Mobile Networks and Applications
https://doi.org/10.1007/s11036-018-1031-1

Data Driven Cyber-Physical System for Landslide Detection

Zhi Liu[1] · Toshitaka Tsuda[1] · Hiroshi Watanabe[1] · Satoko Ryuo[2] · Nagateru Iwasawa[2]

© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Natural disaster is one of the most important research topics worldwide. In this paper, a data driven cyber-physical system is introduced to detect landslides. This system is composed of Wi-Sun acceleration sensors, which can detect the acceleration of the nearby environment in 3D domain, and the sensors are linked with the router (act as 'sink' node) via Wi-Sun transmission (i.e. IEEE802.15.4g). The details of the detection system are explained and the landslide detection mechanism with low computational complexity is proposed. A traffic reduction method is proposed thereafter to help reduce the data needed for transmission by exploring the intra-correlations of the sensor data. This method can save the energy consumption without degrading the detection performance. Field test is conducted and the results show that the landslide can be detected and amount of data to be transmitted can be reduced, which verifies the system's effectiveness.

Keywords Wireless sensor network · WSN · Landslide · Intra-correlation · Energy consumption · Disaster · Wi-Sun

1 Introduction

Disaster has become a hot research topic worldwide [1–3] and has drawn attention from the government, academia, industry, etc. The disaster leads to tremendous problems to the world, and Japan is one of the countries most affected by natural disasters, including earthquakes, tsunamis and landslides. Hence, how to detect the *unusual* events or the disaster has become an important and urgent research issue. There are a lot of related research on disaster from various aspects, such as the high-accuracy weather forecast, disaster detection and emergency responses after disaster [4].

Among these related research, wireless sensor network (WSN) [5–8] is widely studied for disaster detection. WSN can be regarded as part of the more advanced cyber-physical systems [9], where cyber-physical system connects the cyber world (information, intelligence and communication) and physical world via sensors and actuators. WSNs are also good options for detecting the unusual events in the scene of interest (Note that the place to be monitored is denoted as the scene of interest in this paper.) by analyzing the data samples collected from the networked sensor nodes, where the sensor nodes are deployed in the scene of interest.

Although WSNs have been widely studied in various fields, the following issues still exist: *"wireless sensor node is with limited battery life, hence the transmission capability of the sensor node is limited and this leads to a bandwidth-limited connection network. Meanwhile the battery life of sensor node is limited."* Transmitting the tremendous data samples generated by the sensors thus becomes a heavy burden for the WSNs. Meanwhile, how to rely on the sensed data to pre-detect the landslide (or detect the '*pre-landslide*', which denotes the status that the landslide is about to happen) can help reduce the negative influences of diasters greatly.

This paper studies the landslide detection system using one cyber-physical system. A data-driven landslide detection mechanism is proposed, which can help detect the happening of the landslide and save the energy consumption by reducing the data transmitted to a great extent. Note that the battery life consumption can not be obtained directly in

✉ Zhi Liu
liu@ieee.org

Toshitaka Tsuda
tsuda-toshitaka@aoni.waseda.jp

Hiroshi Watanabe
hiroshi.watanabe@waseda.jp

Satoko Ryuo
ryuo.satoko.56@rtri.or.jp

Nagateru Iwasawa
iwasawa.nagateru.81@rtri.or.jp

1 Waseda University, 3-4-1 Ohkubo, Shinju-ku, Tokyo, 169-8555, Japan

2 Railway Technical Research Institute, Tokyo 185-8540, Japan

Published online: 14 March 2018

the filed test but the energy consumption is directly correlated with the amount of data broadcasted. In this paper, the energy saving performance is measured by the percentage of data that are not transmitted. Specifically, the change of the monitored scene is firstly detected by the Wi-Sun acceleration sensor; and the detected data samples are transmitted to the router via Wi-Sun[1] (IEEE802.15.4g)[2] [10, 11] transmission. The router acts as the 'sink' node and the Wi-Sun transmission (IEEE802.15.4g) [10, 11] supports multi-hop transmission and is with low energy consumption. After collecting the sampled sensor data, the router transmits the sensor data to the cloud server via the public network; and the sensor data is stored in the cloud. The control center can access to the cloud in real-time and hence obtain the knowledge of the monitored scene by studying the sensor data.

To evaluate this performance of the proposed cyber-physical landslide detection system, a small scale field test was conducted. The landslide detection testing system was deployed in *Railway Technical Research Institute* (RTRI) for the purpose of measuring transmission capability of the Wi-Sun sensor module and detecting the landslide in a simulating real-life situation. In this system, there is a slope, placed with seven Wi-Sun acceleration sensors. During the experiment, the landslide was simulated. The Wi-Sun acceleration sensor is developed by *National Institute of Information and Communications Technology* (NICT), Japan, and each acceleration sensor detects the acceleration in the x-direction, y-direction and z-direction (3D domain) periodically. The sensors are connected with the router using Wi-Sun (IEEE802.15.4g). The sensors send back the data every fixed period of time. Then the control center can learn the status of the slope from the collected sensor data. The inherent problem is how to detect the landslide from the sensor data (especially the 'pre-landslide') and how to reduce the energy consumption to extend the life time of the system.

