TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

NGUYEN TRUONG AN

COVERAGE AND NETWORK REQUIREMENTS OF A "BIG DATA" FLASH CROWD MONITORING SYSTEM USING USER'S DEVICES

Master of Science thesis

Examiner: Prof. Yevgeni Koucheryavy and Dr. Dmitri Moltchanov
Examiner and topic approved by the Faculty Council of the Faculty of Computing and Electrical Engineering
on 30th April 2016

ABSTRACT

NGUYEN TRUONG AN: Coverage and network requirements of a "Big data" flash crowd monitoring system using user's devices
Tampere University of Technology
Master of Science Thesis, 48 pages, 01 Appendix page
May 2017
Master's Degree Programme in Information Technology
Major: Communication Systems and Networks
Examiner: Prof. Yevgeni Koucheryavy and Dr. Dmitri Moltchanov

Keywords: network, flash crowds, monitoring, visual and aural information

Over the past few decades, aural and visual monitoring of massive people gatherings has become a critical problem of national security. In order to tackle this problem, a fixed infrastructure is used whenever possible. The aim of this thesis is to propose the system for spontaneous "flash crowd" monitoring in areas with no fixed infrastructure. The basic concept is to engage users with their mobile devices to participate in the monitoring process. The system takes on characteristics of "big data" generators. We analyze the proposed system for coverage metrics and estimate the rate imposed on the wireless network. Our results show that given a certain level of participation the LTE network can support aural monitoring with prescribed guarantees. However, the modern LTE system cannot fully support visual monitoring, as much more capacity is required. This capacity may potentially be provided by forthcoming millimeter wave and terahertz communication systems.

The analysis is conducted mainly using own custom-build simulation environment in C programming language to simulate the formalized problem. The processes and experiments run by using many resources to reduce the time consumption due to large number of calculations. And the graphs are sketched from the numerical results by using MATLAB.

PREFACE

This Master of Science thesis had been done under the guidance of Prof. Yevgeni Koucheryavy and Dr. Dmitri Moltchanov in the Department of Electronics and Communication Engineering in Tampere University of Technology, Tampere, Finland.

First, I would like to thank Prof. Yevgeni Koucheryavy for giving me a chance to do this thesis topic.

I would also like to send my sincere thanks to my supervisor, Dr. Dmitri Moltchanov, has given me opportunity to be under his guidance as well as supported me so far. His enthusiasm, suggestions and advices has been helped me a lot to complete this thesis. Especially, sometimes I get stuck at some points; he guided me to go on the right direction. I really appreciate this.

Last but not least, I would like to thank my beloved family in Vietnam, my friends in TUT, and all of you who have committed and facilitated good possible condition for the completion of my work.

Tampere, 01.05.2017

Nguyen Truong An

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

CDF	Cumulative Distribution Function
GPS	Global Positioning System
DDD	Dull, Dirty, Dangerous
GSM	Global System for Mobile
LTE	Long-Term Evolution
TUT	Tampere University Of Technology
UAV	Unmanned Aerial Vehicle
WSN	Wireless Sensor Network
GSL	GNU Scientific Library
URL	Uniform Resource Locator

1. INTRODUCTION

The question of real-time monitoring of massive people gatherings has always been of critical nature for national security. Conventionally, aural and visual information monitoring is performed using pre-installed infrastructure [17], for instance, via cameras mounted on lampposts, buildings' walls, etc., connected to the Internet access points using wired or wireless technology. Besides, in the areas where no fixed infrastructure is available and/or in case of spontaneous gatherings, helicopters are conventionally used for crowd monitoring [10]. And to have crossover the shorting coming of helicopter-based systems, recently, unmanned aerial vehicles (UAV), particularly, quadrocopters or drones have been proposed for flash-crowd monitoring [3].

Neither pre-installed media-capturing infrastructure nor helicopter- or UAV-based units are able to capture minor details of events at micro-scales, especially in highly dense environments. In addition, these systems are not capable of aural monitoring of the environment without the use of highly expensive directional detectors. In this thesis, a new monitoring system for both aural and visual information for spontaneous flash-crowd environments in the areas having no fixed infrastructures is proposed. Recalling that most modern handheld devices are equipped with relatively sensitive microphones and high-resolution cameras, the idea behind the proposed system is to explicitly or implicitly engage the users to participate in the monitoring process with their handheld devices.

1.1 Conventional Crowd Monitoring Systems

1.1.1 Pre-installed infrastructure monitoring system

For many years, multi camera based surveillance systems as pre-installed infrastructure monitoring systems have routinely used for monitoring highway traffic; monitoring crisis such as tsunami, flooding, hurricane, nuclear disaster, earthquake; and monitoring pedestrians. These systems have gained increased importance in order to increase the safety, security of a nation, and especially security of people all over the world. They also have applications in various domains like home or bank security; and in public places like airports, shopping malls, railway station, pedestrians' square, etc.

Indeed, pre-installed infrastructure system helps to monitor a given area of interest. Multiple cameras equipped with audio sensors and visual sensors are used to cover a large area. For a larger area, more number of cameras needs to be installed. The cameras can be placed at different heights and orientations on lampposts, buildings' walls, and tree's trunks, etc. They are set up in such a way that there is less overlap between their field of view. Figure 1 gives an illustration of the setup for the real-time surveillance system.

There are management servers or cluster system in order to receive the information from the cameras. The management system has a role to processing, inspecting and analyzing the receiving data for different purposes. The cameras are able to capture the information and send the raw data directly to the servers via wires or via the Internet access point. Therefore, the connections are more stable and the number of packet loss will be reduced.

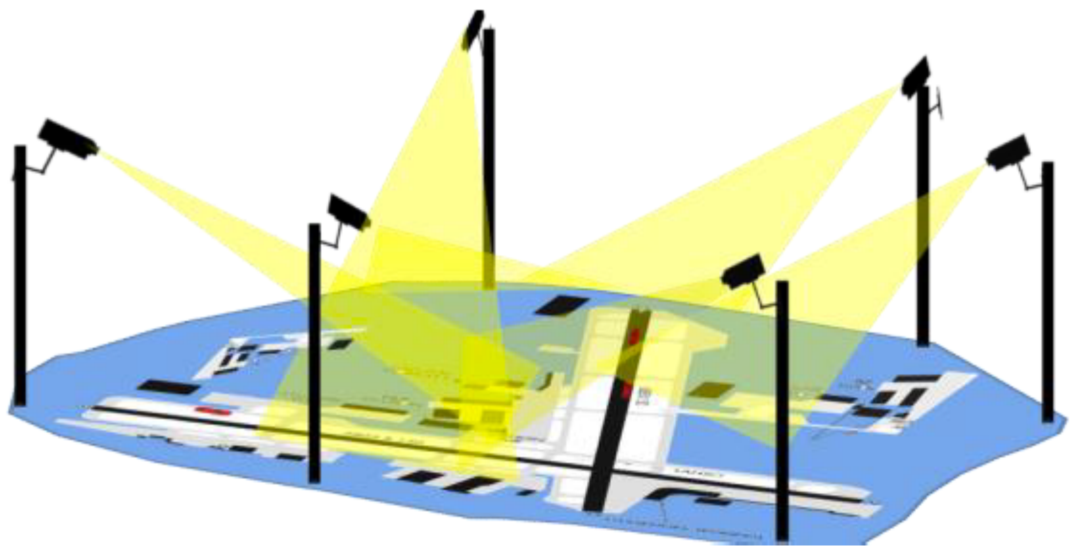

Figure 1. *Pre-installed monitoring system using multi cameras*

The main advantages of this approach include feasibility of optimal coverage planning for both aural and visual information. The existing works on coverage problem use the k-coverage as its model and coverage percentage as its optimal metric. There have been many optimal algorithms to solve this coverage problem with this model and metric. For instance, Particle Swarm Optimization (PSO) method is used to solve the coverage problem of camera networks in the target field with a high real-time quality [18].

On of the other side, it brings additional costs of infrastructure and requires apriori knowledge of area of interest making it not suitable for monitoring of spontaneous gatherings, so-called "flash crowds".

1.1.2 Helicopter-based monitoring system

In some cases, there are some areas where no fixed infrastructure is available or in case of spontaneous gatherings; helicopter-based monitoring systems are usually deployed for such as crowd monitoring or wildfire monitoring, etc.…

The forest fires cause a significantly economical loss of many countries. The damage to eco-system and the big impact on wild life and climate change needs to be considered in the long term. In the worst case, it affects directly to human lives or human health in the large area. Most of the cases are caused by natural calamity. However, poachers are also the reason and they are further threatening problems to national resources. According to a research of ICARUS Research Group from Technical University of Catalonia, a helicopter system called Red-Eye is developed to detect, monitor, control and analyze the information of the land forest [19]. The system includes the air squad and the ground control squad. The air squad has helicopters equipped with thermal cameras and their role is to fly over the area and generate a map of specific hot spots. The visual information is also processed on-board. Then the selected information will be transmitted to the operating team on the ground by communication module. After analyzing the information, the fire-fighter squad can choose the right decisions in a short time. The monitoring architecture is based on Service Oriented Architecture (SOA). The design of Red-Eye system includes five main components (Figure 2):

- A helicopter and its on-board sensor and computer system: has visible cameras, infrared and multi-spectral sensors.
- Airborne command and control station.
- Mobile ground command and control station.
- Ground team information terminals.
- Air/ground communication infrastructure.

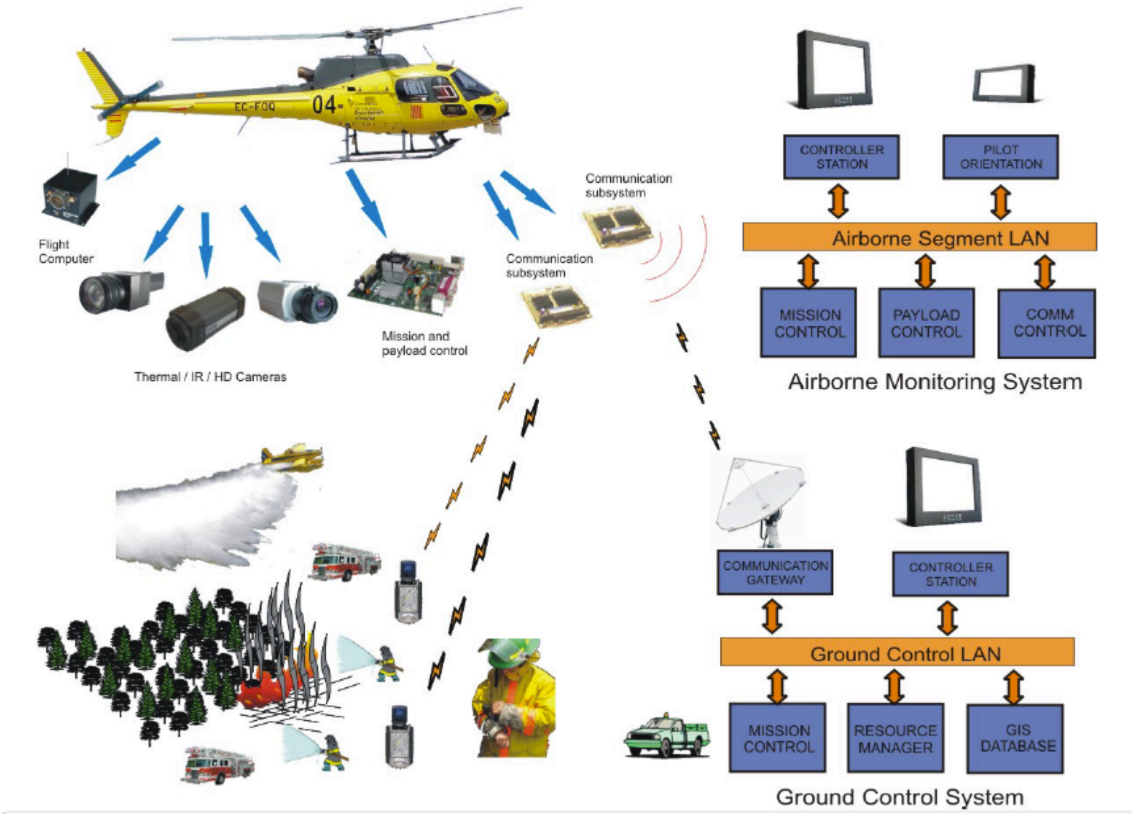

Figure 2. *Overall architecture of Red-Eye system: air and ground segments*

With current technologies, those systems like this can have many advantages not only in wildfire monitoring but also in ecological monitoring of the ocean and terrain surface. For example, the multiple demanding data can be collected even with bad weather conditions such as rain, smoke… For national security issues, those systems also used for monitoring the living things in those areas. However, such an approach is limited to visual information only and due to rather high-flying attitudes may not provide detailed information even when many advanced cameras and sensors are used. The specific aural information is not able be obtained.

1.1.3 UAV-based monitoring system

In order to alleviate the shortcoming of helicopter-based monitoring systems, recently, unmanned aerial vehicles (UAV), particularly, quadrocopters, have been proposed for flash-crowds monitoring.

UAV is known as drone by the International Civil Aviation Organization (ICAO) equipped with sensors, automatic controller, data processing units, communication systems. It can be classified into autonomous aircraft or remotely piloted aircraft. Different UAV has different size, payload and design specification but they have some common components [20]:

- RC aircraft: radio-controlled aircraft

- Avionic system: collecting in-flight data, performing automatic control laws, executing mission-oriented tasks, and communicating with the ground station.
- Ground station: monitoring the flight states of the UAV and communicating with the avionic system.
- Manual control: including the pilot and a wireless joystick.

UAV was first developed for military special operations that are dull, dirty or dangerous (DDD), for instance, monitoring battlefield, missile decoy, monitoring radioactive area. After that, UAV becomes more popular and due to its benefits and reasonable price, it is also for civilian uses. Nowadays, UAV can be used for covert role (policing and fire-fighting), research role (disaster alert, pollution monitoring) and economic reason (crop monitoring).

Burkert and Fraundorfer presented an approach for monitoring pedestrian groups using UAV-based system [3]. In the critical or normal public areas, there are a lot of large-scale events such as festivals, sport events or demonstrations, it is very important to monitor whether an abnormal scenario occurs, especially dangerous scenarios occur in moving crowds. In order to do that, occlusion-free airborne camera platform built-in UAVs systems are proposed to use to capture and utilize image data. Then, the datasets will be analyzed into group behavior models to detect number of scenarios.

The UAVs chosen for the experiment were a squad of Asctec Falcon 8 with mounted Panasonic DMC Lumix LX3 cameras as it can be seen in Figure 3. These UAVs have about 15 minutes battery life including takeoff and landing in average. After that, the battery could be changed or recharged quickly. The height attitude of these UAVs could be up 85m over the area of interest.

