Bridge Maintenance, Safety, Management, Digitalization and Sustainability –
Jensen, Frangopol & Schmidt (eds)

Simultaneous WET and WIT with Q-learning for wireless bridge sensors network

Muhammad Tauseef Mushtaq, Veronico Giuseppe, Nicola Cordeschi, Alfredo Grieco & Gennaro Boggia
Politecnico di Bari, Bari, Italy

ABSTRACT: Wireless Powered Communication Networks play a crucial role in Internet of Things scenarios which operate with low energy sensors. In this regard, simultaneous wireless energy transmission (WET) and wireless information transmission (WIT) are proposed for continuous data transmission with low energy sensors. These sensors can be deployed on bridges that are hit by disasters. However, these low-energy networks demand a random data transmission and hence require access points to decode the packets accordingly. Hence it is challenging to design a slot selection algorithm that provides a better throughput. On this subject, we present a Q-Learning-based algorithm that harnesses the qualities of deep reinforcement learning and enhances the throughput. In our algorithm, we enhance the slot allocation of each user with the influence of a few physical layer properties. Results show an improved slot allocation mechanism leads to better throughpu

1 INTRODUCTION

Wireless-powered sensor networks are a future technology which enables future smart cities and other industrial systems. When discussing smart cities, it’s essential to consider smart buildings and infrastructures, such as bridges and roads, alongside intelligent autonomous vehicles. This leads us to emphasize the importance of smart-bridges, which serve as critical infrastructure for urban areas. During an unfortunate event or flooding, gathering essential data about bridges is imperative. The information can be catagories in two ways; how much damage bridge have suffered?, and how much water/vehicle load bridge can still manage? It’s important to highlight that routine information gathering for bridges can also be accomplished through these networks. This creates a need to create a portable or remote network that can collect this key information. However, these networks’ performance is limited by the amount of battery a sensor occupies which is deployed on the bridge. Our work is related to WP1 topic of Bridgsite; “Collect Bridge Information”. Our proposed work is related to low-cost sensors with edge computing capacity to optimize the volume of information to transmit for Bridgesite project. Low-cost communication systems require minimum amount of data transfers between a Hybrid Access Point (HAP) and deployed sensors subject to lower utilization of energy and other resources. HAP are responsible for both wireless energy and information transmission. That is why to acquire optimised and up to data low-cost bridge wireless networks require a random access protocol such as Irregular repeated slotted ALOHA (IRSA) or it’s variants to achieve these goals. Later on, our paper is focused on efficient data acquisition from low-cost bridge sensor networks.

That is why recharging (WET) and acquiring information (WIT) are very important. Portable communication scenarios such as disaster-hit areas, industrial Internet of things, or Mission-critical communications rely on the sensors deployed in the field with the ability to harvest energy. Healthcare, smart cities, and transportation are other systems that require

DOI: 10.1201/9781003483755-408

energy harvesting. Hence the idea of Wireless Power Communication (WPCN) has widely spread to maximize throughput using simultaneous WET and WIT. In WPCN a dedicated HAP is responsible for simultaneous energy and information transmission. HAP also ensures that there are no collisions and success full data transmission of each user is possible without any error. literature suggests that the network layer has not enjoyed tremendous improvement during the past while the physical layer has taken a huge leap in catering to such scenarios. The physical layer benefits 5G in indoor scenarios or UWB frequency. Similarly, beamforming also plays a critical role in low-energy networks with its directional signal capabilities. Hence, there needs an improvement on the resource allocation part to rep the fruit of much more efficient sensors and access points. For the scenarios mentioned above, IRSA has played a crucial role. IRSA randomly selects slots for replicas of each user providing no control over access points. Recently, a study has been done to improve random access performance.

This includes the introduction of machine learning algorithms for random access, the introduction of mini-slots, and time-offset. Time Offset (TO) increases the probability of successful full packet transmission by allowing users to send their data with a random offset of a constant length Hu, Y. 2022. Similarly, another method is to introduce the mini-slots before the transmission Ayoub, I 2021, Dumas, C. 2021. Mini-slots are small jamming signals from the users to inform HAP and other users about the slots where they will send replicas. Deep reinforcement learning-based methods also use a Q-learning algorithm to select the slots for each user Li, Y 2022, Choi, H 2019. These slot selection decisions are based on the battery level of each user and hence require an access point to interfere in the decision.

