CISTER
Research Centre in
Real-Time & Embedded
Computing Systems

Book Chapter

Human Sensing Network Architecture and Challenges in Smart Cities

Tanmay Chaturvedi
Kai Li*

*CISTER Research Centre

CISTER-TR-180111

2017/09/13

Human Sensing Network Architecture and Challenges in Smart Cities

Tanmay Chaturvedi, Kai Li*

*CISTER Research Centre

Polytechnic Institute of Porto (ISEP-IPP)

Rua Dr. António Bernardino de Almeida, 431

4200-072 Porto

Portugal

Tel.: +351.22.8340509, Fax: +351.22.8321159

E-mail: tanmay_chaturvedi@ntu.edu.sg, kaili@isep.ipp.pt

http://www.cister.isep.ipp.pt

Abstract

Crowdsensing, people tracking and urban space utilization are critical tasks for understanding human behavior in smart cities. Human sensing applications provide valuable feedback to the urban developers and architects to optimize public facilities and re-invent spaces. In this chapter, a Renewable Wireless Sensor Network (RWSN) architecture for human sensing is presented to study the spatial and temporal information of urban space utilization and pedestrian flow. The individual components, pertaining to human sensing and RWSN network architecture, are also explained.

Chapter

Human Sensing Network Architecture and Challenges in Smart Cities

***Tanmay Chaturvedi**, *Kai Li*+**

*Transport Research Centre, Nanyang Technological University, Singapore

+CISTER Research Centre, Porto, Portugal

Abstract

Crowd sensing, people tracking and urban space utilization are among the crucial tasks for understanding people behaviors in smart cities. Human sensing applications provide valuable feedback to the urban developers and architects to optimize public facilities and re-invent spaces. In this chapter, a Renewable Wireless Sensor Network (RWSN) architecture for human sensing is presented to study spatial and temporal information of urban space utilization and pedestrian flow. The battery-powered sensor nodes harvest ambient energy and sensory data is transmitted wirelessly to a cloud server without any human intervention or on-site investigation. All these elements translate to a part of big data and Internet of Things (IoT). Furthermore, prominent challenges pertaining to human sensing and the sensor data analysis is discussed and possible solutions are presented to overcome those challenges.

Keywords: Smart City; Human Sensing; Renewable Wireless Sensor Network; Urban Space Utilization;

1. Introduction

In recent years, the popularity of mobile and pervasive computing stimulates many research efforts on wireless human sensing. An increasingly common requirement of human sensing is to extract information regarding the people density, their space utilization pattern, and moving trajectories in an environment. Many interesting questions could arise, for example: how many

Corresponding Author Email: *tanmay_chaturvedi@ntu.edu.sg, +kaili@isep.ipp.pt

people visit a particular place, say a food joint in a shopping mall? How many workers work on a construction site? Which pathway do people prefer to choose when it is raining or sunny? How much time do people spend on a playground or a jogging track? Such sensing information helps researchers, urban planners and government entities understand the human behavior corresponding to different external parameters and helps service providers to optimize public facilities in residential communities.

In smart cities, all human-related information is collected through a network of wireless sensors that are over the city. Wireless sensors remotely monitor a space and eliminate the need for manual labor to collect any information. The recent developments in low-cost devices and advancement in Internet of Things (IoT) has enabled to deploy low-cost wireless sensor network (WSN) in urban areas. As a result, various applications based on urban WSNs have been proposed, such as air quality monitoring [1], structural health monitoring [2], traffic monitoring [3], disaster monitoring [4], etc. The human sensing networks help to deeply understand pedestrian flow, human mobility, and space utilization. A more accurate sensing methodology would require more complicated sensing instruments which would inevitably give rise to certain challenges. For example, camera-based sensing system accurately differentiates humans from other background objects, which furthers helps in the computation process for any smart city-based analysis. However, the camera-based systems have limited coverage to detect people due to a fixed location and angle, which are not suitable for large-scale human tracking. Moreover, camera-based systems require transmission of the huge data payload to a remote server, which increases the power consumption. Different from the above devices, low-cost electronic sensors consume less power and are less computation-intensive. They may run out of battery, which can be recharged by harvesting ambient energy, e.g., solar, the wind, vibration, etc.

There have been many developments in the renewability of battery-powered devices. As the technological trend moved from wired to wireless, the demand for energy-efficient devices peaked up. Battery-powered devices are capable of collecting information over a region which could even be non-accessible to humans. But, with a small form factor and renewability, the deployment of wireless devices becomes easier and cost-effective. Though the devices are battery-powered, low-power microcontrollers and high optimization in electronic circuitry have enabled technicians to drastically reduce the power consumption and increase battery life. Further, to renew the limited battery reserve, solar energy has been widely utilized as a major energy source.

In this chapter, a Renewable Wireless Sensor Network (RWSN) is investigated, which is used for determining pedestrian flows, human mobility and space utilization across any region. In particular, a human sensing network architecture for a smart city is discussed, and the potential of WSN in tracking the human flow and understanding the spatial-temporal usage of certain point of interest (POI) is explored. Further, the discussion is extended by focusing on the challenges in the network architecture for smart cities.