Figure 3. *UAV AscTec Falcon 8 model with mounted Panasonic DMC Lumix LX3 cameras*

Here is the specification of this UAV model:

- Size: 770 x 820 x 125 mm
- Max. takeoff weight: 2.3 kg
- Max. payload: 0.8 kg
- Flight time including payload: 12-22 minutes
- Max. range: 1000 m
- Max. air speed: 16 m/s
- Tolerable wind speed: 15 m/s (GPS: 12 m/s)
- Data connection: 2 x 2.4 GHz, 10-63 mW
- Transmission power: 5.8 GHz, 25/100 mW
- Power unit: Rechargeable LiPo batteries with 6250 mAh
- Operating temperature: -5^0 – 35^0 (ideal condition for professional use)
- Remote control (RC): Mobile Ground Station, Waypoint navigation
- Inertial guidance system & sensor: AscTec AutoPilot or AscTec Trinity with 1000 Hz update rate
- Flight modes: GPS mode, Height mode, Manual mode
- Safety modes: Direct landing, Come home straight, Come home high
- Certification: CE, RoHS

Hence, at the medium people densities, monitoring time of a specific place could be longer and the images of individuals were more visible and distinguishable. Compare with using helicopters or airplanes, the camera mounted on them could view a specific area only for a few seconds during flight even though they could view a large field of vision due to the high altitude of 1000m or more. Thus, in spite of much lower cost of use compared to helicopters and potentially lower altitude allowing to achieve better resolution, such system are still not suitable for aural information monitoring even though the generated noise is much lower than that of helicopters. The limited flying time requiring frequent and automatic recharge as well as the need for manual navigation adding to the operational costs are additional shortcomings of the system. Figure 4 are shown the comparison about resolution of image captured by UAV Falcon 8 and helicopters. Moreover, both systems can only be used at good weather conditions.

Figure 4. *Image comparison of pedestrians from the UAV Falcon 8 and the airborne 3K-camera system*

1.2 New monitoring system for spontaneous flash-crowd environments

In this thesis, a new flash-crowd monitoring system for both aural and visual information using users' modern handheld devices in the areas having no fixed infrastructure is proposed. The users with their built-in sensitive microphones and high-resolution camera handheld devices can explicitly or implicitly participate in the monitoring process.

This is an advantage for using handheld mobile devices in this approach because it is more and more popular nowadays. According to a report from The Statistics Portal, a prediction about number of smartphone users worldwide from 2014 to 2020 has been made. It is illustrated in the following graph (Figure 5):

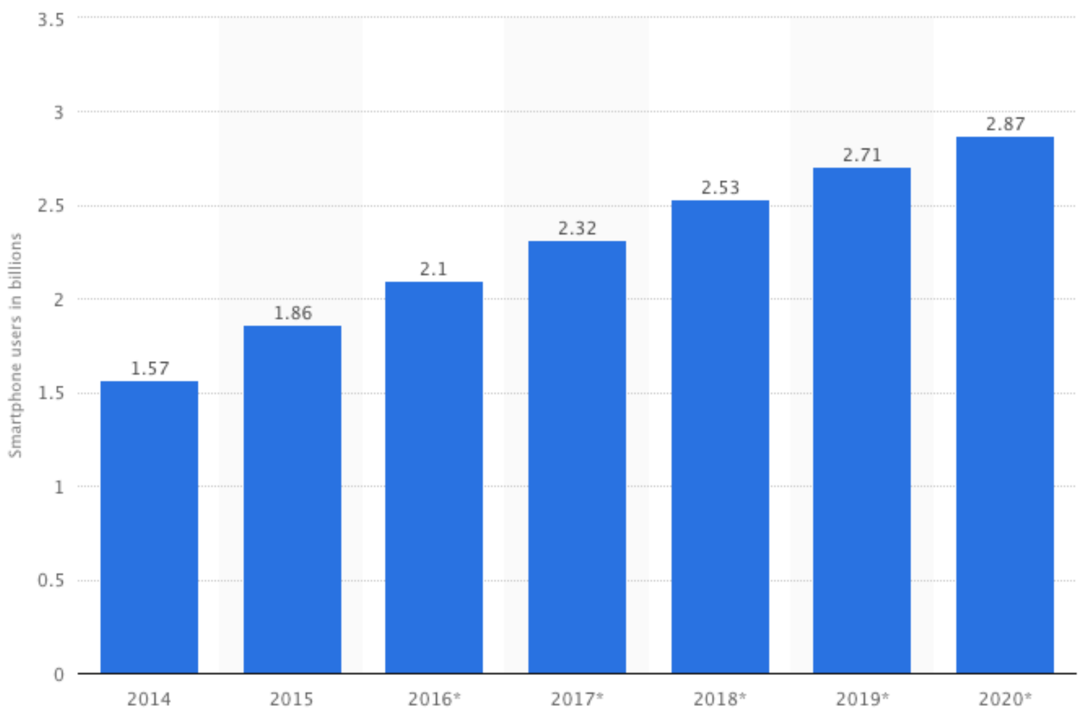

Figure 5. *Number of smartphone users worldwide from 2014 to 2020 (billions)*

It can be seen from the graph that the number of smartphone users is predicted to reach 2.1 billions in 2016 and it grows to around 2.87 billions in 2020. According to Net Market Share in July 2014, there are currently two most popular smartphone operating systems that are Google's Android (45%) and Apple's iOS (44%) (as shown in Figure 6). And in Q4 of 2016, number of android devices occupies 81.7% of whole market share according to Gartner report. Therefore, the end-user application for flash crowd monitoring system can be considered to base on these two mobile operating systems.

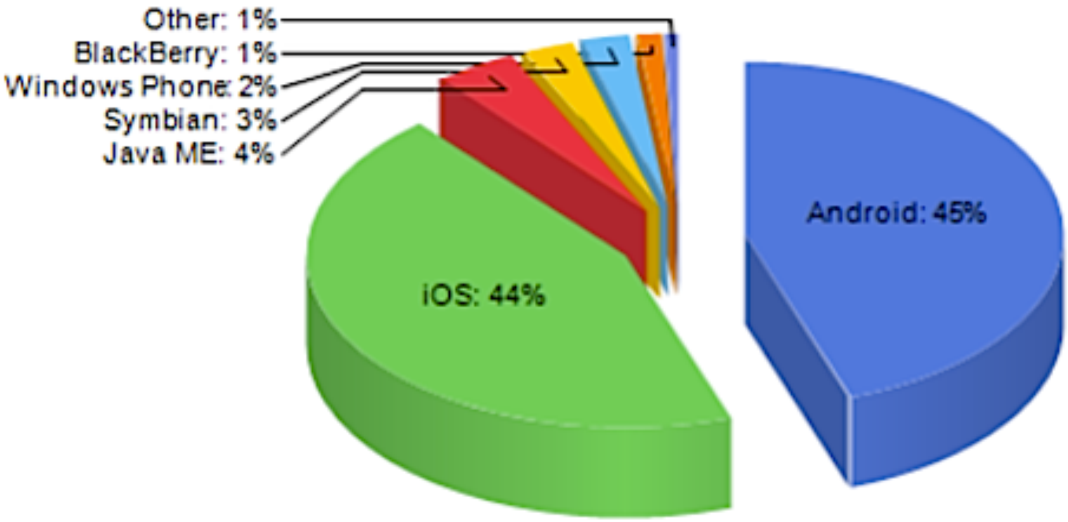

Figure 6. *Mobile operating system market share*

By downloading and installing an application, a user may explicitly engage himself/herself to the monitoring process. It can be seen in Figure 7 that people use their smart devices to record the video with cameras and microphones. They may explicitly involve in monitoring process for flash crowd monitoring purpose. Assuming the uniform distribution of users over the monitored area, the coverage metrics are obtained including the cumulative distribution function (CDF) of the covered area, mean and quantiles for both aural and visual information. For visual information, humans located in the area are explicitly taken into account the blocking of camera view. Then these metrics are translated into the rate required from the network for various audio and video codecs. In addition, the proposed system is compared to the optimal infrastructure-based deployment. The numerical results allow making the following conclusions:

- The capacity of modern LTE system is sufficient to provide audio monitoring of the areas of interest with prescribe coverage metrics;
- Novel wireless communications systems, such as those operating in millimeter wave or terahertz frequency bands, are needed for visual monitoring.

Figure 7. *Illustration of users engage themselves to flash crowd monitoring process using their smart devices*

In addition, due to the shortcomings of using infrastructure-based, helicopter-based, or UAV-based monitoring system, the new proposed system can overcome their issues. The comparison is made according to the following aspects: aural information, visual information, performance, lifetime, detailed of covering nodes and cost.

2. SYSTEM DESIGN

In this section, there is the concept of the proposed flash-crowds monitoring system including the specification of modern handheld devices and ways of collecting information. Then defining the metrics of interest including both coverage and network rate requirement are proceeded. Finally, the problem is formalized and the related work pertaining to the subject of interest is also reviewed.

2.1 The Concept

2.1.1 Handheld mobile devices

Nowadays, a high percentage of handheld mobile devices are equipped with integrated media capturing equipment including microphones and cameras. And these devices are used for the proposed flash crowd monitoring system. Depending on implementation, users can be explicitly or implicitly engaged into the monitoring process. Explicit engagement presumes an application that users download and run on their mobile devices. In implicit engagement scenario, users are not notified in advance about their involvement in the monitoring process. For obvious ethical reasons, this thesis will not be considered the latter as viable solution.

There are different specifications of handheld mobile devices due to different models, but they are often share the common feature such as audio and video capability, operating system, software applications, virtual keyboard, messaging features, camera and microphone features, internet access, flash light.

Here is an example of smartphone specification with Sony Xperia M5 running Android OS (according to Sony Mobile specification) as a mid-range smartphone model:

- Network technology: GSM / HSPA / LTE
- Display: 5.0 inches IPS LCD capacitive touchscreen, 16M colors
- SIM: Nano-SIM
- OS: Android 6.0.1 (Marshmallow)
- Chipset: Mediatek MT6795 Helio X10
- CPU: Octa-core 2.0 GHz Cortex-A53
- GPU: PowerVR G6200
- Memory:
  - Internal: 16 GB, 3 GB RAM

- External: microSD up to 256 GB
- Camera: 21.2 MP, f/2.2, phase detection autofocus, LED flash, touch focus, face/smile detection
  - Resolution: 2160p@30fps
  - Secondary camera: 13 MP, f/2.0, auto focus, 1080p@30fps
- Microphone: built-in 2 microphones, one from the front and one from the back of smartphone
- Sensors: Accelerometer, proximity, compass
- Battery: 2600 mAh
  - Video record time: up to 12 hours

For this purpose, we focus mostly on the built-in camera, microphones, and network capability of a smartphone as the must-have features. Figure 8 shows an overview of a common handheld device with its components:

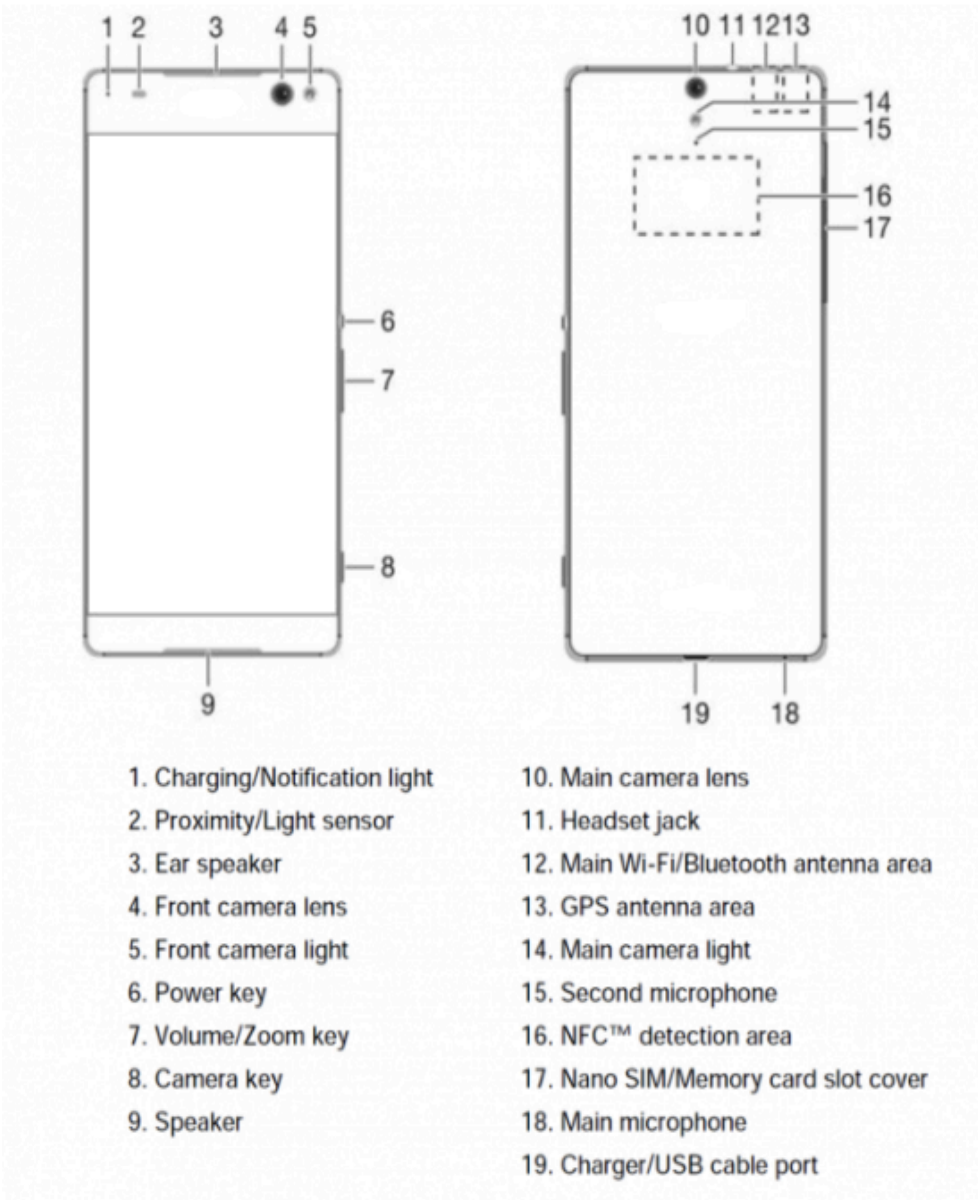

Figure 8. *A common smartphone with its components*

There are usually included two microphones appearing in a smartphone nowadays in order to recording voice and audio for videos. The technology for the sound is advancing day by day to achieve better performance, for example, to have better clarity and ease. Besides, microphone technologies are developed to adapt of highly features smartphones as well as enhancing manufacturability, reducing size and cost. They are designed for audio processing such as echo cancelling, noise cancellation, wind noise filtering, beam steering, 3-D sound and other interesting effects. Sound signal processing will become pervasive as phone producers plan to further differentiate their products with more sophisticated audio features. For many years, manufacturers have been striving to improve the performance of electret condenser microphones (ECMs), including sensitivity, signal-to-noise ratio (SNR), and reflow soldering. Analogue to digital converter (ADCs) ICs, designed for microphones based on micro electromechanical systems (MEMs), are now significantly contributing to improved microphone performance. Thus, handheld devices' microphones are moving from analogue to digital technology.