In this paper, the landslide detection mechanism with low computational complexity is proposed by using the weighted L^1-*norm*. The mechanism can inform the control center the occurrence of landslide and 'pre-landslide' based on two chosen notification parameters generated from each group of the captured data, i.e. *mean* and *local variance*. We also study the collected sensor data samples and calculate the correlations between the notification parameters in different directions (i.e. in x-direction, y-direction and z-direction), using *Pearson product-moment correlation coefficient* (PMCC), a standard way of calculating the correlations between parameters. We find a near linear intra-correlation inside these data. And then we propose a scheme to reduce the redundant data traffic while maintaining the detection ability[3], and we name this the data-driven traffic reduction scheme. The field test is conducted and the performances of the landslide detection mechanism with and without using the traffic reduction method are calculated. The contributions of this paper lie in the following aspects:

1. A new landslide detection mechanism with low computational complexity is proposed, which can help notify the landslide and 'pre-landslide'.
2. The intra-correlations inside the sensor data samples are studied and a data-driven traffic reduction mechanism is proposed, which can help reduce the transmitted data while maintaining the detection performance.
3. The real system is built; and the performances of the proposed scheme are evaluated in terms of detection performance and amount of data to be transmitted based on the real world data.

The rest of this paper is organized as follows: Section 2 introduces the related research in the field of the WSNs and traffic reduction. Section 3 introduces this cyber-physical system and details of the field test. Section 4 introduces the detection mechanism and traffic reduction scheme. The detection performances are introduced in Section 5. Section 6 concludes this paper and introduces the future work.

2 Related work

This section introduces the related research in the field of WSNs and sensor data reduction methods in literature. The differences between the existing research and this paper are explained at the end of this section.

2.1 Wireless sensor networks

Cyber-physical system connects the cyber world (information, intelligence and communication) and physical world via sensors and actuators. WSN can be regarded as part of the more advanced cyber-physical systems [9]. WSN [5–7, 12, 13] is a mature research topic, and WSNs are used in many scenarios [14–17]. They are good options for detecting the unusual events in the scene of interest by analyzing the data samples collected from the networked sensor nodes, where the sensor nodes are deployed in the scene of interest and the network transmission used to connect them can be ZigBee [16], Wi-Fi [17]. Local sensor data processing or decision making is also feasible if the sensor nodes have

[1] http://www.wi-sun.org/

[2] http://www.ieee802.org/15/pub/TG4g.html

[3] Please note that the method could also be applied to other WSNs, especially which uses the same sensor

certain level computation capabilities, and this could help reduce the network traffic.

The research of WSNs can be classified into two categories according to their major focuses. The first category conducts the research on how to build WSNs for specific purposes [18–20]; and the second category studies the key theoretical issues of the WSNs such as routing [13, 21–23], MAC layer design [24, 25], and energy efficiency [26, 27]. Note that the works in the second categories also have their application scenarios and some researches jointly consider the issues of the first and second categories. Regarding works in the first category, the sensor system can be quite different because of the different applied sensor nodes or transmission methods, hence each sensor system usually needs to be studied and optimized individually. The second category mainly targets one or several technical issues of the WSNs and solves the involved problems. The simulation is usually used as the tool to evaluate the performances. Specifically, [18] introduces the WSNs used for monitoring the environment and tracking the disaster. [19] introduces a cyber-physical system for girder hoisting monitoring based on smart phones and elaborates how the system is constructed. [20] proposes a framework of the monitoring system for teenagers, using wearable sensors and mobile phone for the acquisition of physical, behavioural and emotional attitude of them. This system is validated by secondary school students. [21] proposes visibility-graph-based geographic protocols for the shortest path routing in sensor networks with good performance. As a typical work in the second category, [23] introduces a hole plastic scheme, and this scheme can help relieve local minimum problem faced by geographic routing.

2.2 Sensor data reduction

The tremendous data samples have become a serious problem for WSNs, which are limited in terms of battery life and bandwidth, hence how to reduce the traffic becomes a critical issue. Usually the sensor states include transmitting, receiving, listening and sleeping. Different state means different energy consumption, and energy consumption during data transmission is the highest. Meanwhile, the sensors used in the proposed disaster detection system (and many other systems) have fixed frequency of detection, transmission, etc. Hence by sending less data, the needed transmitting time can be shortened and the energy consumption can be reduced. The network transmission can also benefit from the less data transmission requirements. This issue is studied from different aspects [28–32]. The popular applied methods include the in-network processing and unnecessary data samples filtering. In-network processing is a kind of edge computing, and the data is processed at the node before forwarding to the next hop. Filtering out the unnecessary data samples before forwarding also saves the network bandwidth. These works involve drawbacks that the control center is not able to decide whether the system is working or not, since it may not receive information even when it is working. Moreover, different WSNs may utilize different types of sensors and transmission technologies, which may make the existing reduction method proposed in one scenario be unapplicable in another.