2. RWSN NETWORK ARCHITECTURE

In this section, firstly, the RWSN architecture is presented. Next, sensor nodes, hardware components, wireless data transmission, and data analysis are discussed.

The RWSN is an extension of WSN, where the transfer of information takes place through wireless sensors powered purely by renewable energy. This enables the deployment of these sensors in remote areas where human intervention and human access are limited. Figure 1 presents the network architecture consisting of seven key components, which are (1) People Counting Sensor Node, (2) Environmental Monitoring Sensor Node, (3) Wi-Fi Sniffers, (4) Relay, (5) Gateway, (6) Cloud Server, and (7) Base Station.

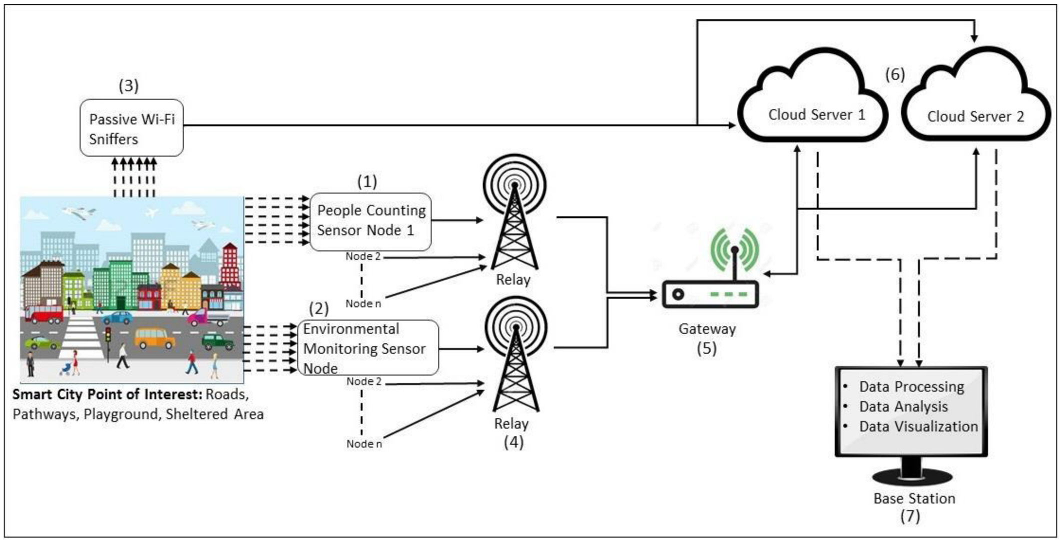

Figure 1. RWSN Network Architecture for Human Sensing.

- *People-Counting Sensor Node*: As proposed in [5], the people-counting node consists of two ultrasonic sensors placed side by side in parallel. Ultrasonic sensors are used to detect people passing by up to the range of 2 meters. By using two parallel sensors, the direction of motion can be identified. For example, if the left sensor detects the person first followed by a right sensor detecting it, the direction of motion can be classified as left, and vice-versa. Using this principle, the directivity of moving person can be known at an instance. This helps to understand the flow of people across a region. In working, the people-counting sensor will collect the number of moving people for a total period of 5 minutes, referred as the sampling period hereafter. The 5-minute accumulated data will be sent together as one data payload to the Gateway. This transmission scheduling is vital to ensure the longevity of sensor nodes, as the sensor nodes are powered by a battery and the battery is recharged by a solar panel.

- *Environmental Monitoring Sensor Node*: The environmental monitoring sensor node has the capability to collect weather information. An example of the node is shown in Figure 2, which consists of a temperature sensor, pressure sensor, barometer sensor, ultra-violet (UV) sensor, luminous intensity sensor, humidity sensor, motion sensor, rain sensor, and noise sensor. All of these sensors are integrated and interfaced with a controlling unit (microcontroller/microprocessor). In order to make the whole sensor setup self-sustainable, renewable solar energy is provided through solar panels, as shown in Figure 2. The placement of solar panel over the sensor box plays an important role in ensuring that maximum solar energy is harvested. Therefore, the solar panel is placed at the top of the sensor box facing the sky, which ensures the sun rays falls on the overall surface area of the solar panel. Likewise, in order to achieve maximum accuracy in data collection, the environmental sensors are strategically placed over the sensor box. The luminous intensity sensor and the ultra-violet sensor is placed over the top of the sensor box. Temperature and humidity sensor are placed underneath the sensor box for maximum sensing accuracy. The rain sensor is resistive, meaning the resistance of the sensor changes according to the presence of water droplet and its amount accumulated over the sensor surface. Hence, the rain sensor is placed at sideways.