In addition to microphone technology, camera technology also plays an important role for the monitoring processing via handheld devices. Standalone cameras are in a large range of types, which depend on the purpose of use: super-zooms are good for photographing far objects such as capturing birds, wild animals or airplanes. The medium format provides the detail needed for magazines and posters. Macro cameras are for capturing closely, and broadcast cameras are used for streaming video. Along with these, smartphone cameras are designed to be the most versatile without being complex. Their sensors are small because of the small footprint of a handset; lens's focal lengths and apertures are fixed to reduce the number of moving parts, the lens has a wide-angle to be the most useful for normal shooting situations, and there's not a lot of extra hardware to accompany the sensor and lens. Megapixel count, sensor size, pixel size and focal length of handheld devices are described:

- Megapixel count: Many marketing materials, specifications focus a lot in megapixel of a camera which people are very familiar with. The reason is that is every easy to keep track: a higher megapixel count reflects to more details that can be used for creating, cropping and zooming images. On a smartphone camera, zooming can be particularly important due to fixed-focus lenses where there is a large megapixel count so that detail of photos is preserved. Having a high megapixel count is good, but it does not tell how a camera performs overall. A typical tradeoff with having a sensor packing many millions of pixels is a small pixel size, but conversely having too few pixels makes images look bad through a lack of detail. All camera manufacturers understand this trade-off, which is why smartphones with sensors packing is often between 5 and 20 megapixels recently.

- Sensor size: Sensor size tells a lot of the other important values related to a camera system, such as the f-number, focal length, and its crop factor. Fortunately, smartphone manufacturers figure all of those things, and the light gathering properties of the sensor are the things should be considered more. In theory, a larger sensor has more area for the light to fall on, equating to a greater ability to gather the light, assuming the mega-pixel count stays the same. The size of a smartphone sensor is typically given as a fractional number in inches (for instance, 1/2.3", 1/3.06"), which may appear to give the diagonal dimensions of the sensor, but it actually does not. Instead it refers to a type of sensor with a diagonal size around one third smaller. For instance, a 1/2.3" sensor typically has a diagonal of 7.66 mm, rather than the expected 11.04 mm one indicated by the number itself.

- Pixel size: Sensor size is useful for getting an idea of how much space in the smartphone camera module is consumed by the sensor, but less useful for estimating total light collection as megapixel counts vary between smartphones. This is where pixel size, giving a direct measure of how large the individual photo detectors are in the CMOS sensor. Pixel size for smartphones fits into a narrow ranges between one and two micrometers (µm) in either the horizontal or vertical direction. For many years ago, each pixel collects the light. The larger pixels are, the more light captures. The technology behind the design of the CMOS sensor can affect the light gathering properties of each individual pixel, but the easiest way to compare is just by focusing on size. A camera with 1.4-micrometer pixels captures twice the light (per pixel) of one with 1.0-micrometer pixels, calculated by comparing difference in total area. Another way of saying this is that the 1.4-micrometer sensor is one stop brighter.

- Focal length: Focal length is the distance between the lens and the sensor, which determines the field of view and magnification. The actual focal lengths for most smartphone cameras are not very useful for people familiar with photography terms because of the small sensors. Therefore 35mm-equivalent focal lengths are usually mentioned instead. 35mm-equivalent focal lengths reflect what focal length the camera's lens would need to have if it were to produce equivalent images on a DSLR with a 35mm-format sensor. It is then easy to bracket the camera lens into different types based on the knowledge about lenses for traditional 35mm cameras: 18-35 mm designates a wide-angle lens, 35-60 mm are for normal lenses, and over 60 mm are for long-focus (as known as tele-lens or zoom lens). All smartphones fall into the wide-angle lens bracket, typically somewhere around 24-30 mm; the larger the number, the less wide angle the lens is. It is shown in figure 9 about field of views of a smartphone camera.

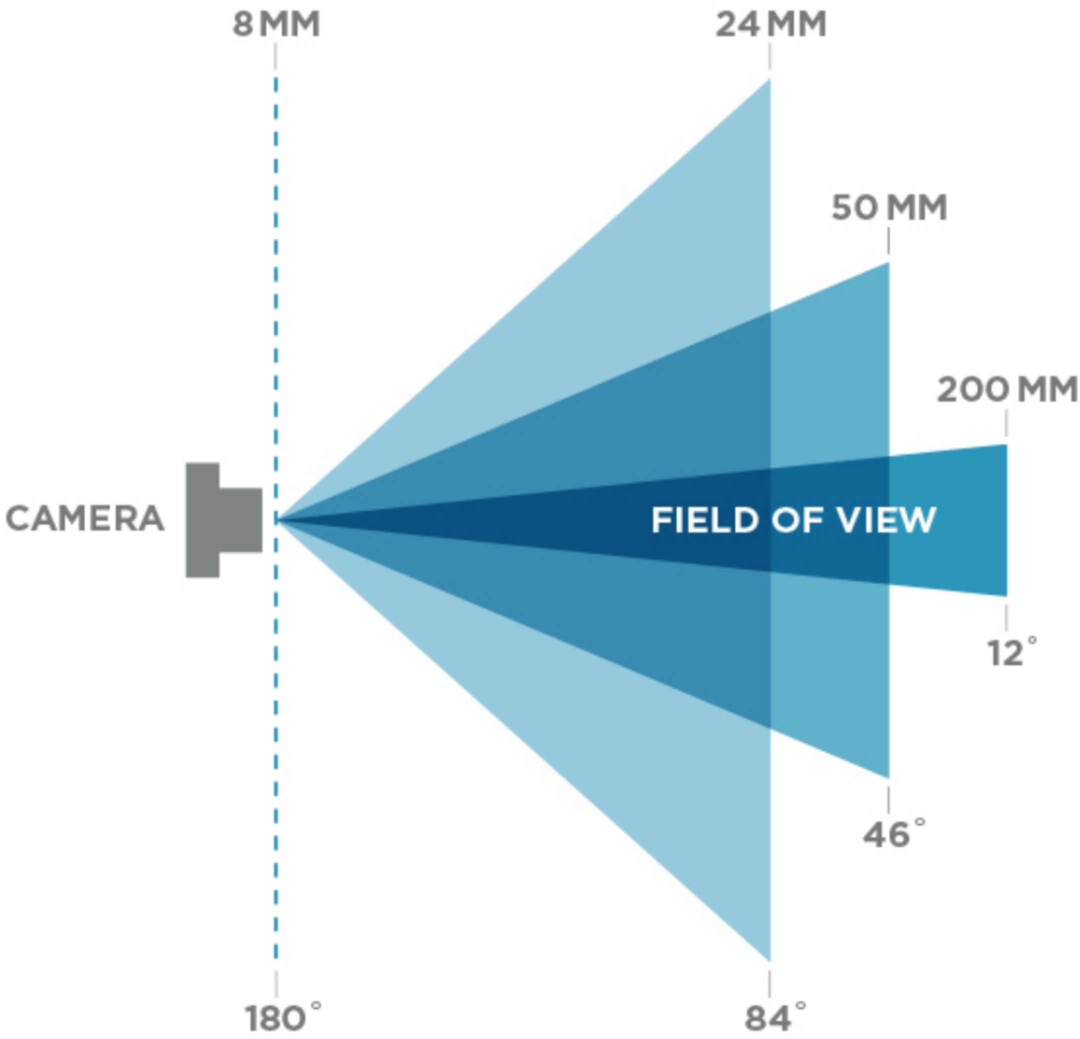

Figure 9. *Field of view of a smartphone camera*

The benefits of the proposed system compared to infrastructure-, helicopter-, or UAV-based monitoring systems are that media is captured at much closer distances inside the flash crowds and that there are potentially a large number of devices providing the coverage. In addition to aural and visual information, modern smartphones equipped with numerous advanced sensing capabilities can also provide other types of information including remote sensing and telemetry. The same principle can be used for environmental monitoring applications.

2.1.2 Ways of collecting information

There are two ways of gathering information from devices participating in environment monitoring. A smartphone-based application could itself provide the logic for information analysis gathered by the audio and video sensors. There are a number of shortcomings associated with this approach. First, the devices shall be extremely powerful as in most case the information needs to be processed in real-time. The question of the use of resources not only concerns the processing power and memory but also be related to the high battery usage by applications performing real-time data processing. Although there might be additional incentives to participate in the monitoring campaign except for the "good will" of a user, the aggressive use of limited resources may prohibit the widespread use of the application. Further, there are security concerns as smartphone-cased information processing requires that the knowledge of the monitoring task to be available at the user devices. Finally, local information processing may not be useful as a single node may not have enough of data to make conclusive decisions. Indeed, the strength of the proposed system is in the ability to get information from many sources located nearby. Thus, the information shall be delivered first to the certain remote server for further centralized data processing.

The mobile devices participating in the monitoring process are expected to use the resources of cellular system uploading the data to the remote server. For specific applications such as flash crowd monitoring, the density of nodes willing to simultaneously use the cellular connectivity can be extremely high and may easily overload the network not only preventing it from handling the data of security application but serving normal connections as well. At the same time, in certain cases there is no need to have more than few nodes to monitor a certain point in space simultaneously. As a result, the external monitoring system shall be capable to turn off remote sensing capability of some nodes.

In this thesis, it is only to concentrate on the flash crowd monitoring systems. In this context, the problem is formalized as following: for a random placement of users on the landscape what should be the density of nodes providing coverage for a certain type of media such that a percentage of area is covered with probability of x. Once this question is answered, the amount of wireless network resources need to monitor the area of certain dimensions is paid attention.

2.2 Related Work

2.2.1 Coverage

Based on the description of the system and metrics of interest, one could notice that the problem at hand reduces to finding coverage of a space in R^2 by sets of special configuration. This problem has been extensively studied in the literature. Recent advances in this area are mostly associated with coverage of wireless sensor networks (WSN), where the set of interest is a communications range of a node having circular form. Several advanced results have been reported so far, including simple and elegant solution proposed by Lazos and Poovendran [9], where they use the integral geometry, particularly, the notion of kinematic density, to provide simple closed-form results for k-coverage problem in WSNs under any distributions of audio sensors. Figure 10 illustrates a heterogeneous sensor network with sensors covering the deployment area. Recalling the system model one may observe that this methodology is directly applicable for aural information.

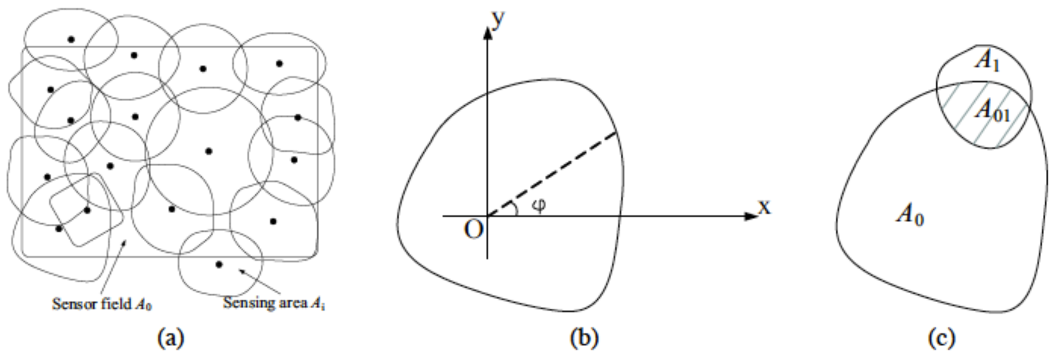

Figure 10. *(a) A heterogeneous sensor network with sensors covering the deployment area. (b) A convex set A and the corresponding quantities defining the kinematic density. (c) Two convex sets A_0, A_1 intersecting, and the common area A_0*

Denote A_0 (F_0, L_0) is the monitoring area with perimeter L_0 and area F_0. Assume that N nodes are distributed according to K (A_0) distribution over sensing area (A_0) in a way that they cover parts of interesting field. Each node has a sensing field A_i (F_i, L_i), (I = 1…N) where L_i, F_i are the perimeter and area of sensing area respectively.

Based on the kinematic density and motion of sensor nodes, the stochastic models of coverage area in this case is given by two models:

- The fraction of A_0 that is not covered by any sensor when N sensors are randomly deployed or the probability that monitoring area A_0 is not 100% coverage.
- The probability that a random selected point of A_0 is covered by at least k ($k >= 1$) sensor(s).

The fraction of A_0 that is not covered by any sensors when N sensors are randomly deployed is given by this equation 1 [9]:

$$p(S=0) = \prod_{i=1}^{N} \frac{2\pi F_0 + L_0 L_i}{2\pi(F_0 + F_i) + L_0 L_i}$$

The probability that a randomly selected point of A_0 is covered by at least k sensors is given by this equation 2 [9]:

$$p(S=0) = \frac{\sum_i^{\binom{N}{h}} \left(\prod_{j=1}^{N}(2\pi F_{T_{i,j}}) \prod_{z=1}^{N-h} \Theta(i,z)\right)}{\prod_{r=1}^{N}(2\pi(F_0+F_r)+L_0L_r)}$$

$$\Theta(i,z) = 2\pi F_0 + L_0 L_{G_{i,z}}$$

$$p(S \geq 1) = 1 - \sum_{h=1}^{k-1} p(S=h)$$

$$= 1 - \frac{\sum_i^{\binom{N}{h}} \left(\prod_{j=1}^{N}(2\pi F_{T_{i,j}}) \prod_{z=1}^{N-h} \Theta(i,z)\right)}{\prod_{r=1}^{N}(2\pi(F_0+F_r)+L_0L_r)}$$

Where $T_{i,j}$ is a matrix in which each row I is a k-permutation of [1…N], $G_{i,z}$ is a matrix in which each row I contains the elements of [1…N] that do not appear in the i^{th} row of $T_{i,j}$.

Based on these two equations, the coverage of aural information can be resulted.

2.2.2 Visual coverage in stochastic deployments

Assessing performance of visual coverage in stochastic deployments is much more complicated. The reason is that individual objects (humans) block the viewing field of cameras. Figure 11 provides a simple illustration of the complex region visible from three cameras in presence of a single blocker.

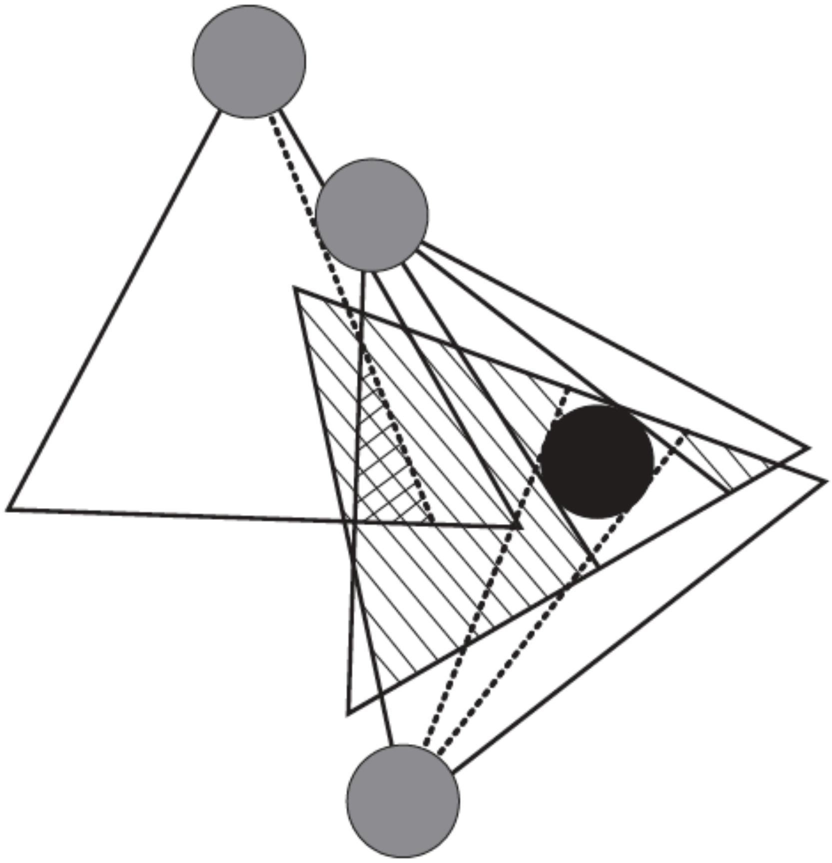

Figure 11. *An illustration of the complex visibility region in presence of a blocker in a certain point in a field*

It also highlights the redundancy associated with the monitoring process. The problem of visibility in the random field of blockers has been addressed in the context of search in forests and, more recently, in context of extremely/tremendous high frequency electromagnetic wave propagation (EHF/THF) in crowded environments [1,5]. As demonstrated by recent measurements, millimeter wave frequencies render themselves quite sensitive to “blocking” caused by obstacles such as humans, vehicles, etc.…

Figure 12 shows an illustrated scenario, according to [5], followed by spatial model including a Tx located at a particular height h_T above the ground and a Rx located at the height h_R.