Unlike these works, the present paper investigates the intra-correlation of the generated acceleration data of the sensors in the x-direction, y-direction and z-direction. Given that the detection parameters used in this paper are highly (almost linear) correlated, we propose a novel scheme to reduce the redundant transmission without degrading the detection performance. Moreover, we propose a new detection mechanism with low computational complexity to help detect the landslide in advance of occurrence (i.e. 'pre-landslide').

[32] studies the train traffic monitoring system, and reduces the data traffic by investigating the correlation between the sampled data. The present research is an extension of [32] by considering the landslide detection method in a totally different scenario; and the objectives, detection mechanism are different from the previous paper. Compared with [33], who also studies such system and proposes a detection method for the landslide detection, a novel traffic reduction method is proposed in the present paper. A low computational complexity detection mechanism considering the reduced traffic is also proposed. These distinguish this paper from the researches in literature.

3 System overview

This section introduces the landslide detection system, Wi-Sun acceleration sensors and shows the small scale field test.

3.1 System overview

The framework of the landslide detection cyber-physical system is shown in Fig. 1. The scene of interest is monitored by Wi-Sun acceleration sensors, which detect the acceleration of the nearby environment in the x-direction, y-direction and z-direction every fixed period of time. One typical Wi-Sun acceleration sensor in the market is shown in Fig. 2. Wi-Sun (IEEE802.15.4g) transmission is used to connect the Wi-Sun acceleration sensors and the sensor data is forwarded to the router. The router acts as the sink node and is connected to the cloud via the public network. The sensor data samples are stored in the cloud server. The control center or operation center can access the data samples stored in the cloud and hence obtain the knowledge

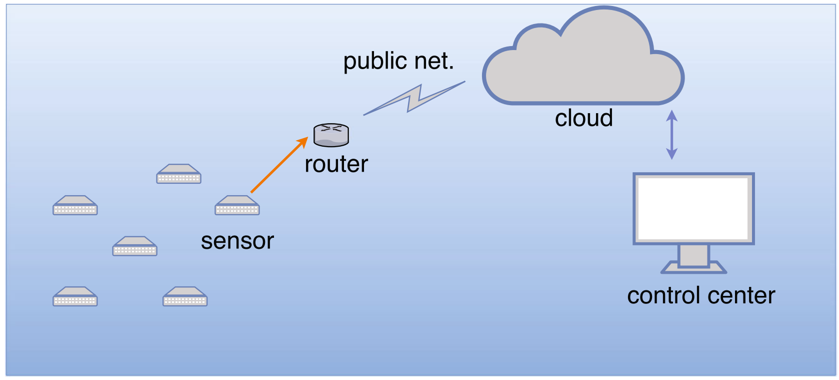


Fig. 1 Framework of the cyber-physical system for landslide monitoring

of the status of the monitored scene. Then the decision of whether there is a landslide happening or not is made by the control center according to the pre-designed detection mechanism.

The adopted Wi-Sun transmission (IEEE802.15.4g) supports multi-hop transmission and is with low energy consumption. IEEE802.15.4g is to create a Physical layer (PHY) amendment to 802.15.4 with the goal to provide a global standard that facilitates very large scale process control applications. This amendment defines an Orthogonal Frequency Division Multiplexing (OFDM) PHY operating in the license-exempt bands below 1 GHz, e.g., 902-928 MHz (USA), and enhancements to the IEEE 802.11 MAC to support this PHY, and provides mechanisms that enable coexistence with other systems in the bands including IEEE 802.15.4 and IEEE P802.15.4g. These applications could be smart-grid network capable of supporting large, geographically diverse networks with minimal infrastructure, with potentially millions of fixed endpoint. This Wi-Sun transmission uses the 920MHz ISM band in Japan, 915-928 MHz in Brazil and 902-928 MHz in US and many other regions.

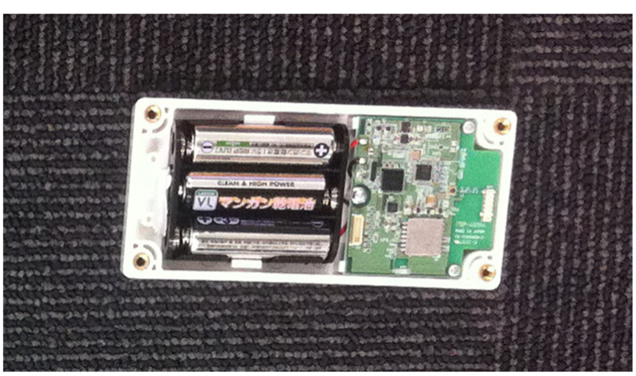


Fig. 2 Illustration of the Wi-Sun acceleration sensor, which can detect the acceleration values in the x-direction, y-direction and z-direction. The sensor also has the transmission module to broadcast the sensed data