Nonetheless, all the methods lack incorporating the physical qualities of the system such as channel quality, and distance between UE and Access point. Despite the introduction of mini slots, time offset, and ML on the network layer for random access, there is a dire need to make the network layer more aware of the environmental effects. To the best of the authors' knowledge, this type of analysis is not done before. Hence, we present a better and improved decision process based on Q-Learning which incorporates channel quality, battery level, and distance between UE and HAP.

In this paper, we consider a nonlinear energy harvesting scenario with simultaneous capabilities of WET and WIT. It is uncommon to incorporate the effect of efficiency of energy harvesting in the WPCN environment at the network layer. In this regard, most of the decisions involved on the network layer are related to slot selection. Therefore, not only an

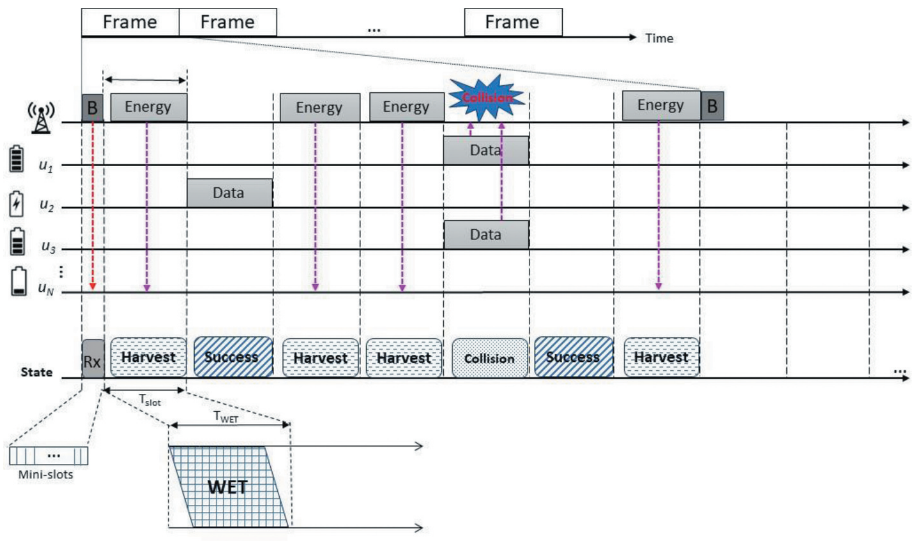

Figure 1. System model.

intelligent slot selection mechanism is required but it should also consider the efficiency of energy harvesting. For this purpose, we introduce a nonlinear energy harvesting model and calculate the harvested energy. By the end of each frame, the network layer has realistic information about the battery level of sensors. Finally, we can exploit this information or intelligent decision-making of slot selection for not only WET but also for WIT.

The rest of the paper is as follows. Section 2 describes the system model of a wireless-powered IoT network. Section 3. presents Q-learning algorithms which can manage the number of replicas along with energy harvesting slots. In section 4, we discuss the results obtained from these algorithms while section 5 concludes the work and provides future suggestions.

2 SYSTEM MODEL

We consider an IoT scenario with energy harvesting capabilities as shown in Figure 1. The system consists of sensors deployed in the field. We name them as $u_1, u_2, \ldots, u_N$. As per the requirement of IoT systems, sensors are always in saturation mode, which means they always have data to transmit. However, sensors' ability to send data is limited by various factors. These factors depend on the physical conditions of the environment such as distance, channel quality, available power, and line of sight. Packet transmission consists of different frames as shown in Figure 1. To send these packets, a sensor needs to select slots for their multiple copies of packets called replicas. Therefore, a sensor needs to decide intelligently for the number of replicas and placement of these replicas in slots of each frame. In our model, sensors are directed intelligently to choose the number of replicas. Available battery and channel are major parameters that influence this decision. It is worth mentioning here that these decisions are collectively made for each user at HAP using a machine learning algorithm called Q-Learning. Details of this algorithm will be provided in the next section.