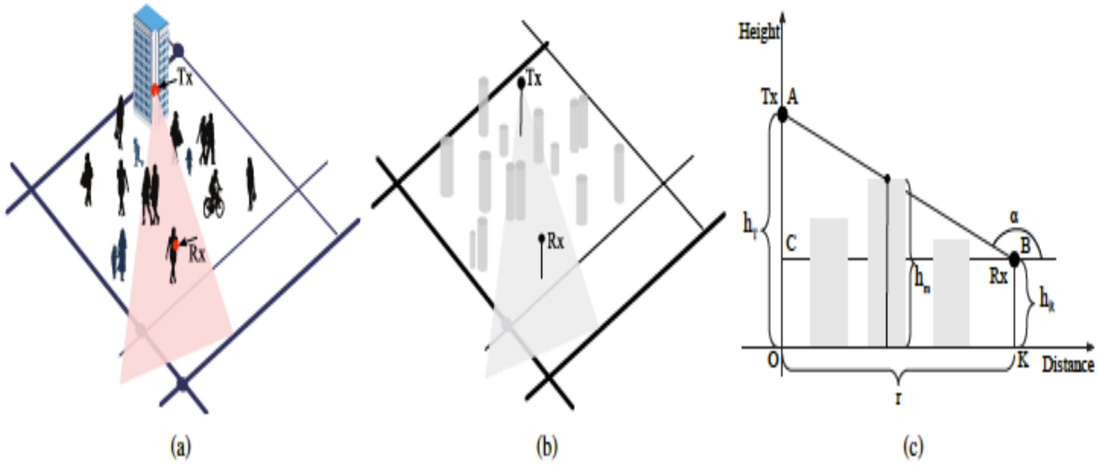

Figure 12. *The considered scenario for analytical modeling*

The potential blockers (humans) are distributed over a specific area. The blockers is modeled as cylinders with a height H and the base diameter of D. Both H and D are random variables (RVs). The distribution of the height for men and women is Normal with the mean and the standard deviation is provided. The mixture of users is closely approximated by the Normal distribution H ~ N(μH , σH). In theory, any distribution could be used to provide a result based on the current methodology. The random variable of D is assumed to be uniformly distributed between d_{min} and d_{max}. The centers of cylinder bases follow a Matern hard-core point process on the plane with the intensity λ_I. The length of the Rx is assumed to be lm.

The model can be extent to incorporate height [1]. For a link OX of length R in R^2, the height of the base station is H_B, and H_U is the height of the user using mobile devices. Ignore the loss of generality, assume that $H_B > H_U$. The height of k-th blockage is H_k according to probability density function f_H. Assume that W = 0, for instance, use line segments process to describe random buildings. The case rectangle process can also be extended to incorporate building height in the similar way.

Denote K as the number of blockages that effectively block the direct propagation of the link OX when considering the height of blockages. Note that even if the projection of a building on the ground crosses OX, in practice it might not be height enough to totally block the link as shown in Figure 13.

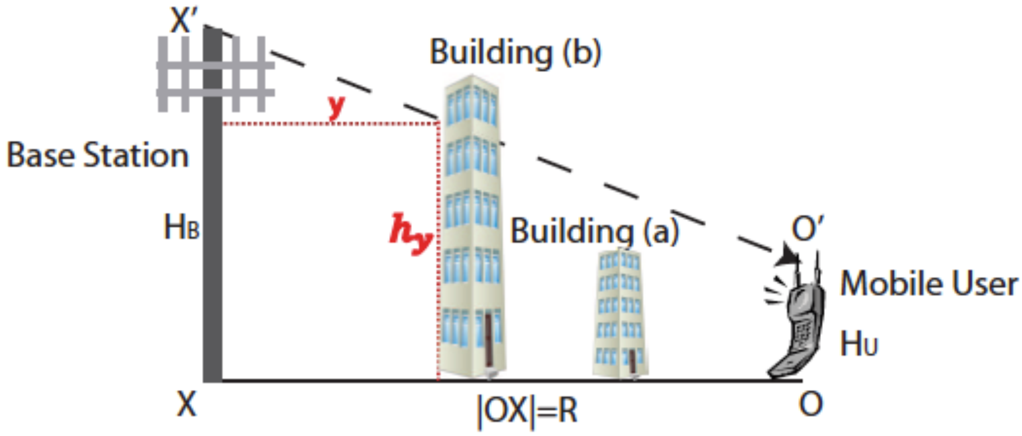

Figure 13. *The transmitter locating at X has a height of H_t, while the mobile receiver has a height of Hr. Not all buildings which cross OX blockage the actual propagation path O'X' in R3, such as building (a) in the figure. If a building intersecting OX at a point y away from the transmitter X effectively blocks O'X' if and only if its height is larger than h_y as building (b) in the figure*

Assume that the building intersecting the link OX at the point which is at a horizontal distance y away from X. The building blocks the direct propagation path O'X' if its height $h > h_y$, where h_y can be computed as:

$$h_y = \frac{yH_{\mathrm{B}} + (R-y)H_{\mathrm{U}}}{R}.$$

After that, the intersection between the building (b) and the link OX is uniformly distributed across the link, which indicates y is uniformly distributed on [0, R]. Thus, the probability of blocking O'X' is:

$$\begin{aligned}\eta &= \frac{1}{R}\int_0^R \mathbb{P}\left[h > h_y\right] \mathrm{d}y \\ &= \frac{1}{R}\int_0^R \left(1 - \int_0^{\frac{yH_{\mathrm{B}}+(R-y)H_{\mathrm{U}}}{R}} f_H(h)\mathrm{d}h\right) \mathrm{d}y \\ &= 1 - \int_0^1 \int_0^{sH_{\mathrm{B}}+(1-s)H_{\mathrm{U}}} f_H(h)\mathrm{d}h\mathrm{d}s,\end{aligned}$$

Since η is only determined by the distribution of the heights, which is independent of K, and K' can be viewed as the result of independent thinning with a parameter of η. So it is also Poison, and $E[K'] = \eta E[K]$.

Also note that incorporating the height of blockages only introduces a constant scaling factor η to the results that ignore height. Thus, the results can be readily modified to account for the heights by incorporating the η factor appropriately.

However, in all those studies the metric of interest was the probability that a certain point in a field of blockers is visible, not the total visible area.

3. PERFORMANCE MODELING

In this section, it is to introduce the system model using visual information are the media of interest. Further, the proposed simulation environment for performance analysis of the considered monitoring process is also described.

3.1 System Model

The type of a sensor affects the coverage area of a single node. There is two types of sensors need to be concentrated: aural and visual sensors.

3.1.1 Visual Information

For visual sensors, such as cameras capturing video or still images, the field view is by default of sectoral shape with radius r_V. For convenience, it is modeled as an isosceles triangle with the height to the base r_V and apex angle α. To include a random orientation of cameras, assume that the bisect of the apex angle is uniformly distributed in $(0, 2\pi)$. Humans that fall into coverage field including those participating in the monitoring process block view as shown in Figure 14.

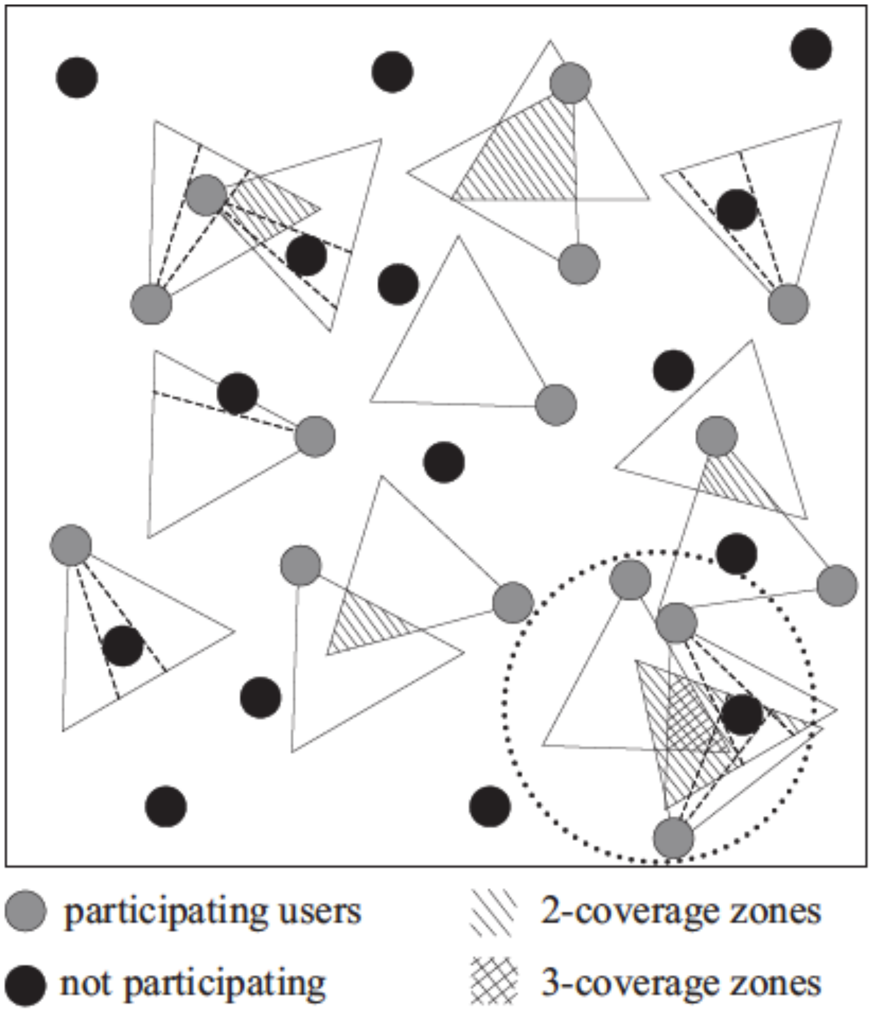

Figure 14. *The illustration of visibility in the dense crowd*

Considering the process of visual flash crowd monitoring system. Since the height of user devices is assumed to be comparable with the height of blockers (humans), it can be limited the interests to two-dimensional scenario. Fixing a certain time instant t, the snapshot of a system is also illustrated in Figure 14. The area being covered is assumed to be 100 by 100 meters. The humans are represented by circles on the landscape of diameter d. There are overall N + M humans in the area comprising a crowd to be monitored. N humans are assumed to follow a conditional Matern process with parameter d in the area [4, 14]. Matern's hard-core processes are valuable point process models in spatial statistics. In order to extent the field of application, Matern's original models are generalized both as point processes and particle processes. The distribution of the number of humans over the landscape is required to quantify the effect of penetration losses. Therefore, the centers of humans are generated and deployed based on Matern process can be replaced by the equivalent Poisson process with a wide range of intensities λ_I. Further, observe that for different values of h_T, h_R, and the distribution of the human height H, there is not all the blockers (humans) affect the experiments. The number of humans should increase as the x-coordinate grows from 0 to r.

In other words, no two users could be closer than at the distance 2d to each other as in practice human bodies do not overlap. M additional humans participate in the monitoring process and they also follow conditional Matern process with parameter d. Thus, the overall number of potential blockers for viewing field of camera is M + N.

Comparing with above deployment, the optimal deployment for M humans participating in monitoring process is also designed. In the explicitly case which M users are engaged to the monitoring process so that they can be planned in advance to place in the planned position in the monitoring area as the grid rectangular field. Assumed that the required numbers of humans with cameras are placed in one point such that they cover a circle 2 π . Also considered the case that the camera angle is the same with other camera models in the monitoring field. Therefore, for an example of a camera angle of 60 degrees, one need to place 6 cameras at the same specific point to cover 2 π . There is 2 methods can be applied for this optimal deployment: clover method and hexagonal method (as known as honeycomb pattern). Hence, the minimal number of participating humans with cameras can be obtained. The figure 15 is drawn to illustrate this visual optimal deployment:

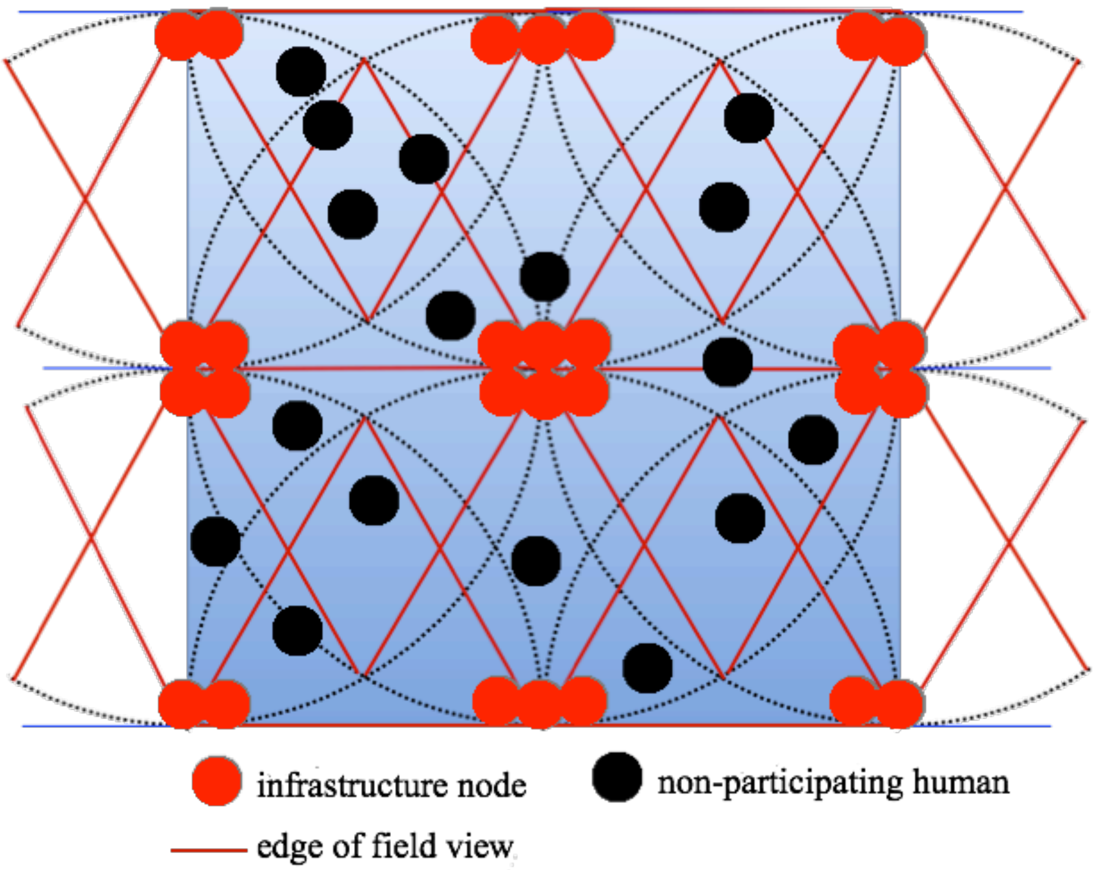

Figure 15. *The optimal deployment illustration for visual coverage*

3.1.2 Aural Information

For audio sensors, such as microphones, the assumption of circular coverage with radius r_A around a user is taken as humans do not block acoustic waves propagation significantly [13]. In normal speech communication, a sound wave is both the end product of the speech production mechanism and the primary source of raw material used by the listener to recover the speaker's message. There are various types of sound, such as human voices, buzzes, hisses and pops… and they are the results from vibration producing air pressures disturbance. Indeed, the effect of acoustic waves through human body is investigated. Acoustic waves can go around the human body within their radius, but it can be reduced when go through the human body. Therefore, we ignored the coverage of audio inside human body while designing the system model.

