14]. From the information geometry's perspective, the outcome of such average KLD minimization is a probability distribution which has the minimum divergence to all accessed distributions (knowledge). Therefore, the node m formalize a consensus based on its own observation and knowledge from its neighbours.

C. Clustering based False Data Detection

Another advantage of the previous KLD based information fusion is the robustness to the false data injection. Consider one of node m 's neighbour, k , is hijacked by a malicious attacker and its sensor measurement $y_{0:t}^k$ are injected with false data. The knowledge from node k is therefore corrupted and shall not be used by node m in information fusion.

By keeping an record of the pairwise KLDs, the node m is able to construct an averaged KLD matrix and perform hierarchical clustering over the knowledge from its neighbours. An illustrative example of a network with 10 nodes are presented in Figure 2. In this example, nodes 1 and 2 are under FDI attack. Consequently, the posteriors from nodes 1 and 2 are different from the posteriors produced by secure nodes. See Figure 2(a) for a two dimensional example. This difference will be reflected in the pairwise averaged KLDs. We visualize the averaged symmetrized KLDs (sym-KLD) between each pair of nodes in the matrix showed in Figure 2(b), where nodes 1 and nodes 2 have visible higher KLDs toward to the knowledge from the rest of the network. Note that a vehicular node does not need centralized communication network to build this matrix. Thanks to its mobility, a vehicular node only need to update the matrix elements correspond to its neighbours at current time. As time passes by, each node gradually completes a stable averaged KLD matrix.

The averaged KLD matrix can be used for hierarchical clustering to detect the nodes under FDI attacks. Figure 2(c) shows the dendrogram of the minimum distance linkages based on the KLD matrix. As we observe, nodes 1 and 2 distinguish themselves with higher minimum symmetrised KLDs. The theory behind this clustering based FDI detection finds its root in statistical hypothesis testing. Given a dynamic process with correlations between dimensions, the KLDs between each pair of posteriors are random variables which follow certain distributions. For example, when Gaussian random field is used to model the spatial-temporal dynamic process, the KLDs follow the χ^2 distribution (Figure 2(d)). By injecting false data into the, the distributions of KLDs are altered, which enables the detection based on the KLDs clustering.

D. Decision Making Support with Reinforcement Learning

In light of the previous discussion, we now investigate the network management strategy for improving the quality of service. In the example of vehicular sensing network, the quality of service is defined by the accuracy of estimating the dynamic process states $\mathbf{x}_t$. Considering the distributed communication with limited range, we now remark the following observations:

- For collecting an informative measurements, a vehicular node intends to explore the urban area where less nodes are deployed;
- For receiving more knowledge from the neighbours, a vehicular nodes intends to move towards to the rest of the network;
- For quickly detect the FDI attack, a vehicular nodes intends to maintain connections with more neighbours.

The first point emphasizes the incentive for a node to move away from the rest of the network; while the second and third observations highlight the importance of maintain stronger connectivity in the network. Thus, there is indeed a trade-off between exploring the area and staying connected with the rest of the network. The task of path planning for a single vehicle to maximize rewards or minimize costs is known as NP-hard. Here the problem is perhaps more difficult in the sense that the nodes in the network need to cooperate. A promising solution is to apply the reinforcement learning. Reinforcement learning is a set of methods to solve the sequential decision problem. The root of reinforcement learning can be traced back to the dynamic programming in 18th century, then famously reformulated and developed by Bellman in the middle of 20th century. Many recent successful stories have show that the continuously developed RL methods are capable of solving decision making problems in complex networked systems [15].

IV. SCALABLE COMPUTATION AND IMPLEMENTATION

While machine learning shows promising capabilities to solve difficult problems with exciting performance, a major concern is about its scalability when applied to large scale networks with limited computation power and runtime budget. In fact, there are plenty of theoretic tools in machine learning for reducing the complexity by applying approximated methods.

Next we discuss a few of them which are suitable for IoT in smart cities.

A. Recursive Learning

One scalability challenge comes from handling large amount of data. Many machine learning algorithms, ranging from the simplest linear regression to non-parametric models such as Gaussian processes, require the inversion operations of the covariance matrix of input data. The matrix inverse has a computational complexity of $\mathcal{O}(N^3)$ where N is the size of the matrix. Various approximation methods can be applied to reduce the complexity such as subset of data, Nyström approximation, and sparse methods. However, most of them require heuristic search for the optimal sub dimensions. In the next we present a recursive method for scalable learning. The recursive learning method follows the same belief updating framework in the previous discussion of Bayesian information fusion. However, instead of using the full likelihood $\ell(y_{0:t}^m, \mathbf{x}_t)$, the composite likelihood methods can be applied. Diving the data sequence $y_{0:t}^m$ into J segments, the composite likelihood is $\prod_{j=1}^J \ell(y_j^m, \mathbf{x}_t)$. Using the composite likelihood in the loss function, we obtain an approximated posterior following the Bayes' theorem. This method is also known as the batch Bayesian learning. Furthermore, the batch Bayesian method can be implemented in a recursive manner, allowing a node to update its posterior online when there is new data segment available. With the recursive learning, the computation complexity as well as the memory usage is significantly reduced.