3.2 Field test

To verify the performance of the proposed detection system, a field test is conducted and the framework is shown in Fig. 3 [33, 34]. This test system uses the framework as shown in Fig. 1. There is a slope in this real field test system, placed with seven sensors. Each sensor is marked with an ID number as shown in Fig. 3. The sampling period of the acceleration sensor is 0.05 second. Each sensor broadcasts 18 consecutive samples to the router every 2 minutes. The sensor does not broadcast all the data samples to save the critical battery life. Sensors are connected to the router using Wi-Sun transmission. The router connects the remote cloud via the public network as illustrated in Fig. 1. The control center can access to the cloud and hence monitor the scene as in reality by looking at the data in the cloud. The data samples record the detected acceleration values in the x-direction, y-direction and z-direction. Each data is with one time stamp. The data samples can be represented by (x_t, y_t, z_t, t), where x_t, y_t and z_t represent the acceleration values in the x-direction, y-direction and z-direction, respectively. t is the instance number. The sensors are not synchronized, which means each sensor has its independent broadcasting time. The landslide trail experiments are conducted and last for 272 minutes.

To simulate landslide caused by water level rising, water is poured from the bottom of the slope (or the embankment) until the water level reaches half of the slope, then the rainfall with the speed of 10mm/h is simulated. Note that this experiment is conducted on a small scale, but our proposed scheme also works for other sizes or scales, since hierarchical structure is used [35] as the wireless sensor network structure. There are also many other popular ways solving the scalability issue in literature can be used, hence the proposed scheme also works for large scale cases. Figure 4 shows the image of the monitored scene. The photo was taken every five minutes. This is the first photo of the landslide and we cannot find ”landslide” in the previous photo.

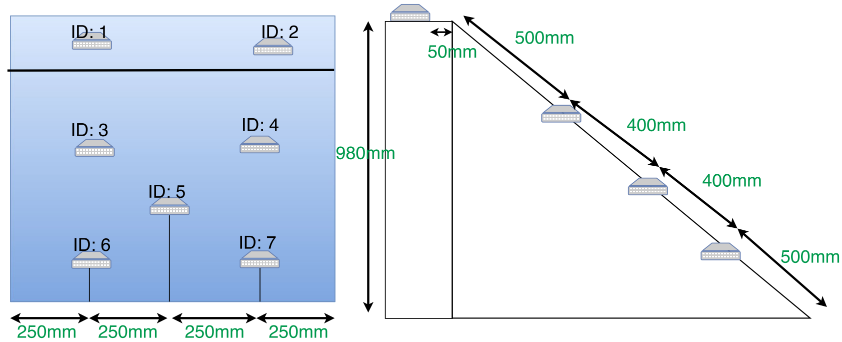


Fig. 3 Illustration of the sensor deployment in the field test system

The raw sensor data stored in the cloud is shown in Fig. 5, in which the data is captured by sensor 6 (6 is the ID of the corresponding sensor). The x-axis denotes the time instances (unit is second); and y-axis is the acceleration value. The unit of the acceleration is mg, which is $G/1000$ and G is the standard gravitational acceleration. The experiment is conducted from 11:57:00.00am, and we mark this time as *t=0, where t denotes the time index in the figures*. It can be observed from the sensor data that the acceleration is non-zero when everything is normal, i.e. each sensor has its own original point, and (0,0,0) is not set to be their original position, (0,0,0) means the acceleration value in x-direction, y-direction and z-direction are 0,0,0, respectively. The x-direction, y-direction and z-direction of each sensor is decided by the sensor's detection component. The acceleration value of every sensor keeps almost the same when there is no landslide (unusual) happening; but the captured acceleration value becomes larger when the landslide occurs. Note that there is no difference in the marker color in Fig. 5(a).

The collected data are then corrected to make sure that when there is no unusual event, (0,0,0) is the sensor acceleration value. We name this process 'correction'. Meanwhile, the collected data is noisy, which is unavoidable due to the characteristics of the sensors. However, from the data shown in Fig. 5, we can observe that the noises are bounded. Then a threshold-based de-noising scheme can be used to help remove these noises. The general idea of this threshold-based de-noising scheme is: firstly, threshold is set in x-direction, y-direction and z-direction, respectively, denoted as δ_x, δ_y and δ_z. Note that the major focus of this paper is how to reduce the data traffic while maintain the disaster detection performance, we directly use the proper threshold for the detection. The improper threshold selection can degrade the system performance and the threshold should be adjusted according to the conditions of the scene. Usually when the acceleration is larger than the upper bound of the noise (which can be observed from the captured sensor data), we can claim there is unusual happening, and the acceleration is within some range when everything is normal. We can then calculate bound of the mean and local variance during the normal cases and then set some threshold value for the system. As more data is collected, the threshold value can be further modified to make the system work better. Then if the absolute value of the data sample in the corresponding direction does not exceed the pre-defined threshold of the corresponding direction, we claim everything is normal and the data samples should be 0, if the data samples are not. In other words, if $abs(x_t) < \delta_x$, $x_t = 0$, where t is the time index. x can be substituted by y and z to conduct the de-noising in y-direction and z-direction, respectively.