After providing an overview of the system model, we explain each parameter that is involved in our system. First, we consider a nonlinear energy harvesting (NLEH) model for our system as it is closer to real scenarios. Before defining our non-linear model for energy harvesting, we need to define the following parameters. We define distance dn as the distance between HAP and nth user equipment. During the operation of Wireless Energy Transmission (WET), the power of HAP is denoted by P. Each user's equipment has a battery with limited capacity and energy harvesting capabilities. We define μ as the amount of energy required by UE to transmit one data packet. Hence, we can assume that if a user has a battery level available of $\mu = 2$ then this user can send a maximum of two packets in one frame. It is also safe to assume that if a user does not have enough energy, then it has to wait for the WET phase and harvest enough energy to be able to again send data packets. WET capabilities of the system depend on multiple factors such as distance, channel quality, and power of HAP. Therefore, in our system channel quality is as important as battery level. We model our channel as a Rician channel. We assume an IoT scenario in which HAP enjoys a LOS with UE. Therefore, we consider the Rician channel which provides us with a LOS component along with NLOS components (multi-path). Considering this, the Rician channel hn between nth UE and HAP with K-factor κ, can be modeled as:

$$h_n = \sqrt{\frac{\kappa}{\kappa + 1}} \bar{h}_n + \tilde{h}_n \in \mathbb{C}, \tag{1}$$

where $\bar{h}_n$ is the LOS component, which represents the large-scale fading and $\tilde{h}_n \in \mathbb{C}(0,1)$ represents NLOS small scale fading due to multi-path propagation. After modeling the channel and given the distance between HAP and UE, we evaluate the nth gain as:

$$\gamma_n = \left| \sqrt{\beta (d_n)^{-2}} w_n^H h_n \right|^2, \tag{2}$$

where w_n^H is the beam-forming vector. Whereas β represents channel power gain at a reference distance of 1m which we define as follows:

$$\beta = \frac{c^2}{16\pi^2 f_c}, \tag{3}$$

here the carrier frequency is denoted by f_c. In our system, we consider Ultra-Wide Band (UWB) frequency for harvesting which ranges from 5.98 GHz to 8.80 GHz and c represents the speed of light.

The above equations show that our system model considers the channel, power gain, and energy transmission limitations of HAP more realistically. Hence, next, we define our NLEH model as:

$$E_n = \frac{\alpha_o P \delta \gamma_n}{\alpha_1 P \gamma_n + \ \alpha_1}, \tag{4}$$

where $\alpha_o = 0.399$, $\alpha_o = 0.826$ are positive constant obtained from Lacovelli, G. 2023. Considering that each slot is dedicated to either WET or WIT, we can say that each WET is exactly the time dedicated to one slot. We define this time with δ. It is worth noting with the above definitions that UE will need several slots to charge its battery. Hence it is more evident that intelligent use of battery is necessary in such scenarios.

3 JOINT ACTION LEARNER (JAL) WITH NON-LINEAR ENERGY HARVESTING (NLEH)

Joint Action Learner (JAL) was first introduced in ref. It takes advantage of the learning capabilities of a deep reinforcement learning-based algorithm called Q-Learning. In this regard, we take another step forward and make the decision process of Q-learning more efficient. We introduce the channel quality parameter along with an enhanced battery charging model. Starting with the state of the algorithm, we define it as $S_t = \{s_{t1},\ s_{t2}, \ldots,\ s_{tN}\}$. The state is the combination of available power and channel quality of each user such that $s_{tn} = (p_{tn},\ c_{tn})$. Where pt n is the power available at the sensor and ct n is the channel quality of the sensor. The algorithm decides the number of replicas for each user and decides the specific actions for each defined as $A_t = \{a_{t1}, \ldots,\ a_{tN}\}$. With each user, we also relate a reward matrix defined as R_t. The reward is based on the action taken by each user. Considering these basic parameters of JAL, we provide the following points to better understand the workings of JAL.