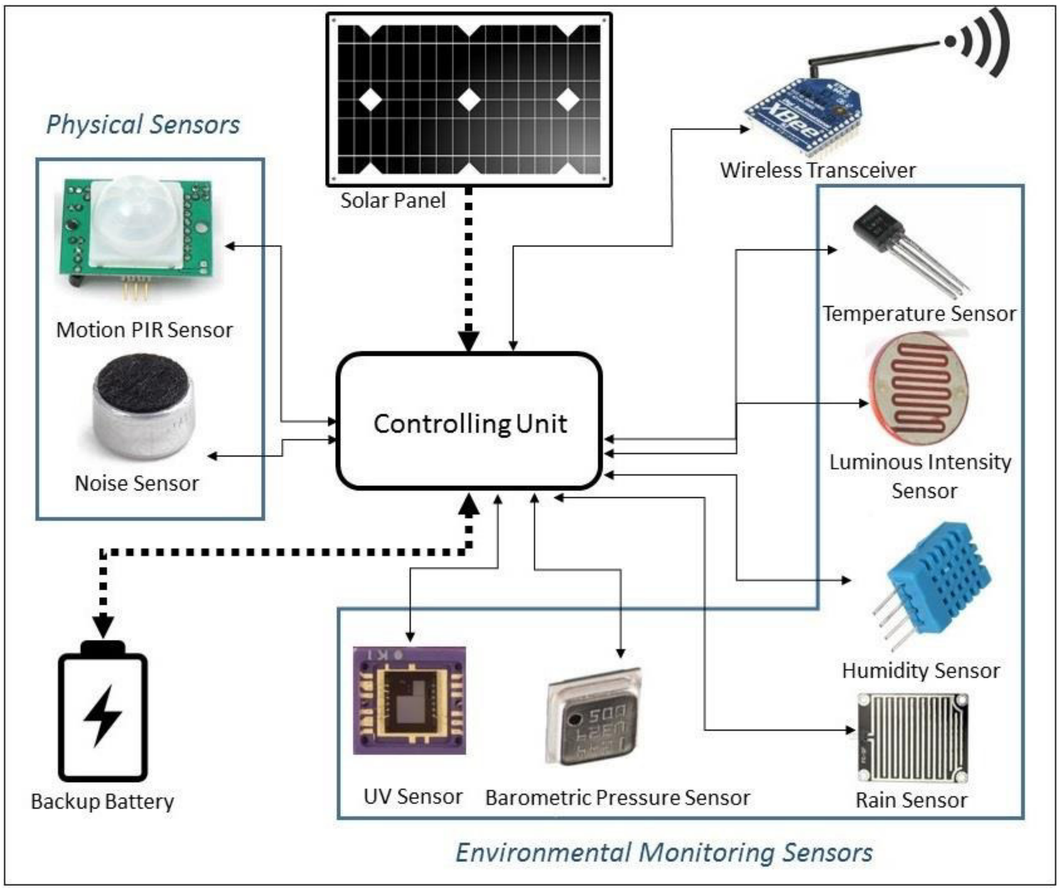

Figure 2. Renewable Environmental Monitoring Sensor Node.

In terms of sensor data acquisition, the controlling unit schedules the data collection of the motion sensor and noise sensor differently than other sensors, as shown in Figure 3. Since motion and noise values are continuously changing over time, these values can be collected throughout the time interval of, say 5 minutes, and is accumulated in one data packet each. Since the other environmental factors like temperature, luminous intensity, etc. do not change abruptly over time, only 1 instantaneous value of these sensors are collected in a 5-minute interval.

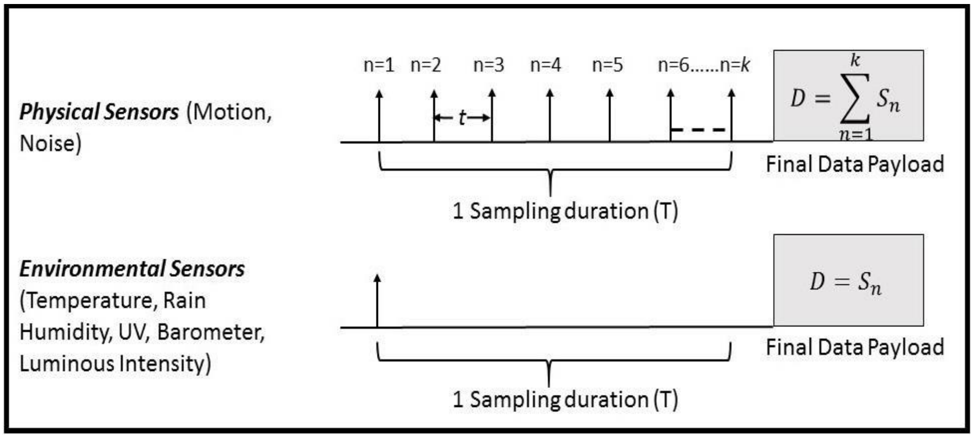

Figure 3. Sensor data acquisition methodology from different sensors.

The final data payload, as collected in case of physical sensors like motion and noise sensors is defined as:

$$D = \sum_{n=1}^{k} S_n \tag{2.1}$$

where D is the summation of total k data accumulated over one sampling duration, taken in regular intervals of t, where t is the time duration between 2 individual data recorded by physical sensors within 1 sampling duration (T). S_i is the instantaneous sensor value in the i^{th} time interval. The time difference, t between two individual data within one sampling duration, T is calculated by:

$$t = \frac{T}{k} \tag{2.2}$$

In terms of environmental sensors, the final data payload D is equal to one instantaneous sensor value collected within a sampling duration.

$$D = S_n, \{n = 1\} \tag{2.3}$$

The sampling rate for wireless data transmission remains T for all the sensors. However, this sampling rate can vary according to the residual energy in the sensor box. For example, when the residual energy is lower than a pre-set threshold value, the data transmission frequency will reduce by some pre-defined amount to ensure that the data is recorder for a prolonged period.