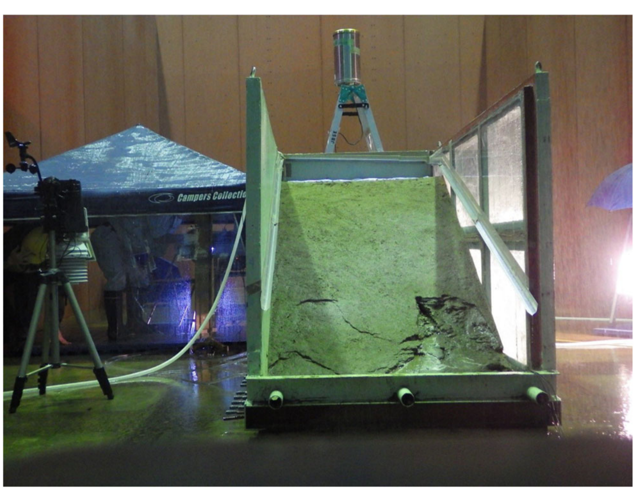
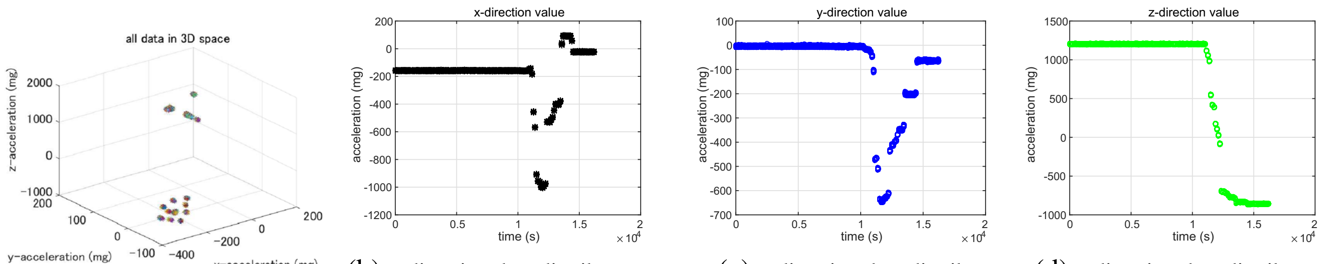


Fig. 4 Photo illustration of the happening of the landslide in the experiment

The sensor data after the correction and de-noising are shown in Fig. 6. From this figure, we can notice the scene is normal from the very beginning, but the acceleration value becomes significantly larger when the time index is 10920 (i.e. x-axis value t=10920). The large acceleration indicates the happening of the landslide and confirms the effectiveness of the proposed scheme. Note that some sensor node does not move or has very small acceleration value, which is smaller than the selected threshold after generating the significantly large acceleration value for a duration of time. The possible reason is that the acceleration sensors become statable again after the landslide or even during the



(a) 3D distribution. (b) x-direction data distribution. (c) y-direction data distribution. (d) z-direction data distribution.

Fig. 5 Illustration of the collected raw data generated by the sensor with ID=6 during the field test. The x-axis is the time in second, where the $x = 0$ stands for 11:00:00am

landslide. The duration of the process between the start of the landslide and stabilization of the sensor nodes is usually larger than the period between two consecutive broadcast transmissions (2 minutes in this experiment), therefore the control center can obtain enough data to capture the happening of the landslide if the network transmission does not fail. Similar to Fig. 5a, there is no difference in the marker color in Fig. 6a. Next we introduce how we perform the landslide detection and reduce the sensor data traffic by exploring the intra-correlations inside the data samples.

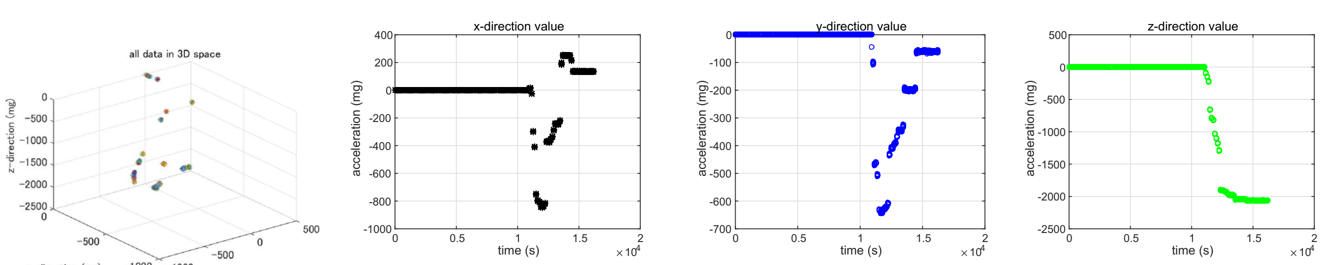
Figure 7 illustrates the sensor data during the landslide happening period. We can notice that the acceleration is larger than normal but not significant for a period of time before the disaster happens. Since the acceleration value is not significantly large, we can not call this disaster but this is different from the normal cases. We name this period to be *pre-landslide* period. This period is quite important since this may lead to the disaster, and we should pay much attention to the scene's changes during this period.