- Each UE is equipped with a battery capable of storing 3μ units of battery, where μ = 0.5Wh.
- Users need a minimum of a μ amount of battery to send one packet. Hence, a user can send a maximum of three packets with a fully charged battery.
- With each energy harvesting slot, UE will collect power from HAP and update its battery accordingly.
- We assume perfect CSI at HAP. Hence, it is safe to assume HAP can send a perfect beam of energy towards the users.
- User to HAP distance and channel condition limits harvesting capability of each user. A user experiences low battery recharge during WET when it does not experience LOS with HAP and vice versa.
- In our simulation, we consider each user to be at a varying distance from HAP. Distance ranges from 0.8 meters to 2 meters. These values are obtained from experimental results in ref. We also assume that the distance between UE and HAP remains constant for the whole duration frame.
- Time for each slot is δ = 0.05Sec. While the parameters of JAL have values ρ = 0.5, ξ= 0.7.
- HAP also keeps track of battery level by calculating the harvested energy using Eq. 4.
- Each user also encodes its battery information along with the data. This helps the HAP in syncing the battery information.

By considering all the above information, we now define our algorithm in Algo.1. We reduce the computational cost of each user by keeping track of channel and battery information at HAP. Hence HAP is responsible for scheduling the packets in slots based on the available battery unit and channel quality. Q-Learning updates and optimizes the Q-values using Bellman's equation defined as,

$$Q(s_t, a_t) = (1 - \alpha)Q(s_t, a_t) + \ \alpha(r(s_t, a_t) + \ \gamma \ \boldsymbol{max} Q(s_{t+1}, a_t)), \tag{5}$$

where $r(s_t, a_t)$ is the reward collected after taking a specific action. After learning through JAL, HAP directs the sensors for the next transmission. A summary of our JAL algorithm is provided below in the Algo. 1.

Algorithm 1 JAL with Non Linear Energy Harvesting (JAL-NLEH)

Initialize Parameters ρ, ξ
Initialize $Q(S, A)$ randomly
for `For each frame` t **do**
  HAP collects s_i^t
  Obtain the joint state $S_t = \{s_1^t, s_2^t, ...s_M^t\}$
  Generate random number x
  if $x < \epsilon$ **then**
    Select a joint action randomly
  else
    Select a joint action by solving
    $A_t(S_t) = argmax_{A \in \mathcal{A}} Q(S_t, A_t)$
  end if
end for
for device $i \in M$ **do**
  Randomly select data slots for action $a_i^t \in A_t$
  Harvest energy using Eq. 4
  Collect the reward r_i^t
  Observe the next state S_i^{t+1}
end for
Calculate the joint reward: $R_t = \sum_{i=1}^{M} r_i^t$
Obtain the next joint state: $S_{t+1} = \{s_1^{t+1}, s_2^{t+1}, ..., s_M^{t+1}\}$
Find $max_{A \in \mathcal{A}} Q(S_{t+1}, A)$
Update the Q-value using Bellman's equation

Figure 2. Algorithm 1.

4 RESULTS AND DISCUSSION

In this section, we first present a comparison of idle slots for IRSA and JAL-NLEH. While we prove that JAL-NLEH performs better in terms of creating more slots for energy harvesting, we focus our remaining results on the performance of JAL-NLEH in different scenarios. We see how varying the channel quality and available battery can affect the number of replicas.

Before describing the results, we present the following scenario. First, the HAP initializes the Q-Table with random values from 0 to 6, then for each frame, collects the state for each user. The battery state is initialized at 0.05Wh and the quality of the channel is measured using envelop. The HAP decodes the packets according to Successive Interference Cancellation (SIC), and then each user receives a reward that is: 0 if none of its packets successfully decoded at HAP, 1 if any packet decoded HAP but collision occurs with another user, 2 if they were successfully decoded. This reward system ensures that HAP instructs users to send more replicas when channel quality is low. HAP looks for idle slots in each frame and uses them for WET. After performing WET, it obtains new channel quality and remaining power for each user and updates its joint state.

Figure 3 shows a comparison of idle slots for IRSA and JALNLEH. Here the number of users is 8 while the number of slots in a frame is also 8 while the Rician channel K-factor is $\kappa = 8$. Hence, we can assume that competition between users for slots is high. As we can observe JAL-NLEH manages to create more idle slots which are crucial for energy harvesting as compared to IRSA. This is because IRSA does not have the advantage of battery and channel information while JAL-NLEH has many parameters such as channel, and battery information of users to work with. A comprehensive analysis of idle slots, decoded packets, and collided packets is shown in Figure 4. Here, we also consider the extreme scenario where competition for slots is very high. Our algorithm not only aims at reducing the number of collisions but also manages to create more idle slots for energy harvesting. It is worth mentioning that the probability of having more than 6 collisions is almost negligible while at the same time probability of idle slots has been also on the higher side.