- *Wi-Fi Sniffers*: A Wi-Fi sniffer is used to collect Wi-Fi probe requests from smartphones in its radio coverage. The passive Wi-Fi sniffer can be built on portable single board computers (SBCs) like BeagleBone and Raspberry Pi. With these SBCs, up to 4 USB dongles can be connected together. However, for this particular sensing applications, 2 USB Wi-Fi dongles can do the job. One dongle will sniff the Wi-Fi probe requests while the other can help to relay the processed information to the base station.

Since every smartphone has a different physical address (MAC address), Wi-Fi sniffer is able to distinguish between different probe requests, and hence, the number of people at a particular time can be counted. Moreover, if multiple sniffers are deployed strategically, the flow of people can also be tracked which is helpful in generating further insights. Note that the Wi-Fi sniffers work in passive mode, and it does not collect any personal information. Though the probe packets might differ between smartphones of different manufacturers, generally, the packet contains the smartphone's Source's MAC address, Destination's MAC address and the type of frame.

- *Relay*: In wireless communication, the transfer of information to the base station relies heavily on the network link strength and network connectivity. The presence of buildings, vehicles and electromagnetic interference reduce the signal strength and coverage of the wireless network. Thus, a relay acts as a repeater which is used to increase the coverage of mesh network. The relay nodes make the wireless network highly reliable by sharing the burden of sensor nodes, which increases the lifetime of WSN [6]. Based on the application, the relay could be set to transmit the data collected from sensor nodes to the base station. However, the relay has to stay in the listening mode forever if it is not synchronized with the sensor nodes.

- *Gateway*: As the renewable wireless sensors are designed to be less power consuming, they do not generally contain Wi-Fi capabilities to directly relay data to the cloud server.The lack of Wi-Fi capabilities of low power sensor nodes ensures that they can be deployed anywhere, irrespective of the presence of an internet connection in the vicinity. Thus, the gateway acts as a link between the cloud server and the in-place wireless sensors. Gateway also helps to buffer the messages in case the cloud server is busy. In addition, if cloud server does not provide an acknowledgment to the gateway, it re-transmits the data to the cloud server, to avoid potential data loss.

- *Cloud Server*: In all the remote sensing applications, the role of the cloud server is imperative. It acts as a link between the sensor nodes and the remote base station (BS) and facilitates the access of the data sets from anywhere. As presented in Figure 1, multiple cloud server can be setup to help reduce packet congestion. When one of the servers is busy or down, the remaining backup servers can take over and do the necessary job.

- *Base Station:* The sensor data collected on the cloud servers is sent to Base Station (BS). BS possess more computational power than the wireless sensor nodes and is equipped with a more stable energy source, like the power grid. The role of BS is to process the sensed data, perform temporal and spatial analysis, and provide insights through interactive graphical visualization. At the BS, sophisticated software programs and APIs are running to perform specific tasks on the collected data.

  Multiple numerical and statistical computational software like MATLAB and R provides functionality for data processing and visualization. Data visualization provides a more vivid analysis of the computations and helps to derive insights on the collected big-data. Figure 4 presents a broad overview of the type of computations performed at the BS.

3. HUMAN SENSING WITH RWSN

In smart cities, human sensing involves identifying the presence of people, understanding their day-to-day activities, tracking their flow of movements, and determining how different external conditions affect their behaviors. In the following sections, three different human sensing models are discussed. People counting model utilizes multiple ultrasonic sensors to count the people passing by across a road or pathway in the POI area. Human space utilization model utilizes sensor-fusion of motion and noise data to determine how much an area is utilized over the time. Weather classification model utilizes multiple environmental sensors to collect weather data, determine different weather conditions based on that data and presents insights on the changes in human behavior based on changes in weather conditions across a POI.

3.1. PEOPLE COUNTING MODEL

In people counting model, people mobility and crowd number monitoring are based on the notion of *Right to Left* or *Left to Right* and the notion of wireless devices around respectively. The notion of people counting is denoted as following:

$$D_{(t)} = \{RL_{(t)}, LR_{(t)}\} \tag{3.1}$$

where $D_{(t)}$ is denoted as total moving directivity recorded at time, t. $RL_{(t)}$ is denoted as the number of motion recorded moving from right to left, while $LR_{(t)}$ is denoted as the number of motion recorded moving from left to right.

The individual parameters used in the people-counting model are calculated through the data recorded from ultrasonic sensors. The people-counting sensor nodes can be ideally placed near the pathways, lift lobbies, sheltered areas or any particular POI. Moreover, to increase the accuracy of data collection and precisely identifying the pedestrian flow, the people-counting sensor nodes can be placed along the Passive-Wi-Fi sniffers. There would be, however, a slight discrepancy between the data collected from Wi-Fi sniffers and people counting nodes which will be further discussed in the following section.