4 Landslide detection and traffic reduction

This section introduces the detection method and the traffic reduction mechanism by exploring the correlation inside the data samples.

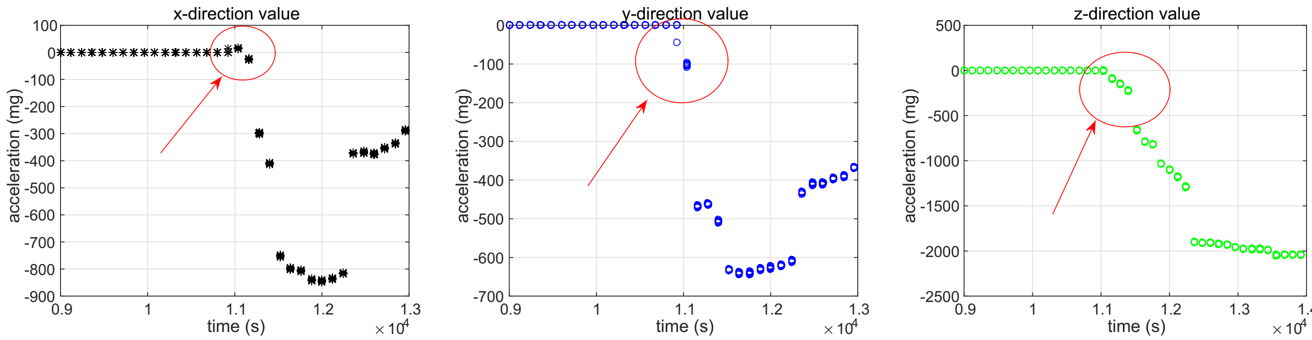
4.1 Unusual notification parameters

The captured acceleration value of sensor nodes is normally bounded. We call the event that leads to significantly different acceleration values from the normally generated acceleration data 'unusual' event, which is the landslide in this landslide detection system. N consecutive data samples are broadcasted by each sensor every couple of minutes (in the experiment, each sensor broadcasts 18 data samples every two minutes). During transmission window i, the data samples transmitted by sensor node j are represented by $d_{i,j,x}^{1}, ...d_{i,j,x}^{N}$. Here x means the data samples are the data in the x-direction. It can be replaced by y and z, so that the data samples in y and z direction can be denoted as well. *Mean* and *standard deviation* are usually used as the unusual notification parameters instead of using the raw data to perform the unusual detection. This is to avoid the false alarm caused by the transmission or sensing error, which can lead to high acceleration value at some time instances. Given the sensor node is very sensitive. By looking at less data, better understanding of what is happening can be known since the landslide may affect only a portion of all the data samples captured. Hence this system uses the *mean* and *local variance* as the unusual notification parameters, which can help omit the negative influence of the data noise and keep the high detection sensitivity.



(a) 3D distribution. (b) x-direction data distribution. (c) y-direction data distribution. (d) z-direction data distribution.

Fig. 6 Illustration of the data distribution of sensor with ID=6 after the correction and de-noising during the experiment. The x-axis is the time in second, where the $x = 0$ stands for 11:00:00am



(a) x-direction data distribution. (b) y-direction data distribution. (c) z-direction data distribution.

Fig. 7 Illustration of the data distribution of sensor with ID=6 during the landslide happening. The data shown is after the correction and de-noising during the experiment

mean $m_{i,j,x}$ of these N samples can be calculated as:

$$m_{i,j,x} = \frac{\sum_{k=1}^{k=N} d_{i,j,x}^{k}}{N} \tag{1}$$

and *local variance* $v_{i,j,x}$ can be calculated as follows:

$$v_{i,j,x} = \sqrt{\sum_{k=1}^{k=N} \sum_{r=max(1,k-\gamma)}^{r=min(N,k+\gamma)} \frac{|d_{i,j,x}^{r} - d_{i,j,x}^{k}|^2}{Nn_r}} \tag{2}$$

Here, γ defines how many data samples are considered in the calculation of the *local variance*, and in the experiment we set $\gamma = 2$. n_r is related with r and it defines how many adjacent data samples are considered in the calculation. The number of samples within the range $[max(1, k - \gamma), min(N, k+\gamma)]$ are the value of n_r, where k is the index. The maximum value of n_r is $2\gamma + 1$.

These two equations are defined for the data samples in the x-direction, the y-direction and z-direction notification parameters $m_{i,j,z}$, $v_{i,j,y}$ and $m_{i,j,z}$, $v_{i,j,z}$ can be calculated similarly by replacing the x-direction acceleration value using the corresponding value in the y-direction and z-direction, repetitively. Note that the sampled data faces the risk of transmission failure for different reasons. The control center hence may not have all the N data samples. In this case, the control center uses what is delivered to the cloud to calculate the unusual notification parameters instead.