Of course, there is no free lunch for reducing the runtime. The approximated posterior will have higher uncertainty (measured by for example the confidence interval) than the exact posterior. In statistical machine learning, the state-of-art result quantifies the extra uncertainty introduced by the composite likelihood method. This theoretic result can be used as a guidance for balancing the trade-off between runtime and service of quality.

B. Ensemble Learning

Another scalability concern arises from applying machine learning methods to large scale networked systems. For instance, in the vehicular sensing network, handling data from a large scale network is expensive in terms of both computation and communication.

Ensemble learning is a paradigm of machine learning when different models are used to produce the inference result together. Examples are mixture of experts and model average methods. Indeed, the previously presented consensus filtering and clustering based FDI detection can also be seen as a type of ensemble learning methods, where each node produce its local knowledge of the dynamic process and then combine the knowledge from neighbours.

The advantage of combining knowledge instead of raw sensor data is that, each node is free to choose its own model and inference method which depends on the available computation resource, energy level, and runtime budget. As long as coherent quantifications of uncertainties are reported

together with the locally produced knowledge, they can be weighted properly during the combining procedure.

V. CONCLUSIONS

In this paper we presented an unifying view of machine learning assisted cross layer network management for IoT in smart city, together with scalable computation and implementation methods. Using the vehicular sensing network as an example, we quantify the information gain of sensing and communication in the network, and present the distributed and secure information fusion method. We discuss an example of detecting FDI attack via hierarchical clustering based on the KLD matrix, and the potential solutions of various network management problems via reinforcement learning. Considering the limited computation, communication, and energy resources, we further presented two scalable machine learning paradigms which can be implemented with the Bayesian information fusion.

As the integration of information systems and our physical society goes further, we believe the future of intelligent networks assisted by machine learning and AI technology is promising. On the other hand, more and more challenges such as data privacy, network security, and energy efficient computing will reveal themselves. Those opportunities and challenges encourage both academic and industries to build more connected, intelligent, and sustainable cities.

REFERENCES

- [1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things journal*, vol. 1, no. 1, pp. 22–32, 2014.
- [2] 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.
- [3] Y. Zheng, F. Liu, and H.-P. Hsieh, "U-air: When urban air quality inference meets big data," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2013, pp. 1436–1444.
- [4] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1644–1652, 2017.
- [5] P. Koopman and M. Wagner, "Autonomous vehicle safety: An interdisciplinary challenge," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 1, pp. 90–96, 2017.
- [6] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems security: A survey," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1802–1831, 2017.
- [7] J. Wang, C. Jiang, K. Zhang, T. Q. Quek, Y. Ren, and L. Hanzo, "Vehicular sensing networks in a smart city: Principles, technologies and applications," *IEEE Wireless Communications*, vol. 25, no. 1, pp. 122–132, 2018.
- [8] Y. Gao, W. Dong, K. Guo, X. Liu, Y. Chen, X. Liu, J. Bu, and C. Chen, "Mosaic: A low-cost mobile sensing system for urban air quality monitoring," in *INFOCOM*, 2016, pp. 1–9.
- [9] B. Maag, Z. Zhou, and L. Thiele, "A survey on sensor calibration in air pollution monitoring deployments," *IEEE Internet of Things Journal*, 2018.
- [10] J. Dizdarevic, F. Carpio, A. Jukan, and X. Masip-Bruin, "Survey of communication protocols for internet-of-things and related challenges of fog and cloud computing integration," *arXiv preprint arXiv:1804.01747*, 2018.
- [11] C. Yang, L. Feng, H. Zhang, S. He, and Z. Shi, "A novel data fusion algorithm to combat false data injection attacks in networked radar systems," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 1, pp. 125–136, 2018.
- [12] D. J. MacKay and D. J. Mac Kay, *Information theory, inference and learning algorithms*. Cambridge university press, 2003.
- [13] P. G. Bissiri, C. C. Holmes, and S. G. Walker, "A general framework for updating belief distributions," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 78, no. 5, pp. 1103–1130, 2016.
- [14] G. Battistelli and L. Chisci, "Kullback–leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability," *Automatica*, vol. 50, no. 3, pp. 707–718, 2014.
- [15] Y. F. Chen, M. Liu, M. Everett, and J. P. How, "Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 285–292.