3.2.HUMAN SPACE UTILIZATION MODEL

The usage of an area at a particular time t can be defined as Space Utilization. When people utilize a space, they generate some movement and noise. Human movement and noise can be actively sensed from Environmental Monitoring Sensor Node, which then can be utilized to define the space utilization at a POI.

With the advancement in Sensor-Fusion techniques, it is now more reliable to model the relation between multiple individual parameters and extract meaningful information out of them. Hence, to define space utilization across a region, the combined effect of noise and motion will have to be considered. Space utilization can be thus, defined as:

$$S_t = f\{(M_t), (N_t)\} \quad (3.2)$$

where S_t is the space utilization function at time t, (M_t) is the individual contribution of Motion and (N_t) is the individual contribution of Noise in the space utilization calculation. Merging Motion and Noise values provides a more accurate definition of space utilization, but does not guarantee a definite interpretation of space utilization. Therefore, in order to investigate the combination of both noise and motion data, investigation of the individual contributions of values obtained from both sensors is needed.

Since motion and noise values have different measurement units, they cannot be directly merged to extract space utilization information. Therefore, each value has to undergo normalization function to generate values in the range {0.0, …., 1.0}. The normalization function $Norm$ for any value X is defined as:

$$Norm(X) = \min\left(\frac{X}{NormValue}, 1\right) \quad (3.3)$$

where $NormValue$ is the normalization threshold value set differently for each individual sensor. The normalization threshold will be application specific and based on the type of sensor data. For example, the motion values could lie between zero to a small finite number in a sampling period, but noise values might not be zero anytime because of random electronics errors and could reach very high during thunderstorm and etc. Normalization threshold is then determined accordingly to normalize the existing values and remove the outliers.

Now, combining the equations (3.2) and (3.3) gives a more accurate representation of space utilization across a POI at a particular time t:

$$S_t = f\{Norm(M_t)\,, Norm(N_t)\} \quad (3.4)$$

3.3. Human Behavior with Weather Conditions

Weather plays an important role in influencing human behavior. Changing weather conditions affect the pedestrian flow and space utilization. For example, people prefer to walk in open areas when the weather is cloudy or non-raining and walk under sheltered roofs or pathways when it is hot or raining. It is thus important to model a relation between urban space utilization and weather, as the relation will give further insights about the way humans behave in a particular weather.

Different classification of weather conditions is possible, such as Raining, Non-raining, Hot, Cold, Humid, Non-Humid, Hot and Humid, Dry, Cloudy etc [7]. The classification and identification of such weather conditions are possible with environmental monitoring sensors like humidity sensor, temperature sensor, barometric pressure sensor, luminous intensity sensor, rain sensor and UV sensor.

The fusion of space-utilization model and people counting model with the weather classification model helps to understand the human mobility and urban space utilization in different weather conditions. Specifically, the weather classification model filters out datasets based on the weather conditions. Figure 4 presents a detailed architecture of pre and post-processing of sensor data at the BS to understand human behavior based on different conditions.

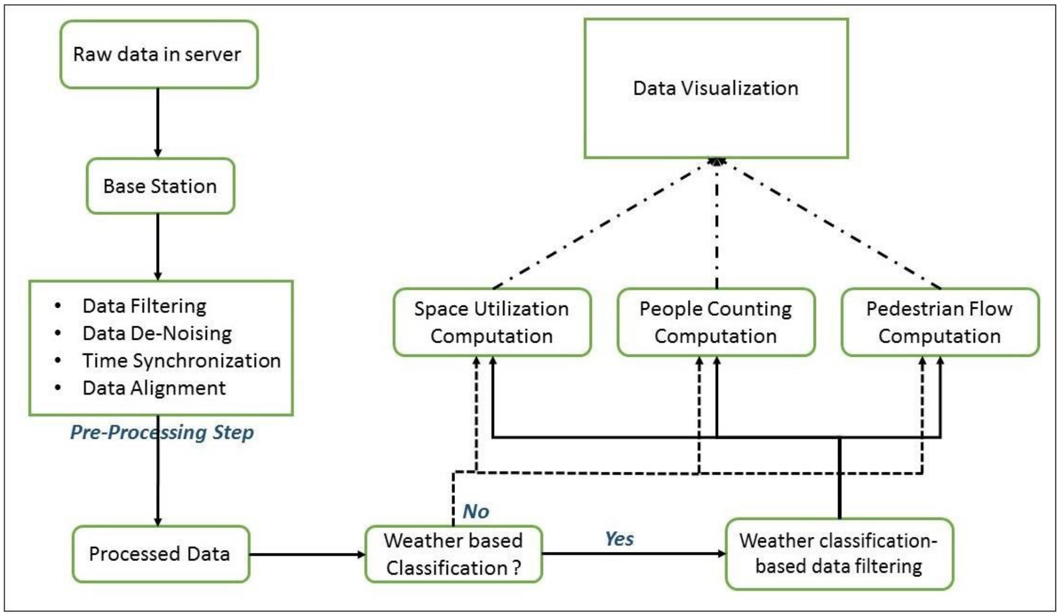

Figure 4. Overall human sensing data analysis model.