4.2 Graph-based landslide detection

Based on the calculated unusual notification parameters, the control center decides whether there is a landslide happening or not. The detection rules are defined as follows:

- If the collected sensor data in any direction satisfies the prerequisite that *mean* is larger than the pre-assigned threshold δ_m, and *local variance* is smaller than δ_v, an 'unusual' or 'landslide' is claimed to be happening.
- If the detected acceleration value after the correction and de-nosing is non-zero, but the 'unusual' is not claimed to be happening, we claim this is the 'pre-landslide'.
- Otherwise, control center keeps monitoring the scene of interest.

δ_m and δ_v are two predefined parameter for each sensor deployed and can be obtained by learning the history data. The threshold-based landslide detection method is with low computational complexity and straightforward, and the control center can then understand what is happening exactly in each monitored area as shown in Section 5. However, some false alarms may happen, e.g. some sensors may be triggered when an animal walks into the field or when other non-disaster event happens. Even so, these events tend to only affect one single sensor. To avoid this kind of false alarm[4], we propose a graph-based landslide management mechanism, which can help the control center make more proper decisions. In this management mechanism, a graph is constructed as follows: each sensor is a node in the graph, the sensor node near the train track (or other important facilities) is set with priority 2 and other nodes are with priority 1. If two sensor nodes are adjacent geographically, an edge is added to the graph connecting the two nodes. The edge weight is set to be 0 initially. If the vertices of an edge triggers the alarm, the weight of this edge becomes the sum of both vertices. Then we can notice this detection mechanism actually calculates the weighted L^1-*norm*, and the mechanism works as follows:

- If priority 2 sensor triggers an alarm, the control center is suggested to take immediate actions.
- If priority 1 sensor triggers an alarm, and the sum of the edge weight is larger than a threshold δ' for a

[4]Note that other type of sensors such as video cameras [36] could be used to help reduce the false alarm, this is left as the future work.

period of time T', the control center is suggested to take immediate actions.
– Otherwise, control center will wait and pay attention to what will happen next.

δ' and T' are two constant parameters and assigned by the control center. The principles behind this management mechanism are: when a sensor triggers an alarm, the alarm can be caused by a real landslide, or by passing animals, etc. If the sensor is near the train track or other important facility, we do not want to take any risk and immediate actions will be taken. Hence we mention train track is because this system will be used in the areas near train track. Actually, if the sensor is triggered by an animal, this should also be informed to the control center, since the animal may enter the train track and damage the traffic. If the sensors are not near the train track or other important facility, the landslide decision won't be made immediately, and the control center will check what will happen next since the wrong judgement at this time won't directly damage the railway. Note that if the sensors are triggered by animals passing by, the number of sensors triggered won't be many, and the false alarm of a specific sensor won't last for a long period of time (for example, 2 minutes). Hence, the sum of the weight and the duration time are counted. If both values exceed the preassigned threshold value, immediate actions will be taken; otherwise, the control center will wait and pay attention to what will happen next. Meanwhile, by knowing how many nodes have triggered the alarm, the control center can have better understanding of the situation in the monitored scene. Since each sensor transmits consecutive data samples every fixed period of time (2 minutes in the field test), T' is assigned to be this period plus the time of the delivered consecutive data samples' duration.

4.3 Intra-correlation calculation

The traffic reduction method proposed in this paper relies on the intra-correlations and now we study the intra-correlations between the collected sensor data. PMCC is the standard way of calculating the intra-correlations with the output of a PMCC coefficient. PMCC coefficient 1 and -1 means linear correlation; and coefficient 0 means there is no correlation. PMCC is calculated as follows:

$$\rho_{X,Y} = \frac{E[(X - m_X)(Y - m_Y)]}{\lambda_X \lambda_Y} \tag{3}$$

In this equation, X and Y are the two to-be-tested data set. E stands for the process of calculating the expectation or the mean. m_X is the mean of X and m_Y is the mean of Y. λ_X and λ_Y represent data set X and Y's standard deviation, respectively.

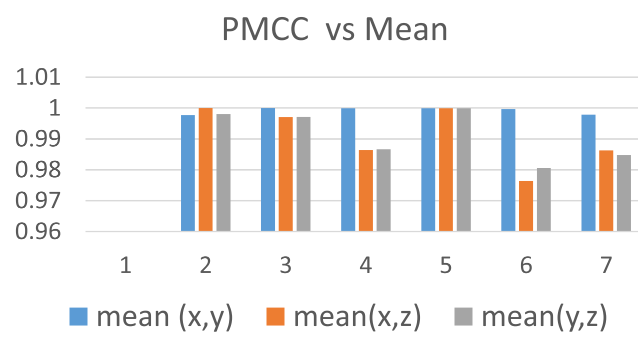


Fig. 8 Illustration of intra-correlation coefficients. mean (i, j) denotes the correlations between the mean of i-direction data and mean of the same sensor's j-direction data, where i,j can be X, Y or Z

We then test the intra-correlation inside the data samples collected by the field-test. Intra-correlation is calculated between the notification parameters of different directions. Since the data are counted in terms of the mean and local variance for the disaster detection, we show the PMCC of the mean, and PMCC of local variance in Figs. 8 and 9, respectively. For sensor 1, no intra-correlation is calculated because all the sampled data collected is 0, i.e. no unusual happens.