The distribution of M+N humans for collecting aural information remains the same with collecting visual information in considering that M users participating in the monitoring process trigger microphone for recording voice or while using recording video. The illustration in the figure 16 is added the circular coverage of microphone sensors.

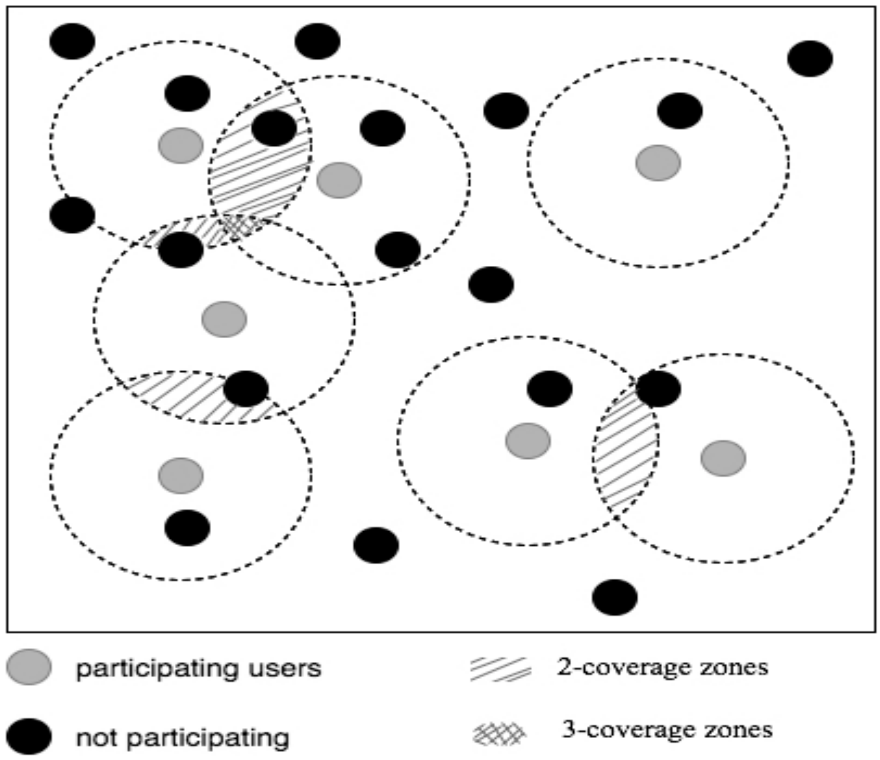

Figure 16. *The illustration of aural with circular coverage in the dense crowd*

Besides, similar with visual optimal deployment of camera, in case of collecting aural information, the estimation of the minimal number of microphones when perfect planning deployment is assumed that 100 percentage of area of interest is covered. The difference is only one handheld device with built-in microphone are needed at a certain point in the rectangular area. It is considered like an infrastructure node (shown as figure 17). However, the number of needed microphones is signification less than cameras in this case. This method is also done for different sensing area of a microphone later in simulation.

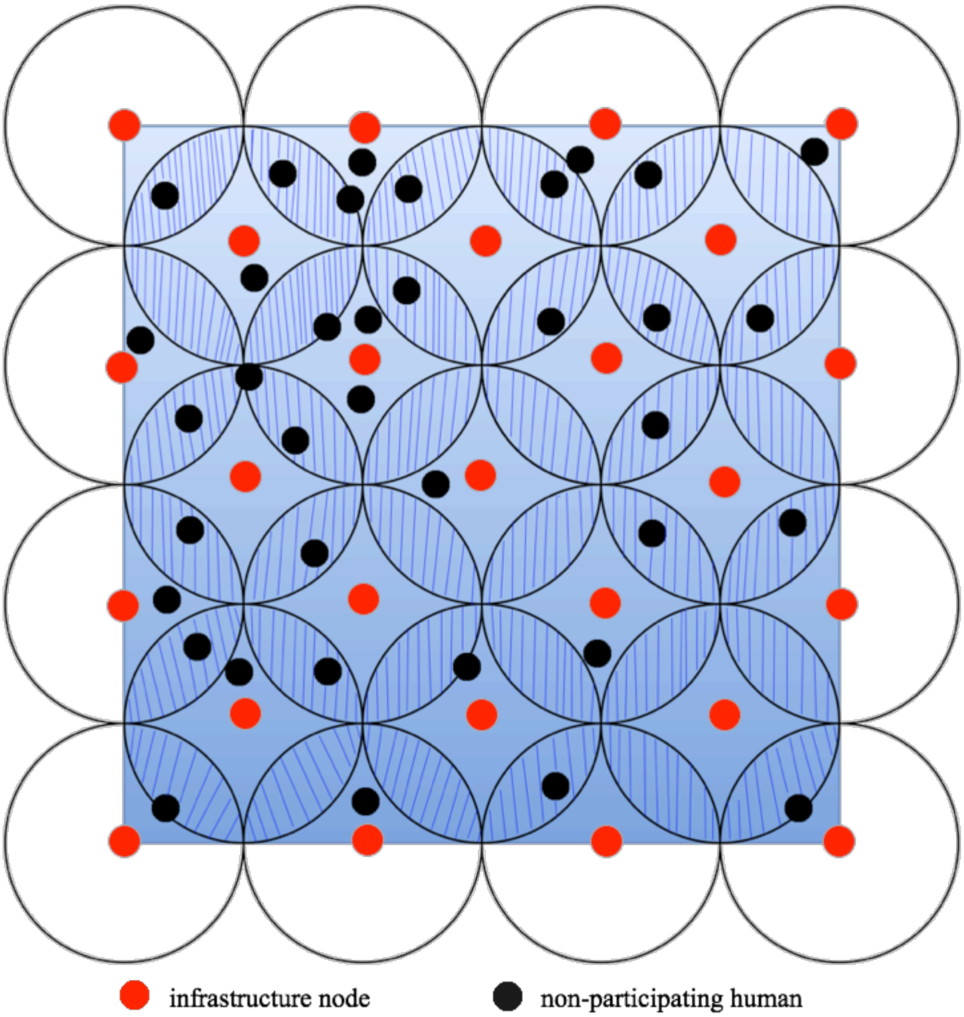

Figure 17. *The optimal deployment illustration for aural coverage in which fully cover a rectangular area with minimum amount of fixed radius circles from microphones audio sensors*

3.2 Simulation Environment

3.2.1 Custom-build simulation environment

To analyze the formalized problem, we have used our own custom-build simulation environment written in C. The choice of high-performance programming language is dictated by the complexity of the coverage area estimation. And GNU Scientific Library (GSL) is also included to the programming for using a wide range of mathematical issues such as random number generators (for generating humans or blockers) and some special functions.

Besides, using simulation methodology will get the advantages of no restrictions results; model structure, algorithm and variables can be quickly changed. However, this kind of method is a very much time consuming method. On the other hands, Analytics methodology gives the straightforward results and accuracy but the drawback of it is to results a restrictive assumptions. Table 1 shows the comparison between analytic method and simulation method in performance modeling.

Simulation		**Analytic**
Measured or invented	**Input Parameterization**	Measured or invented (with certain limitation
Virtualization	**Model Components**	Composed of limited basic building blocks
Anything that can be measured	**Model Outputs**	Equilibrium measures
Arbitrary	**Effort to construct model**	Modest
Typically large	**Computational Cost**	Typically small
Probability / Statistics	**Underlying Concepts**	Algebra to stochastic process
Credible	**Special Properties**	Insight and Optimization

Table 1.*Analytics versus Simulation Methods*

In order to obtaining the numerical results after running the simulation by C programming, Matlab is used for analyzing and processing the aural and visual information from those raw data.

3.2.2 The process

Modeling stochastic patterns of humans in the monitored area is critical for accurate performance assessment. To construct a conditional Matern process with N + M users, it needs to be checked the condition $(N + M) < N^*$, where N^* is the number of humans corresponding to dense circle packing [8]. Further, for each individual human, we first generated its (x, y) coordinates and checked the condition of non-overlapping. If a newly generated human overlaps with already existing ones, coordinates are re-drawn and the process continues up until all the humans are generated. Once M + N humans are generated, M of those are chosen as the ones participating in the monitoring process. They are further assigned audio or video sensor coverage.

Coverage analysis is the most time-consuming procedure. The grid method is used consisting in division of the area of interest into the lattice grid and checking whether nodes of a grid are covered or not [2,11,12,15]. The step of the grid is the parameter severely affecting the trade-off between accuracy of analysis and performance of the simulation framework. In this simulation, it was set to 0.1 of a meter.

The aim of the experiment process is to determine the metrics of coverage assessment. To process this, the program is required to input the required parameters in order to getting the results of covered nodes and uncovered nodes by checking each point of lattice grid. The detailed information can be seen in the source codes with the detailed comments.

- **Input:**

  + Fixed area size: 100 meters x 100 meters

  + PPP's density: 0.1

  + Radius of a blocker (human): 0.5 meters

  + Number of nodes (microphones / cameras): to be increase up to 1000 nodes

  + Number of rounds for running experiments: 1000 rounds

  + Interval x (device the horizontal OX axis into x intervals): 100

  + Interval y (device the vertical OY axis into y intervals): 100

  + Camera field of view (wide angle): 60 degree (for normal 35mm mobile camera lens with the focal length of F = 50mm in horizontal dimension).

  + Aural radius of microphone: 1m / 3m / 5m

  + Visual radius of camera: 5m / 15m / 25m

The output will be written in to files as the data for further analysis.

The programs and algorithms as the solution for the experiments, we divided to separate programs and run them separately for each case of aural and visual information issues. After that, the experiments for calculating the aural/ visual coverage are done with optimal placement of minimal number of nodes (microphones/ cameras) in the area of interest.

There are the main functions of the program:

a. **generateBlockers()**: N blockers (humans) will be generated with their own coordinates (x, y).

```
void generateBlockers(int number_camera, Blocker* blockers,
         unsigned count, int size_x, int size_y, gsl_rng* r){
    /* count = amount of blockers that are generated.
    This value changes every time this program is executed and is
    determined by PPP function. */
    int temp = 0;
    for (unsigned j=0; j<count; ++j){
    /*We get a random value in [0,1). That value is multiplied
    with the size of our area. We have to take in account that
    blocker with radius 50cm cannot be closer to wall than 50cm.
    */
         double randomx = gsl_rng_uniform(r)*size_x;
         double randomy = gsl_rng_uniform(r)*size_y;

         if (randomx < 50){randomx = 50;}
         else if (randomx > (size_x-50)){ randomx = (size_x-50);}

         if (randomy < 50){randomy = 50;}
         else if (randomy > (size_y-50)){randomy = (size_y-50);}
         if(temp<10){
         temp ++;
         }
         //Rounding and casting to integer.
         int x = (int)floor(randomx);
         int y = (int)floor(randomy);

         Blocker newBlocker = {x,y};

         //begin to check overlapping
         if (overlapping(blockers,newBlocker, j) ||
         overlappingWithDevices(cameras, newBlocker,
         number_camera)){
              --j;
              continue;
         } else {
              blockers[j] = newBlocker;
         }
    }
}
```

Program 1. *generate N blockers (humans) with their coordinates*

b. **generateDevices()**: M users participating in the monitoring process will be generated with their own coordinates (x, y).

```
void  generateDevices(int  *number_camera,  double  size_x,  double
    size_y){
    double camera_R, beta, alpha;
    printf("Enter camera info with (view length camera_R, angle of
           camera - beta): ");
    scanf("%lf %lf", &camera_R, &beta);

    //Input cameras (x,y)
    for (unsigned i = 0; i<*number_camera;i++){

          //Random cameras on the field (x,y)
          cameras[i].pos_x = (int)(gsl_rng_uniform(r)*size_x);
          cameras[i].pos_y = (int)(gsl_rng_uniform(r)*size_y);

          cameras[i].camera_R = camera_R;
          cameras[i].beta = beta;
          cameras[i].alpha = gsl_rng_uniform(r)*360;
          cameras[i].beta_rad = ((cameras[i].beta) * M_PI)/180;
          cameras[i].alpha_rad = ((cameras[i].alpha) * M_PI)/180;
          cameras[i].vector_bisector_x = cos(cameras[i].alpha_rad);
          cameras[i].vector_bisector_y = sin(cameras[i].alpha_rad);

          printf ("Camera: %d \t %d \t %lf %lf %lf %lf
          %lf\n",cameras[i].pos_x, cameras[i].pos_y,
          cameras[i].camera_R, cameras[i].beta, cameras[i].alpha,
    cameras[i].vector_bisector_x, cameras[i].vector_bisector_y);
    }
}
```

Program 2. *generate M users with their coordinates*

c. **operlapping()**: this function will check the distance between two blockers using Pythagoras's theorem and return whether new generated blocker overlaps with other blockers or not.

```
int overlapping(Blocker* blockers, Blocker test, unsigned array
    Size){
    int x1 = test.pos_x;
    int y1 = test.pos_y;

    for (unsigned g=0; g<arraySize; ++g){
          int x2 = blockers[g].pos_x;
```

```
        int y2 = blockers[g].pos_y;
        int delta_x = abs(x2-x1);
        int delta_y = abs(y2-y1);
        double distance = sqrt(pow(delta_x,2)+pow(delta_y,2));
        if (distance < 2*HUMAN_R){
            return 1;
        }
    }
    return 0;
}
```

Program 3. *check overlapping between two blockers*

d. **overlappingWithDevices()**: this function will check the overlapping of new generated blocker and the camera/microphones of participating users.

```
int overlappingWithDevices(Camera* cameras, Blocker test, int
number_camera){
    int x1 = test.pos_x;
    int y1 = test.pos_y;

    for (unsigned i = 0; i<number_camera; ++i){
        int x2 = cameras[i].pos_x;
        int y2 = cameras[i].pos_y;
        double distance = sqrt(pow(x2-x1,2) + pow(y2-y1,2));
        if (distance < 2*HUMAN_R){
            return 1;
        }
    }
    return 0;
}
```