Raw data is stored in the cloud server. Base station extracts this data for offline processing and analysis. Data filtering and de-noising remove any outliers or random errors in the data set. Time synchronization step synchronizes all the data sets collected from different sensor nodes according to local timings. Weather-based classification filters out sensor data corresponding to different weather conditions. Next, individual human sensing blocks carry the computations, the output of which is the graphical visualization of analyzed data.

4. Prominent Challenges

In this section, prominent challenges in the human sensing network architecture are discussed. Specifically, challenges in human sensing arise during the computation of human mobility, urban space utilization, data collection, data processing and data analysis.

4.1. Challenges in Human Sensing

- *Space Utilization*: As discussed before, urban space utilization is computed by combining the individual effects of motion values and noise values. However, there are still certain challenges in this approach. Firstly, there are no particulars sensors available, which can determine the space utilization across a region. The reason being, the definition of space utilization may vary as per the research approach. Unavailability of such sensors restricts the evaluation of the algorithm presented in equation 3.4 against the ground truth. The calculation of space utilization, as discussed before, is by far the closest approach in determining the ground truth.

Secondly, as the space utilization is modeled based on motion and noise values, there is a certain level of a discrepancy possible in this approach. Consider the instance when noise value might significantly shoot up, which could be in the case of a thunderstorm or lightning. This would affect the calculation of space utilization drastically, as high noise and a corresponding increase in motion could result in high space utilization, which might not be the actual case. Similarly, a high peak in motion values along with considerably fewer noise values could still result in a high value of space utilization. Since the actual microphone-based sound is not collected, but only the noise intensity, the calculation of space utilization might deviate from the ground reality in some scenarios. The best solution to this challenge of eliminating the extreme sensor values is by using Normalization and then filtering out these high values through a static or dynamic threshold.

Finally, the space utilization is location specific. For example, children while playing in the playground generate sound and motion, but people taking nap or rest at a particular region generate slight noise but no or negligible motion. Comparing the two scenarios and assuming an equal number of people in both the situation, the sensor placed near the playground will detect high space utilization than the sensor placed at the location where people take rest. This discrepancy can only be resolved when the ground truth is known, which is not always the case. Thus, these challenges affect the accuracy in calculation of space utilization.

- *Accuracy and sniffing challenges in Passive Wi-Fi*: Human sensing requires accurate data collection to understand human behaviors. The Wi-Fi sniffing works on the principle of probe request collection from different smartphones. On any smartphone, probe requests are transmitted in bursts, the interval of which depends on the Operating System (OS), Wi-Fi chipset driver, Wi-Fi connection, and on the smartphone status, like Wi-Fi registered and screen off, Wi-Fi registered and screen on, Wi-Fi not registered and screen on, Wi-Fi not registered and screen off.

  The majority of smartphones currently available in the market runs on three major OS, namely iOS, Android, and Windows. Each of these OS has different probe request interval. This difference is largely employed by the OS developers as a measure to reduce the power consumption under different circumstances. The study of probe intervals in different OS and different smartphone status, as presented in Table 1, results in the following data:

Smartphone OS	Wi-Fi Registered		Wi-Fi Not Registered	
	Screen On	Screen Off	Screen On	Screen Off
Android	2.11 sec	2.15 sec	0.8 sec	1 sec
iOS	1200.8 sec	1204.4 sec	70.6 sec	109.8 sec
Windows	1200.8 sec	1204.4 sec	10.9 sec	13.9 sec

Table 1: Average probe request interval in different smartphone OS [8].

As evident from the Table 1, in screen on and Wi-Fi Not Registered status, the smartphones increase the probe request interval to conserve battery power, comparing to the status of the screen on. Moreover, iOS and Windows OS have a long interval of around 1200 seconds when the user has connected to Wi-Fi network. However, Android phone still keeps a short interval of 2.11 and 2.15 seconds. The different probe request intervals are due to a differentiated energy-saving design of smartphones in Wi-Fi registered status.

It is also evident from Table 1 that a longer contact time between the sniffer nodes and the smartphone can increase the probability of receiving probe request. Means, the slower the people moves, the higher the chance that the nodes can capture the probe requests from the smartphones. However, the movement style of people is random. They could be running, walking, jogging, or standing. Therefore, in order to increase the probability of probe request collection, the number of Wi-Fi sniffing nodes has to be increased. This approach ensures that a probe request emerging from a smartphone is captured by at least one of the sniffers, irrespective of the movement style of the individual. Moreover, the same probe request logged in different sniffing nodes can be easily eliminated later in the data processing part.