In Figure 5, We keep our number of slots and users constant while varying the channel quality κ. We can observe two trends here, first is before idle slots are 5, and second, after the idle slots are more than 5. We explain this trend by explaining the battery reservation goal of JAL-NLEH. We can see an increasing trend in the probability of number of idle slots and after 6 number of idle slots, we see a decreasing trend. This shows that it is relatively easy for JAL-NLEH to create 6 number of slots as compared to a higher (more than 6) number of slots. We can also see how the channel affects the probability of idle slots. For example, for the worst channel i.e. κ = 4, probability is lesser than for $\kappa = 6$ *and* 8. Another worth noting analysis from these results is that performance for κ = 6 and 8 is not much different than each other. It is because our algorithm is optimized to create more idle slots and it enhances its performance when the conditions get tough. Another observation one can make is that our algorithm is optimized for channels with moderate values $\kappa = 6$ *and* 8 which are the conditions for outdoor channels.

Finally, in Figure 6, we present a collective result for different settings and $\kappa = 6$. It is observable from the figure that our algorithm always performs much better in conditions that matter the most.

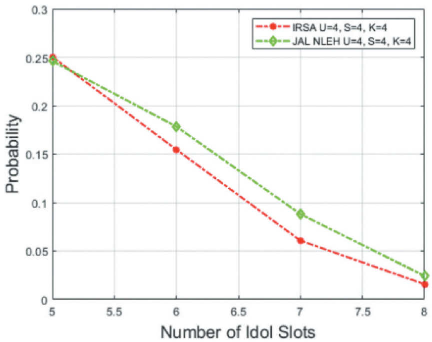

Figure 3. Probability of idol slots with $\kappa = 8$ for IRSA and JAL NLEH.

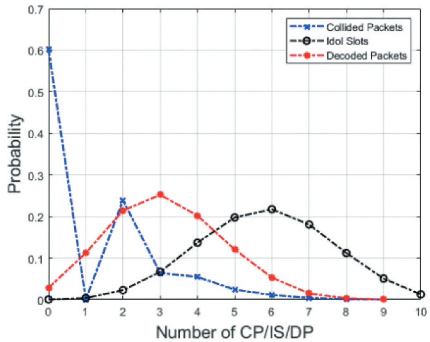

Figure 4. Comparison of Decoded packets, idol slots, and collided packets for $\kappa = 8$ users=10, and slots=10.

5 CONCLUSION AND FUTURE WORK

We have shown the key to achieving better performance in tougher scenarios for random slot selection of bridge sensor networks. We have also shown the importance of information related to physical layer characteristics at the mac level layer. This can help in improving the decision capability of the Mac layer. In this way, we can achieve more efficient with minimum age of information data for bridges which may be hit by disaster or flooding. We can also conclude that traditional IRSA needs improvements and these improvements can be made by enhancing the decision-making ability empowered by machine learning. For future work, we would like to extend this work with a continuous optimization/machine-learning algorithm such as proximity policy optimization (PPO) which does not require quantizing the battery and channel levels hence improving the results.

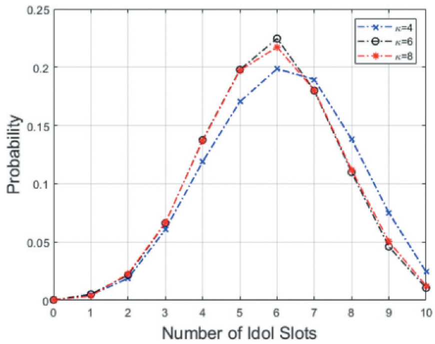

Figure 5. Comparison Of Idol Slots For Different κ, U=10, S=10.

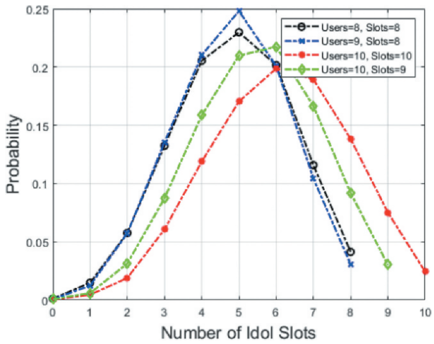

Figure 6. Comparison of different users and slots settings for κ.