Program 4. *check overlapping between generated blocker with participating users*

e. **checkIntersection()**: this function will check if a blocker hits to the LOS. The idea is to calculate the area of a circle (blocker), the length of LOS and thus determine the height of our triangle (the distance from the center of a circle to LOS). This algorithm use vector dot product to calculate the area of the triangle.

```
int checkIntersection(int number_camera, double ix, Blocker*
 blockers, unsigned blockerCount, double size_y){
    for (unsigned h=0; h<blockerCount; ++h){
        //True if target is covered by at least 1 camera
        int checkAllCamera = 1;
```

```
        //Check all blocker. If there is one blocker that blocks
        //the point, that point is not visible.
        for(int m=0;m<number_camera;m++){ //Check all of cameras
            int blocker_x = blockers[h].pos_x;
            int blocker_y = blockers[h].pos_y;

            //Area of the triangle
            double area = fabs( (ix-cameras[m].pos_x)*
                (blocker_y-cameras[m].pos_y) -
                (blocker_x-cameras[m].pos_x)*
                (size_y-cameras[m].pos_y) )/2;

            // -- condition 1 --
            //lengthAB from camera to visible object
            double lengthAB = sqrt( pow(ix-cameras[m].pos_x,2) +
                                pow(size_y-cameras[m].pos_y,2));
            //exceptional for codition 1- distance from camera's
            //central point to blocker's central point.
            double distance = sqrt( pow(cameras[m].pos_x -
            blocker_x,2) + pow(cameras[m].pos_y - blocker_y,2));
            //vectorAB from camera to visible object
            double vectorAB_x = ix - cameras[m].pos_x;
            double vectorAB_y = size_y - cameras[m].pos_y;

            //the angle gamma
            double cos_gamma = (cameras[m].vector_bisector_x *
                vectorAB_x + cameras[m].vector_bisector_y *
                vectorAB_y)/lengthAB;

            // -- condition 2 --
            //cos_gamma>cos_beta_half => gamma < beta_half
            double cos_beta_half = cos(cameras[m].beta_rad/2);

            // -- condition 3 --
            //the distance
            double height = 2*area/lengthAB;

            //check visible sight of camera to the object
            //break if it is not blocked
            if (((height > HUMAN_R || lengthAB<(distance-HUMAN_R))
                && lengthAB<=cameras[m].camera_R
                && cos_gamma >= cos_beta_half)
            ||lengthAB==0){
```

```
                checkAllCamera = 0;
                break;
            }
        }
        if(checkAllCamera){
            return 1;
        }
    }
    return 0;
}
```

Program 5. *Function for check intersection*

f. **calculateAverageArea()**: this function calculates the average rate of the target which is covered by a blocker. The result is relative (i.e. between [0,1]).

```
double calculateAverageArea(double start_x,
double end_x, unsigned* coverArray,
unsigned rounds){
int targetSize = end_x-start_x;
int sum = 0;
for (unsigned q=0; q<rounds; ++q){
sum += coverArray[q];
}

int divisor = rounds*targetSize;
double result = (double)sum/divisor;
return result;
}
```

Program 6. *Function to find the average rate of coverage by a blocker*

g. **writeToFile()**: writing all the results to files

4. NUMERICAL RESULTS

In this section, there are specific coverage metrics including CDF, mean and quantile of the coverage process. Then the wireless network rate requirements associated with certain coverage are demonstrated. The area of interest for all experiments is set to 100 x 100 meters.

4.1 Coverage Metrics

Coverage CDFs (cumulative distribution functions) for different number of participating users and different coverage radius of a single user are shown in figures with different coverage radius.

a. **CDFs for aural information (Figure 18a, 18b, 18c):**
   - Microphone radius, r = 1m (Figure 18a):

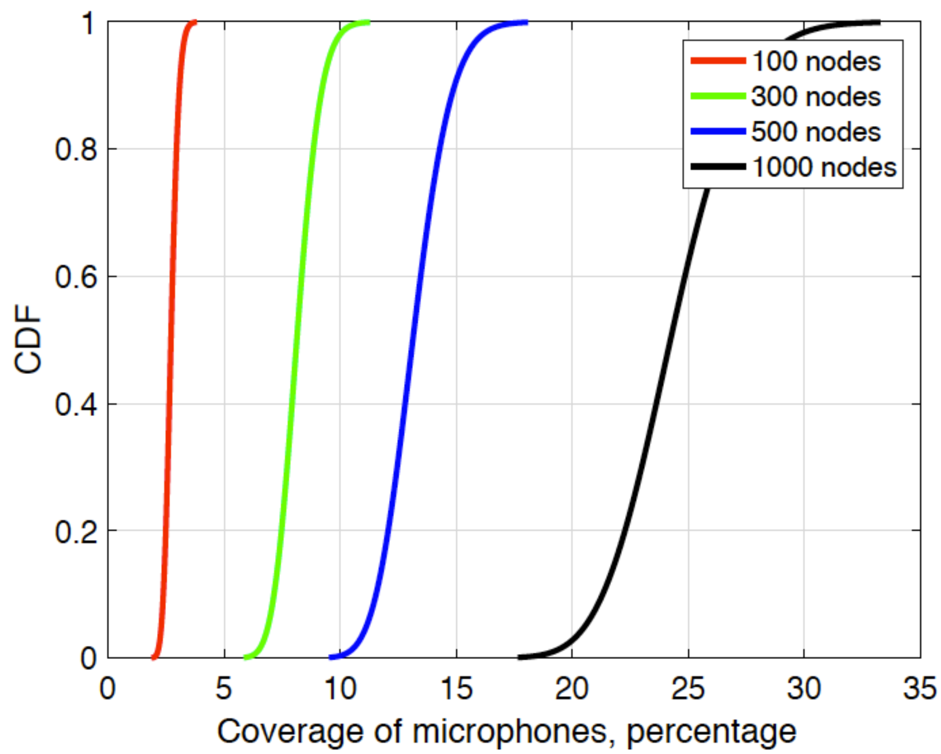

Figure 18. ***a.*** *CDFs of coverage for microphones, r = 1m*

- Microphone radius, r = 3m (Figure 18b):

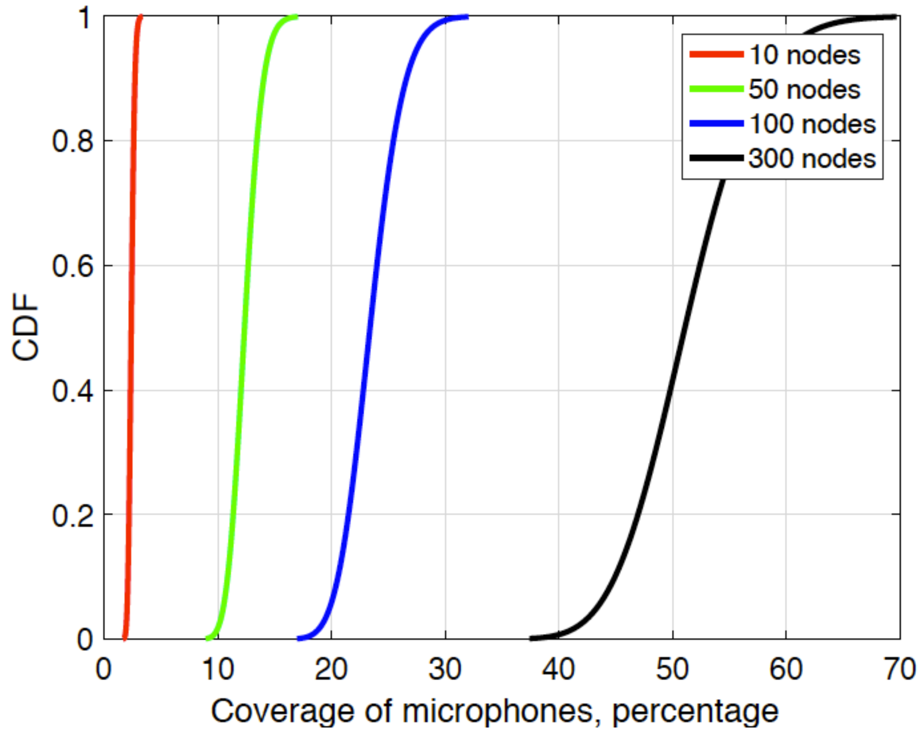

Figure 18. ***b.*** *CDFs of coverage for microphones, r = 3m*

- Microphone radius, r = 5m (Figure 18c):

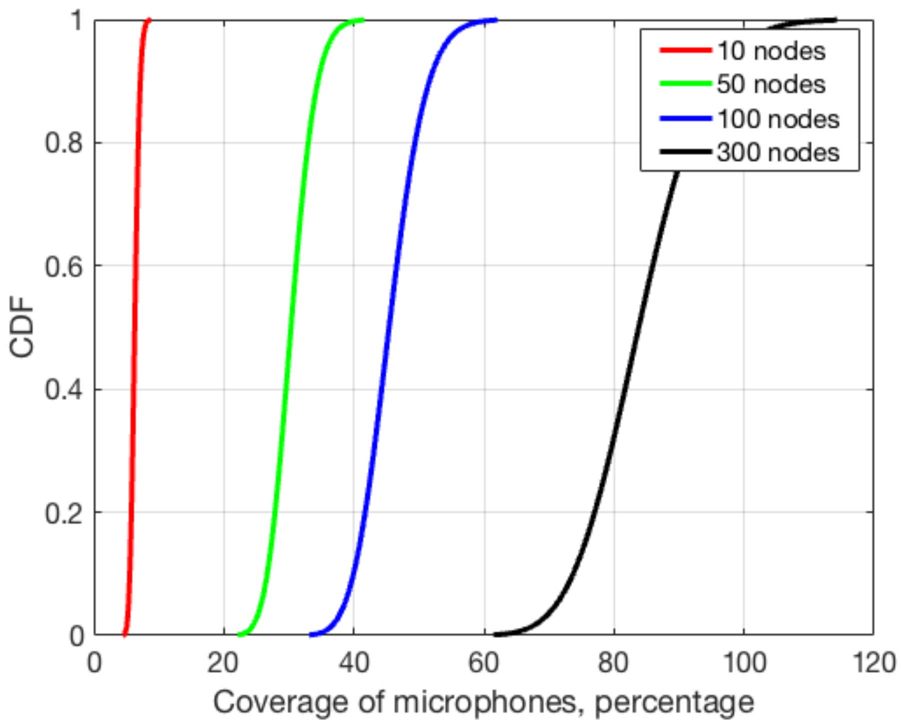

Figure 18. ***c.*** *CDFs of coverage for microphones, r = 5m*

b. CDFs for visual information (Figure 19a, 19b, 19c):

- Camera radius, r = 5m (Figure 19a):

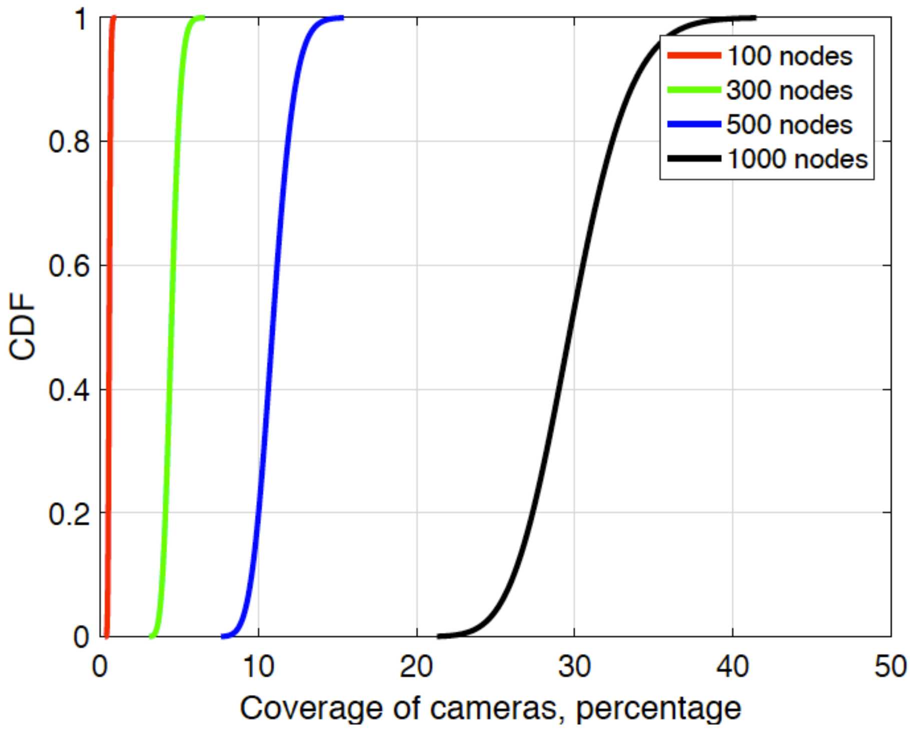

Figure 19. ***a.*** *CDFs of coverage for cameras, r = 5m*

- Camera radius, r = 15m (Figure 19b):

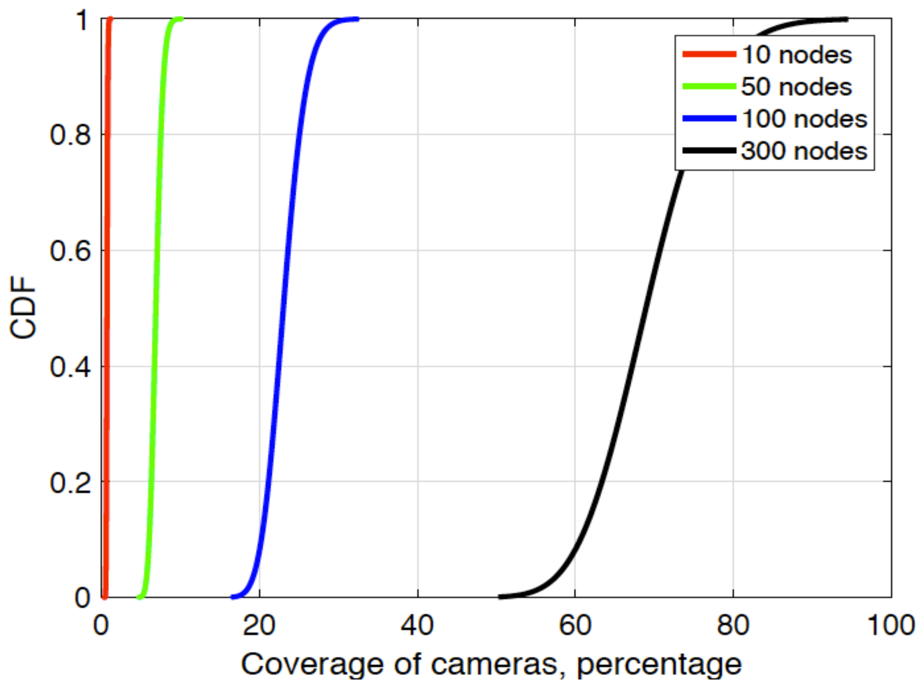

Figure 19. ***b.*** *CDFs of coverage for cameras, r = 15m*

- Camera radius, r = 25m (Figure 19c):