- *People counting accuracy*: People counting nodes utilizes ultrasonic sensors to determine the motion and directions of people. However, two or more people walking in parallel might be detected as only one human from the sensor. This is a limitation of using ultrasonic sensors-based people counter. Moreover, people counting through Wi-Fi sniffers also has some challenges. Firstly, the assumption that every passer-by will carry a smartphone is not always valid. Secondly, as discussed in Table 1, the probe request intervals are different for different smartphones. So, people going at a fast pace might remain undetected due to a longer probe request interval. One way to overcome the discussed challenges and to increase the accuracy of people counting model is to use Wi-Fi sniffers near to people counting nodes to ensure that a maximum number of people are counted at a particular instance.

4.2. Challenges in Data Analysis

- *Data acquisition and transmission frequency*: Sampling time in the acquisition of data plays a very important role in battery-powered sensor applications. In the network architecture presented above, since different sensors have different ways to sense the information, it becomes very imperative to set a sampling rate, which is suitable for all the sensors. The human sensing in a smart city requires non-sporadic and reliable sensor data. If the data goes missing due to any circumstance, the reliability and the accuracy of the network architecture is compromised. This can be broadly understood through some examples. As presented in Figure 3, the instantaneous noise and motion data is accumulated over the period of 5 minutes and then packaged into a single data byte for each of the sensors. This is done because the motion and noise values are sporadic and could randomly increase or decrease according to the indefinite human movement. On the contrary, the values of temperature, humidity, atmospheric pressure etc. do not change instantaneously but gradually. Hence, even a single instantaneous sensor value collected once over a sampling duration could effectively cover the entire information for the sampling period. As also explained in Figure 3, this method helps to reduce power consumption by scheduling the data acquisition from different sensors and helps to extract more meaningful information from limited data set.

- *Packet scheduling and collision avoidance*: Three major challenges arise in the data collection for WSN and IoT networks: massive number of sensor nodes accessing a shared challenge in the service area of the base station (BS), scheduling the data transmission of sensor nodes, and to guarantee a fair channel access to all the nodes [9]. As number of sensor nodes are added to the same wireless network, the contention to access the shared wireless medium increases. This increase in contention results in higher possibilities of packet collision, which further results in data and information loss. To avoid collision between packets, a Medium Access Control (MAC) protocol is necessary which determines how and when nodes access the shared channel. The protocol, as discussed in [9] takes into consideration parameters like residual energy inside the sensor node, the dynamic contention probability and sensor priority to determine the sequence of data transmission from different nodes. This not only saves a lot of battery power of the nodes as they only transmit when they are allowed to, but it also helps the server or BS avoid congestion due to its limited capabilities and also helps to avoid unwanted packet collision and information loss. The categorization and performance of different MAC protocols for WSNs are presented in [10], the discussion of which is beyond the scope of this chapter.

- *Server congestion*: During the instances of server congestion, the backup server is utilized. The gateway receives an acknowledgment from the server every time server receives the data. If a server fails to send an acknowledgment, gateway re-send the data again to the server. In the case of multiple failed attempts to send data to the server, the gateway re-routes the data to the next available cloud server. This helps to avoid data loss. Overall, to optimize the resources, gateway schedules the transmission of data dynamically as per the state of the server.

- *Data processing*: In order to extract meaningful information from the sensors' data and make the human sensing network architecture more reliable, a robust data processing is required. It might look obvious that there would not be any glitches in the data collection once the wireless sensor network is successfully established. There are, however, possibilities, where the data collected might not be: accurate, time synchronized, noise-free, double-sampled, etc.

  There are multiple reasons like malfunctioning of hardware, environmental noise, electrical noise, data packet loss which contribute to the random errors in the collected data. Data filtering is required to eliminate the outliers in the data. For example, the high peaks in the noise values could be the resultant of lightning sounds, honking of cars, the sound of a nearby metro train, etc., which doesn't represent actual noise made by humans. So, a *NoiseThreshold* value could be used to separate the outliers from the meaningful data.

  Data redundancy could occur due to the missing acknowledgment from BS, which results in sensor resending the data packet. It is fairly easy to identify and eliminate the double packets, as they have same sensor value with a different time stamp and are reported to the BS twice in one sampling duration.

Time synchronization between different sensor nodes is important while correlating different sensor values at a particular instance to extract information. Synchronization occurs only when the BS has the ability to re-program the deployed sensor nodes and adjust their internal clock time. This, however, is difficult due to fewer computation abilities of the sensor nodes to conserve their power at maximum. Thus, post-processing algorithms are imperative in this regard. A *TimeSync* module can be incorporated to enable the data analysis framework to process the data from multiple sensors at the same time. A threshold is defined to allow the maximum drift between the timestamps of sensor data and local clock of the server. Say, for example, the threshold is T_{th}, local clock is $T_{localclk}$ and the time-stamp is T_{data}. The data samples, which have $| T_{data} - T_{localclk} | \leq T_{th}$ are all synchronized with the time $T_{localclk}$.

In the final collected data, there could be some instances when random data packets are missing. This generally happens when one or many sensor nodes did not transmit their data packets because of less or no residual energy inside them. Concurrently, it is also possible that few of the other sensor nodes did transfer their packets. This creates an imbalance between the number of packets collected from different deployed sensor nodes. *TimeSync* module plays an important role in removing this discrepancy in data processing, as it helps to eliminate the data packets in the instances when the data was collected from only fewer nodes and not from the rest of nodes.