REFERENCES

Mushtaq. M. T. & Cordeschi, N. & Grieco, L. & Boggia, G. 2023. Optimized Scheduling of Packets with Deep Sensing Irregular Slotted Aloha Influenced by Channel Knowledge for Wireless-Powered IoT Networks. International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), September 2023, Cosenza, Italy.

Lacovelli, G. & Grieco, G. & Petrosino, A. & Grieco, A. & Boggia, B 2023. Fair Energy and Data Rate Maximization in UAV-Powered IoT-Satellite Integrated Networks. IEEE Transactions on Communications, 15 December 2023.

Li, Y. & Chin, K. 2022. Energy-Aware Irregular Slotted Aloha Methods for Wireless-Powered IoT Networks. IEEE INTERNET OF THINGS JOURNAL, 11784–11795 VOL. 9, 14, JULY 2022.

Choi, H. & Shin, W. & Levorato, M. & Vincent, P. H 2019. Harvest-Or-Access: Slotted ALOHA for Wireless Powered Communication Networks. IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, 11394–11398 VOL. 68, NO. 11, NOVEMBER 2019.

Hu, Y. & Seo, J. & Jin, H. 2022. Exploiting In-Slot Micro-Synchronism for S- ALOHA. IEEE TRAN. WIRELESS COMM., 10854–10870 VOL. 21, NO. 12, Dec. 2022.

Hmedoush, I. & Adjih, C. & Mu¨hlethaler, P. 2022. Deep Learning, Sensing-based IRSA (DS-IRSA): Learning a Sensing Protocol with Deep Reinforcement Learning. 23 Sep 2022, https://hal.inria.fr/hal-03744126v2.

Schmickl, S. & Faseth, T. & Pretl, H. 2019. A 1.9-μW 7-GHz IR-UWB Transmitter with RF-Energy-Harvester in 180-nm CMOS for Battery- Less Bio-Sensors. 17th IEEE International New Circuits and Systems Conference (NEWCAS), June 2019, Munich, Germany.

Chen, Z. & Feng, Y. & Feng, C. & Liang, L. & Jia, Y. & Quek, T. Q. S 2023. Analytic Distribution Design for Irregular Repetition Slotted ALOHA With Multi-Packet Reception. IEEE TRANSACTIONS ON VEHICU- LAR TECHNOLOGY, 1360–1365 VOL. 72, NO. 1, JANUARY 2023.

Chu, Z. & Xiao, P. & Mi, D. & Hao, W. & Lin, Z. & Chen, Q. & Tafazolli R. 2022. Wireless Powered Intelligent Radio Environment with Non-Linear Energy Harvesting. IEEE Internet of Things Journal, 18130–18141 VOL. 09, NO. 18, March 2022.

Chen, Y. & Zhao, N. & Alouini, M. 2017. Wireless Energy Harvesting Using Signals from Multiple Fading Channels. IEEE Transactions on Commu- nications, 5027–5039 VOL. 65, NO. 11, August 2017.

Hu, Y. & Seo, J. & Jin, H. 2022. Exploiting In-Slot Micro-Synchronism for S-ALOHA. IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, 10854–10870. VOL21, NO. 12, DECEMBER 2022.

Ayoub, I. & Hmedoush, I. & Adjih, C. & Khawam, K. & Lahoud, S. 2021. Deep-IRSA: A Deep Reinforcement Learning Approach to Irregular Repetition Slotted ALOHA. 10th IFIP International Conference on Performance Evaluation and Modeling in Wireless and Wired Networks (PEMWN), Nov 2021, Ottawa, ON, Canada.

Dumas, C. & Salau¨n, L. & Hmedoush, I. & Adjih, C. & Chen, C. S. 2021. Design of Coded Slotted ALOHA With Interference Cancellation Errors. IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, 12742–12757 VOL. 70, NO. 12, DECEMBER 2021.

Schulman, J. & Wolski, F. & Dhariwal, P. & Radford, A. & Klimov, O. 2017. Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347 (2017).

Bi, S. & Zeng, Y. & Zhang, R. 2016. Wireless Powered Communication Networks: An Overview. IEEE Wireless Communications. 10–18 VOL 23, April 2016.