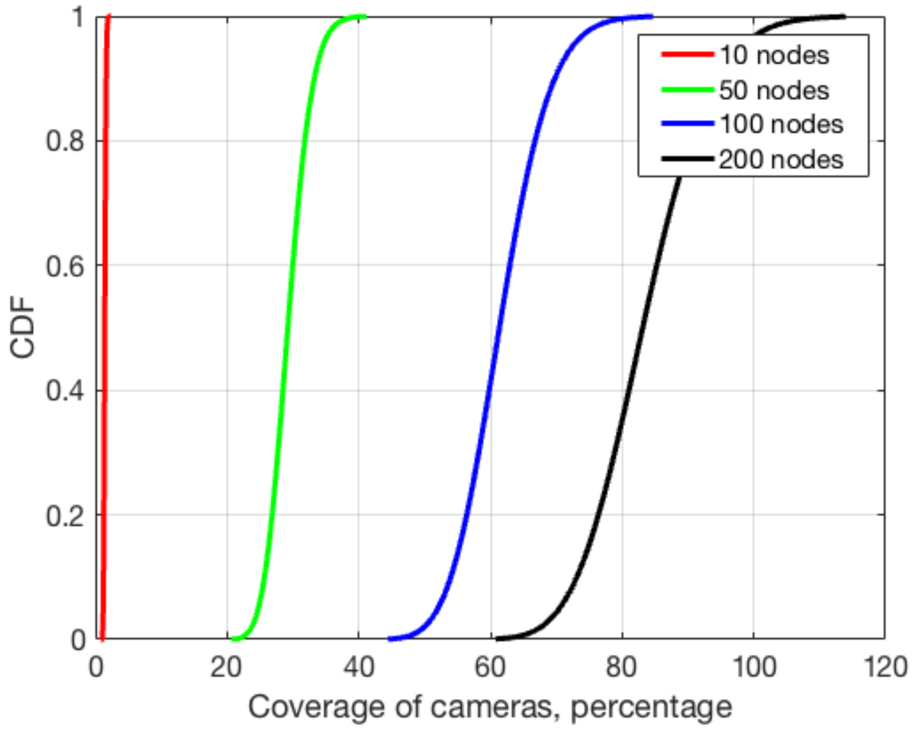

Figure 19. *c. CDFs of coverage for cameras, r = 25m*

The number of non-participating humans was kept constant and equal to 1000. Note that instead of the absolute values, we plot the percentage of the covered area in OX axis. Expectedly, for the same number of participating users, better coverage is provided for larger coverage radius of a single node. Furthermore, increasing the number of participating users provides better coverage. However, as one may observe, even for extremely large number of participating users (for example, 1000 nodes) full coverage is provided with negligible probability for aural information. Thus, to reliably cover 100 x 100 meters area, one needs to significantly more users than 1000, which might be problematic.

Cameras are characterized by significantly larger coverage radius. Thus, as one may observe already 300 participating users each with coverage radius of 15 meters provide non-negligible probability of 80% coverage of the considered area. It is also highlighted that the form of CDFs for both aural and visual information are highly-peaked (it can be seen, for example, visual information for 500 nodes) meaning that they should provide rather strict guaranteed of coverage in the considered random deployment scenario. Finally, it is emphasized that blocking of visibility field in visual information scenario does not qualitatively affect the form of CDFs compared to non-blocking aural information scenarios.

The mean values of the area coverage percentage as a function of the number of participating users and different coverage radius of a single user are shown group by microphones and cameras for aural and visual information.

The mean area coverage by microphones for aural information measurement shown in Figure 20, which is calculated, is based on the experiments.

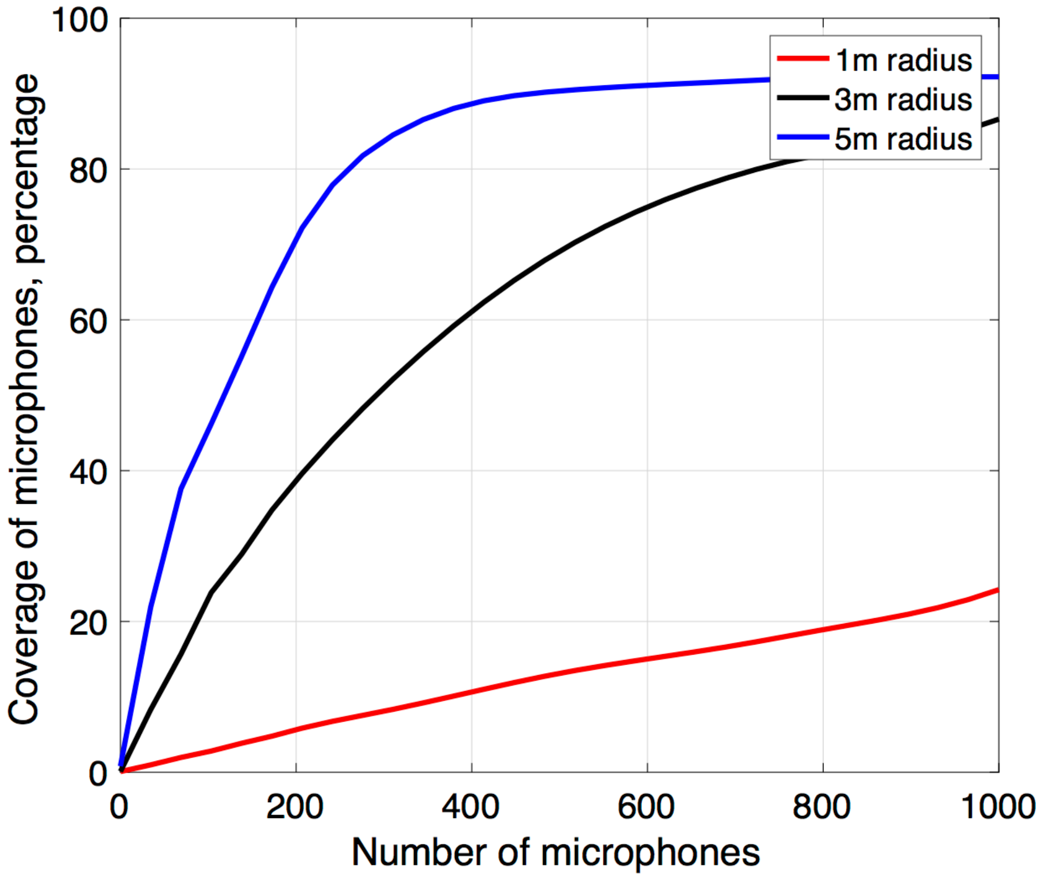

Figure 20. *Mean area coverage by microphones*

On the other hand, for visual information, mean area coverage by cameras is also shown in Figure 21.

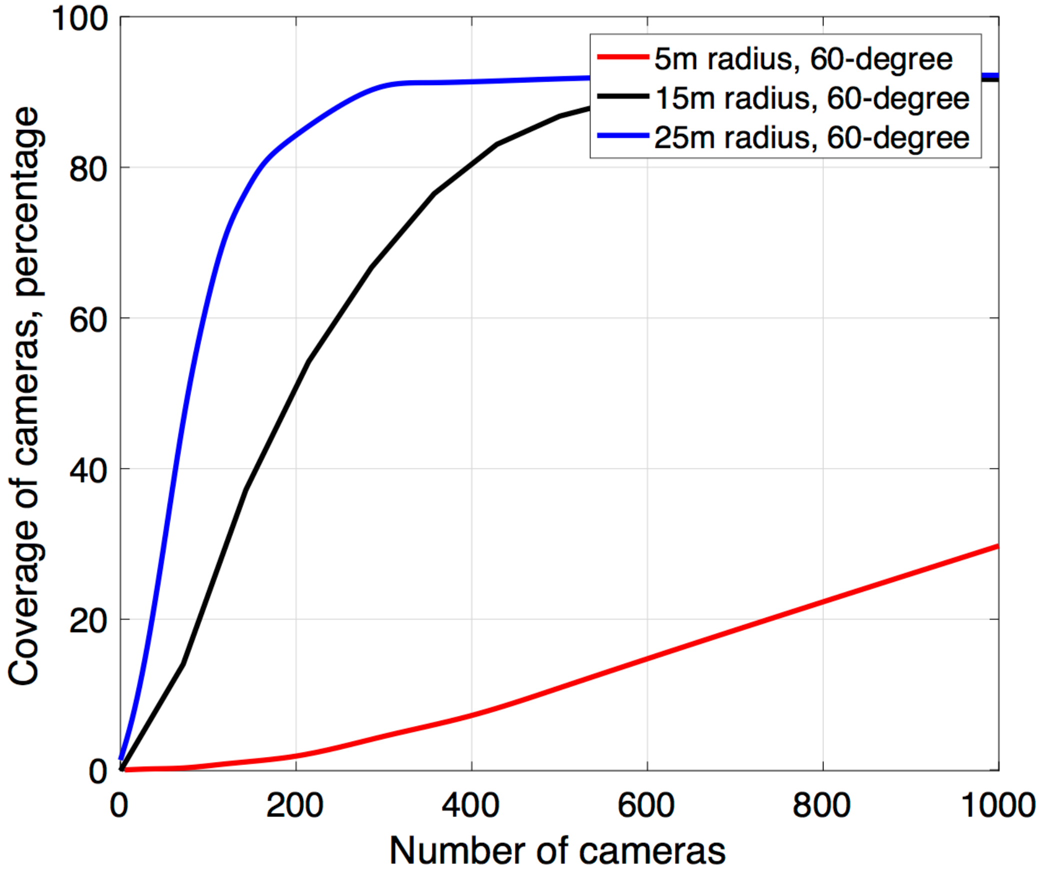

Figure 21. *Mean area coverage by cameras*

The number of non-participating users is set to 1000. One important behavior of this metric is that it does not approach 100% even for extremely high number of users and rather large coverage of a single user (for instance, 25 meters for visual information). This behavior is attributed to completely random choice of the participating users (uniform distribution area). Thus, to provide the mean coverage with close to 100% value is almost impossible for the proposed system and can only be achieved using either infrastructure nodes places in predefined places or drones/helicopters. Another option is to provide a wise choice of participating users selecting those that are located in favorable places.

4.2 Network requirements

It is now to consider the rate requirements imposed on the wireless networks by the proposed monitoring system. Note that depending on the quality of the codec, the coverage area may in generally vary for both aural and visual information. However, this effect is expected to be of minor importance and thus neglected here. Further, parameters such as resolution and compression rate may affect the performance and thus, smaller compression rates and higher resolution are generally preferable.

Audio and video codecs, which is used, have their parameters listed in Table 2, where MOS stands for mean opinion score [6,7,16].

Audio codecs				Video codecs			
Type	**MOS**	**Raw rate**	**LTE rate**	**Type**	**Resolution**	**Raw rate**	**LTE rate**
G.732.1	3.8	53kbps	6.626kbps	H.264	LD 360	0.7Mbps	0.875Mbps
G.726	3.85	32kbps	40kbps	H.264	SD 480p	1.2Mbps	1.5Mbps
G.711.1	4.1	64kbps	80kbps	H.264	HD 720p	2.5Mbps	3.125Mbps

Table 2.*Parameter of audio and video codecs*

The trade-off and dependencies of video and voice streaming are included bitrate requirements, power consumption and perceived quality.

About power consumption, there are transmission energy and encoding energy:

- Transmission energy: it can be seen that different voice/video codecs produce their output in a wide range of data rates. The different between them is the amount of power spent by amplifying the received signal in receiving states.
- Encoding energy: Depending the type of the codec, the energy consumed for encoding is also different. The actually energy depends on the type of signal processor used for encoding.

In addition, for the purpose of video streaming as the visual information recorded from monitoring process will transmit to the processing base control, the quality of video needs to be paid attention. For instance, H.264/AVC Video Coding Standard has been enhanced compression performance and provision of a network-friendly video representation addressing conversational (e.g. video telephony) and non-conversational (storage, broadcast, or streaming) applications, especially for mobile phone applications.

The network requirements in term of the bitrate needed from the network as well as 0.7 and 0.9 quantiles of the coverage process are plotted in figures (Figure 22a, 22b, 23a, 23b) as a function of the number of participating users for different types of codecs and different coverage radius of a single node.

a. Microphones, aural information – coverage quantiles and network requirements:

- Microphones, 0.7 quantiles:

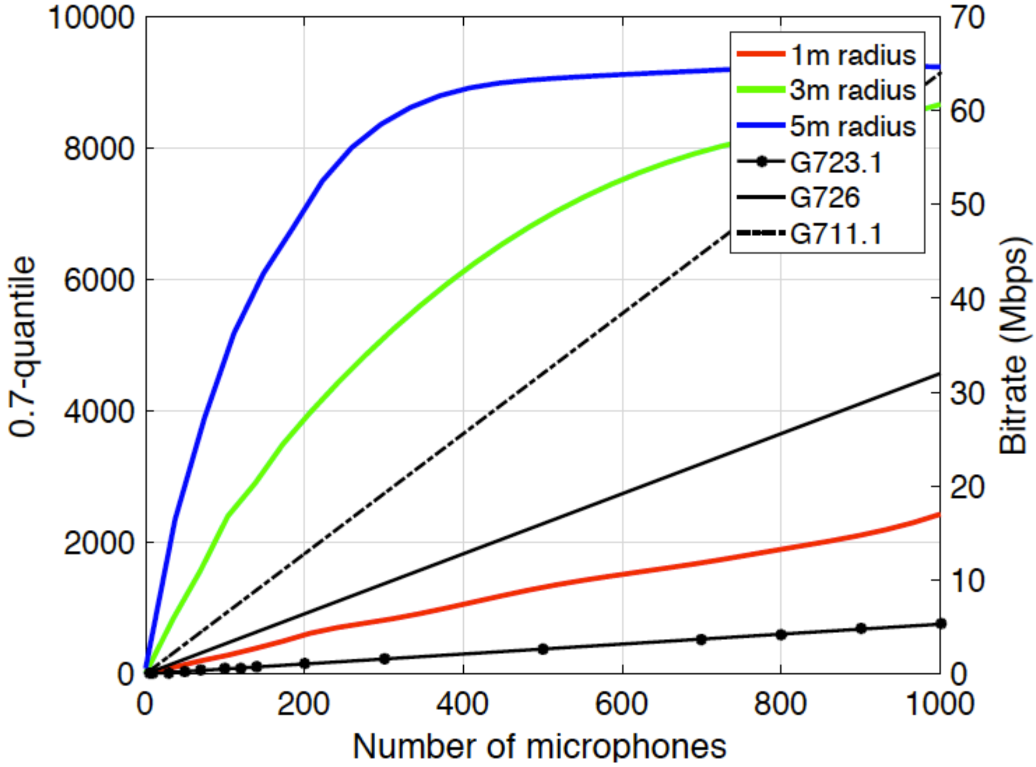

Figure 22. ***a.*** *Microphones, 0.7-Q and network requirements*

- Microphones, 0.9 quantiles:

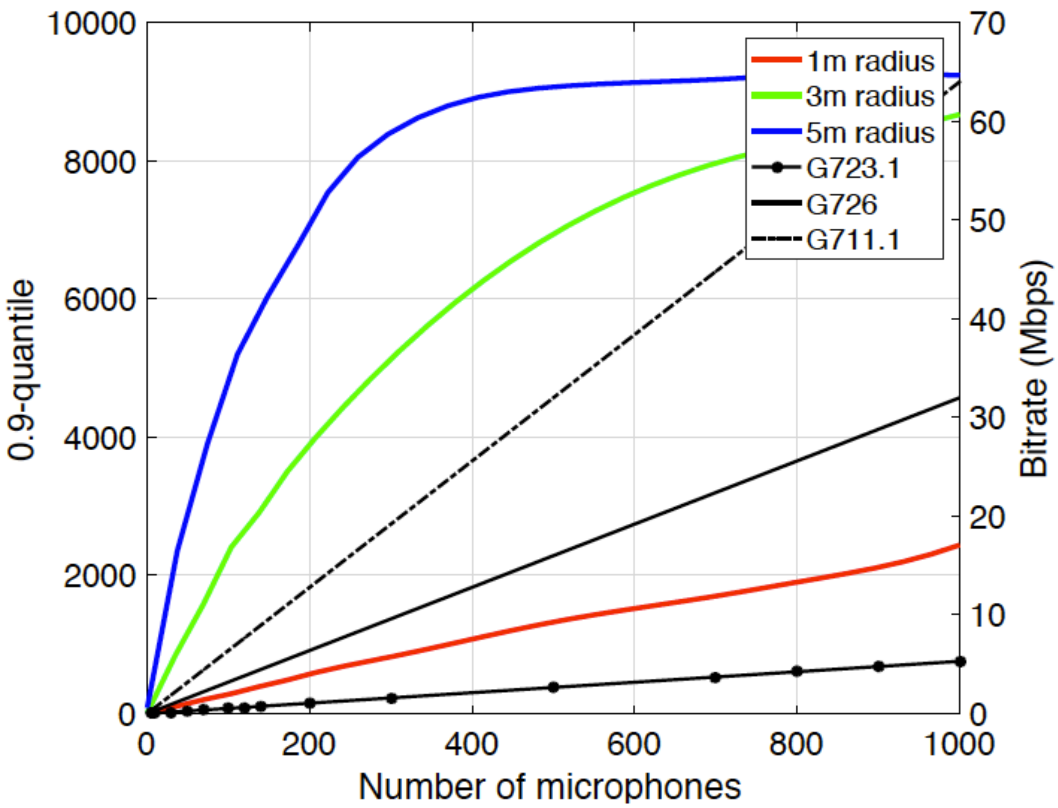

Figure 22. b. Microphones, 0.9-Q and network requirements

b. Cameras, visual information – coverage quantiles and network requirements:

- Camera, 0.7 quantiles:

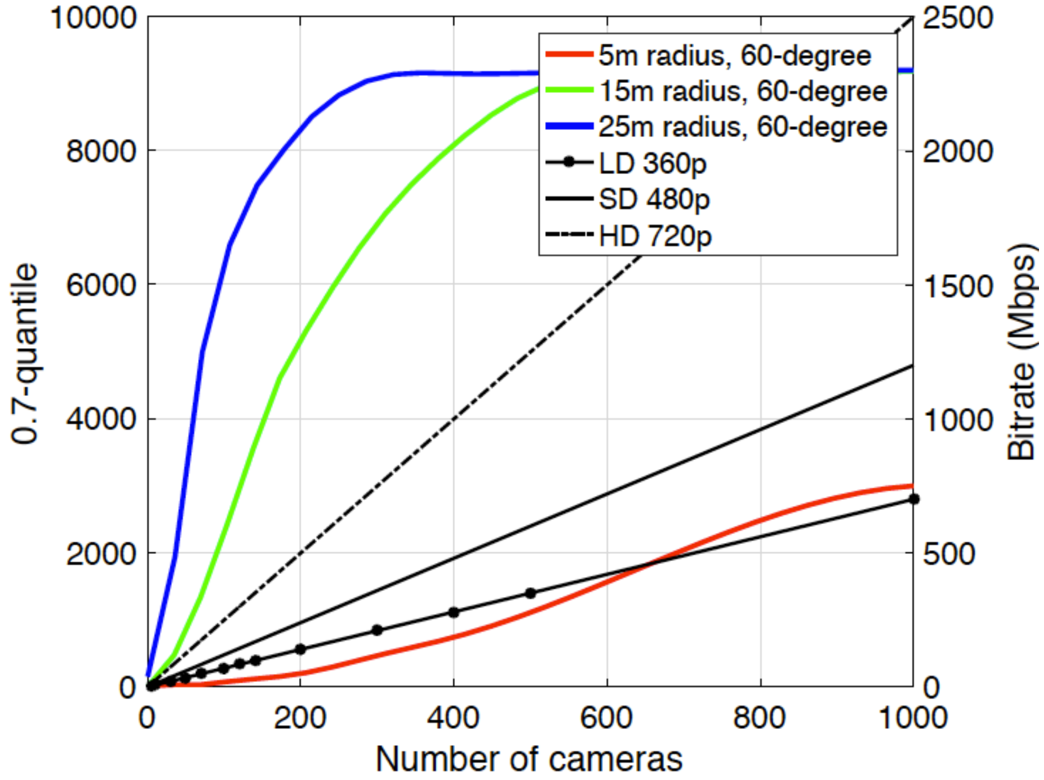

Figure 23. a. Cameras, 0.7-Q and network requirements

- Camera, 0.9 quantiles:

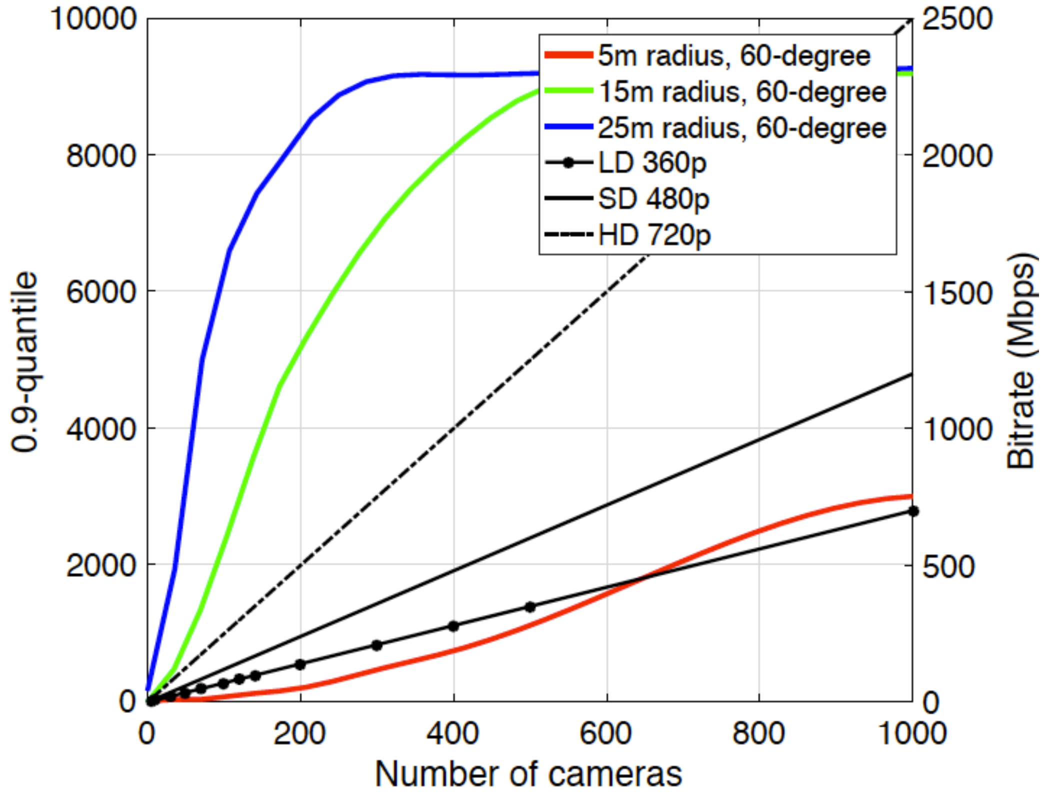

Figure 23. ***b.*** *Cameras, 0.9-Q and network requirements*

As one may observe, the network requirements for aural information approaches the value of 60Mbps for G.711.1 codec (raw rate 64Kbps). For G.723.1 type of a codec the aggregated rate from the nodes is just 5Mbps which can be easily handled by the modern LTE systems. Note that even 60Mbps can be supported by the LTE system.

Expectedly, the bitrates required by the video information are much higher. Even the lowest considered quality LD 360p requires the rate of 500Mbps to satisfy the 0.7-quantile of the area coverage. However, the millimeter wave systems operating at 28, 60 and 72GHz and offering the effective rate of up to 7Gbps are sufficient for visual monitoring even with HD 720p quality.

5. CONCLUSIONS

In this thesis, we proposed and analyzed the flash crowd monitoring system. The idea is to engage a subset of users in the crowd explicitly or implicitly to participate in the monitoring process. Assuming random positions of humans, the system for coverage metrics of interest for both aural and visual information is analyzed. Also, we compared these metrics to those of the optimal infrastructure-based monitoring system. Finally, we calculated wireless network requirements for the proposed system.

The numerical results show that the required density of the participating users needs to be exceptionally high to achieve "almost full" coverage, for instance, 0.9 quantile, for both audio and video sensors. Even though the associated network requirements are exceptionally high, they still can be supported by the forthcoming millimeter Wave systems offering substantial rate boost at the interface. After the work, we notice that the proposed system is the only viable option for detailed monitoring of in-crowd events for both aural and visual information. Taken together, we could say that the proposed system can be effectively used in conjunction with infrastructure-, helicopter-, or UAV-based monitoring systems in order to providing details information about the area of interest inside a flash crowd in either manual or unmanned manner.

BIBLIOGRAPHY

[1] Bai, T., Vaze, R., Heath Jr., R.W.: Analysis of blockage effects on urban cellular networks. IEEE Transactions on Wireless Communications pp. 5070-5083 (Sept 2014)

[2] Brualdi, R., Shanny, R.: A set intersection problem. Linear Algebra and its Applications 9, 143-147 (1947)

[3] Burkert, F., Fraundorfer, F.: UAV-based monitoring of pedestrian groups. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci. XL-1/W2, 67-72 (2013)

[4] Chiu, S., Stoyan, D., Kendall, W., Mecke, J.: Stochastic geometry and its applications. Wiley (2013)

[5] Gapenyenko, M., Samuylove, A., Gerasimenko, M., Moltchanov, D., Singh, S., Aryafar, E., Yeh, S., Himayat, N., Andreev, S., Koucheryavy, Y.: Analysis of human body blockage in millimeter-wave wireless communication systems. In: Proc. of IEEE ICC (May 2016)

[6] Radnosrati K., Moltchanov D., Koucheryavy, Y.: The choice of VOIP codec for mobile devices. Proc. ICN 2014 pp. 1-6 (2014)

[7] Radnosrati K., Moltchanov D., Kouceryavy, Y.: Trade-offs between compression, energy and quality of video streaming applications in wireless networks. Proc. IEEE ICC pp. 1000-1005 (2014)

[8] Kennedy, T.: Compact packings of the plane with two sizes of discs. Discrete and Computational Geometry 35(2), 255-267 (2006)

[9] Lazos, L., Poovendran, R.: Stochastic coverage in heterogeneous sensor networks. ACM Trans. Sensor Netw. 2(3), 325-358 (2006)

[10] Lee, Y., Park, Z., Bunkin, A., Nunes, R., Pershin, A., Voliak, K.: Helicopter-based lidar system for monitoring the upper ocean and terrain surface. Applied Optics 41(3), 401 (2002)

[11] Liu, X., Yang, B., Chen, G.: Barrier coverage in mobile camera sensor networks with grid-based deployment. Journal of Information and Engineering pp. 210-222 (2015)

[12] Pullman, H.: An elementary proof of pick’s theorem. School Science and Mathematics 79(1), 7-12 (1979)

[13] Singal, S.: Radio wave propagation and acoustic sounding. Atmospheric Research 20(2-4), 235-256 (1986)

[14] Teichmann, J., Ballani, F., Van Den Boogaart, K.: Generalizations of Matern hard core point processes, Spatial Statistics 3, 33-53 (2013)

[15] De Vries, P.: Area estimation with systematic dot grids. Sampling Theory for Forest Inventory, pp. 204-211 (1986)

[16] Wiegan, T., Sullivan, G., Bjontegaard, G., Luthra, A.: Overview if the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Technol. 13(7), 560-576 (2003)

[17] Behera, R., Kharade, P., Yerva, S., Dhane, P., Jain, A., Kutty, K.: Multi-Camera based surveillance system. Proc. IEEE WICT (2012)

[18] Zhou P., Long C.: Optimal coverage of camera networks using PSO algorithm. IEEE (2011)

[19] Enric P., Marc S., Juan L., Paolo R., Cristina B.: Helicopter-based wildfire monitoring system software architecture. IEEE (2010)

[20] Cai G., Chen B.M., Lee T.H.: Unmanned rotorcraft systems. Springer Science & Business Media, 2011.

APPENDIX A: FINDING MINIMUM NUMBER OF FIXED RADIUS CIRCLES TO FULLY COVER A RECTANGLE

Given a rectangle, X by Y, find the minimum amount of circles N with a fixed given radius R, necessary to fully cover every part of the rectangle

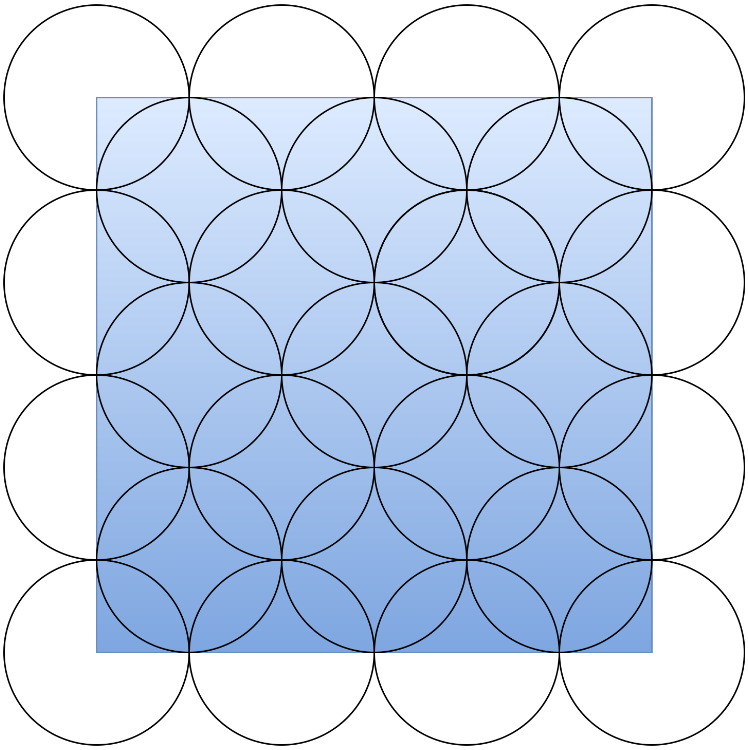

Figure 24. *Illustrate the problem of fully cover a rectangle with minimum number of fixed radius circles*

Denote that R is the radius of fixed circles, D is the diameter of the circles, X and Y are the two edge of the rectangle, N is the number of circles. We have:

D = 2*R; X >= 2*D ; Y >= 2*D

Thus, N = ceil(X/D) + ceil(Y/D) + 2*ceil(X/D)*ceil(Y/D)

In particular case if the remainder for X/D and Y/D equal to 0, then

N = (X + Y + X*Y/R) / D

For example:

Case 1: R = 1, X = 2, Y = 2 => N = 4

Case 2: R = 1, X = 4, Y = 6 => N = 17

Case 3: R = 1, X = 5, Y = 7 => N = 31