- *Hardware challenges*: All electronic sensors are interfaced with a controlling unit (microcontroller/microprocessor) which is responsible for scheduling the data acquisition from sensors, transmitting the data wirelessly through wireless transceiver and recharge its battery through the solar panel. Once the sampling period is set up, the controlling unit will direct its sensors to collect the data instantaneously or continuously in the sampling period, as per the respective sensor type. As earlier presented in Figure 3, physical sensors like noise and motion sensors continuously acquire the data throughout the sampling period, while environmental sensors collect the data for only one instance in a sampling period. This approach helps to minimize the power consumption of the overall sensing node. However, there are other challenges which hinder the effective functioning of the sensing node.

  As the whole sensing unit is renewable, it relies heavily on solar energy for its working. The job of the controlling unit is to basically power the whole hardware unit using solar energy and in the parallel, store the extra energy in the backup battery pack. However, the capacity of the backup battery pack is also limited to few thousand mili-ampere hours (mAH). In the event of less or no sunlight in cloudy weather, the backup battery provides the required power. But, when the sunlight is less for a prolonged time, say for few days, the sensing unit might run out of power once the backup battery power is fully utilized and thus get switched off. This results in loss of data for that particular period. In order to prolong the data collection, an adaptive residual energy-based threshold can be setup, which continuously compares itself with the residual energy in the backup battery. Based on the amount of residual power in the backup battery, the sampling frequency can be dynamically adjusted. For example, if the residual power is only 25% of its maximum capacity, the data transmission frequency could be reduced by some amount, thus ensuring that the data is collected for a prolonged time even with less frequent data packets. Another direct solution to this challenge is to optimize the power consumption of the controlling unit by employing less computation-intensive algorithms.

The controlling unit is also prone to humidity and moisture effects in the course of weather change, the protection of which is ensured by properly sealing the entire sensing unit during its manufacturing and thus reducing the chance of water and moisture reaching near the controlling unit.

5. CONCLUSION

In this chapter, the potential of RWSN for extracting human sensing information is explored. Human sensing is investigated using three different computation models, which are, space utilization model, people counting model, and pedestrian flow model. Environmental conditions affect the human behavior and correspond to change in space utilization and pedestrian flow, which is monitored and carefully examined through the WSN. Certain challenges arise in the RWSN architecture and data analysis process. Therefore, to improve future research, solutions to some of the prominent challenges are proposed.

6. REFERENCES

[1] E. Yaacoub, A. Kadri, M. Mushtaha and A. Abu-Dayya, "Air quality monitoring and analysis in Qatar using a wireless sensor network deployment," *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, Sardinia, 2013, pp. 596-601.

[2] X. Liu, J. Cao, W. Z. Song and S. Tang, "Distributed Sensing for High Quality Structural Health Monitoring Using Wireless Sensor Networks," *2012 IEEE 33rd Real-Time Systems Symposium*, San Juan, 2012, pp. 75-84.

[3] W. Balid, H. Tafish and H. H. Refai, "Versatile real-time traffic monitoring system using wireless smart sensors networks," *2016 IEEE Wireless Communications and Networking Conference*, Doha, 2016, pp. 1-6.

[4] X. Li, X. Cheng, P. Gong, and K. Yan, “Design and implementation of a wireless sensor network-based remote water-level monitoring system,” *Sensors*, vol. 11, no. 2, pp. 1706–1720, 2011

[5] S. K. Viswanath, S. R. Gubba, B. Arunn, C. S. Veerappan, and C. Yuen, “On the design of a cost-effective and lightweight people counting sensor,” in *International Internet of Things Summit*. Springer, 2014, pp. 176–182.

[6] A. Vallimayil, K. M. K. Raghunath, V. R. S. Dhulipala and R. M. Chandrasekaran, "Role of relay node in Wireless Sensor Network: A survey," *2011 3rd International Conference on Electronics Computer Technology*, Kanyakumari, 2011, pp. 160-167.

[7] B. P. L. Lau, T. Chaturvedi, B. K. K. Ng, K. Li, M. S. Hasala and C. Yuen, "Spatial and Temporal Analysis of Urban Space Utilization with Renewable Wireless Sensor Network," *2016 IEEE/ACM 3rd International Conference on Big Data Computing Applications and Technologies (BDCAT)*, Shanghai, 2016, pp. 133-142.

[8] K. Li, C. Yuen, and S. Kanhere, “Senseflow: An experimental study of people tracking,” in *Proceedings of the 6th ACM Workshop on Real World Wireless Sensor Networks (REALWSN)*, 2015, pp. 31–34
[9] T. Chaturvedi, K. Li, C. Yuen, A. Sharma, L. Dai and M. Zhang, "On the design of MAC protocol and transmission scheduling for Internet of Things," *2016 IEEE Region 10 Conference (TENCON), Singapore,* 2016, pp. 2000-2003
[10] Joseph Kabara, Maria Calle, “MAC Protocols Used by Wireless Sensor Networks and a General Method of Performance Evaluation,” *2012 International Journal of Distributed Sensor Networks,* vol. 8, no. 1.