Figure 8 shows the intra-correlation coefficients of the mean value in different directions. From the figure, we can observe that the coefficient is close to 1, which indicates that near linear correlation exists among the data samples in different directions. In other words, x-direction parameters (mean) and y (or z)-direction parameters are highly correlated; y-direction parameters and z-direction notification parameters are also highly correlated. The x-direction, y-direction and z-direction data samples vary almost at the same time in terms of the mean value. Note that PMCC of different sensors may be different, the reason could be the different location, heterogenous nearby environment, and unsynchronized sensing directions (i.e. the coordination system of different sensors may be different).

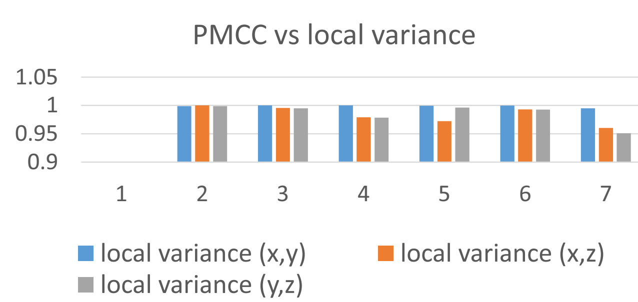


Fig. 9 Illustration of intra-correlation coefficients. Local variance (i, j) denotes the correlations between the local variance of i-direction data and local variance of the same sensor's j-direction data, where i, j can be X, Y or Z.

Figure 9 shows the intra-correlation coefficients of the local variance in different directions. From the figure, we can observe that the coefficient is also close to 1, which indicates the almost linear correlation among the data samples in different directions. In another word, x-direction parameters (local variance) and y (or z)-direction parameters are highly correlated, y-direction parameters and z-direction notification parameters are also highly correlated. The x-direction, y-direction and z-direction data samples vary almost simultaneously in terms of the local variance value. This is reasonable as the acceleration sensor captures the vibration of the nearby environment in the 3D domain, which results in the acceleration value changes in x-direction, y-direction and z-direction at the same time.

Overall, the high correlation inside the data samples leads to a possibility of reducing the data transmitted. How we design the traffic reduction mechanism is introduced in the following section.

4.4 Traffic reduction

From Section 4.3, we can observe that x-direction acceleration, y-direction acceleration and z-direction acceleration value changes significantly at the same time when the landslide happens, and this leads to the high correlations in terms of the unusual detection parameters *mean* and *local variance*. This means it becomes possible to detect the landslide using only one set of the values among the x-direction acceleration, y-direction acceleration and z-direction acceleration values, i.e. the sensors can only send one set of the acceleration values to the control center instead of sending the values of all the three directions over the bandwidth-limited network. In other words, the network traffic can be reduced without losing the detection performance[5].

Given the high intra-correlation of the notification parameters, we propose an algorithm for the traffic reduction with the detection performance guarantee. The general idea of this scheme is that if PMCC coefficient of the u-direction parameters and w-direction parameters (u, w could be x, y or z), is larger than τ, we can omit the data sampling of one direction. τ is pre-defined by the system and determines how much notification performance can be sacrificed, and $0 \leq \tau \leq 1$. The larger the τ is, the higher accuracy the system demands. We let $\rho_{m_{j,x},m_{j,y}}$ stand for the intra-correlation between sensor j's mean of x-direction data and mean of y-direction data. Similarly, we can define $\rho_{v_{j,x},v_{j,y}}$, which calculates the intra-correlation of redlocal variance of x-direction data and y-direction data. By calculating the intra-correlation of the notification parameters in different directions, we can then transmit part of all the data values only to the control center. The detailed method is illustrated in Alg. 1, in which u, w, s can be different sensing directions among the x, y, z direction.

Algorithm 1 Traffic reduction method illustration

```
calculate all the intra-correlation values
if ρ_{m_{j,u},m_{j,w}} ≥ τ, ρ_{v_{j,u},v_{j,w}} ≥ τ, and ρ_{m_{j,u},m_{j,s}} ≥ τ,
ρ_{v_{j,u},v_{j,s}} ≥ τ then                ▷ intra-correlation is ≥ τ
    sensor broadcasts u direction value alone.
else
    if ρ_{m_{j,u},m_{j,w}} ≥ τ, ρ_{v_{j,u},v_{j,w}} ≥ τ, and ρ_{m_{j,u},m_{j,s}} < τ,
    ρ_{v_{j,u},v_{j,s}} < τ then
        sensor broadcasts u and s direction's value.
    else
        sensor broadcasts u, w, and s direction's value.
    end if
end if
```

Then we can calculate how much traffic could be reduced according to the Alg. 1 without degrading the detection performance. Note that this traffic reduction method is designed in an ideal condition with all the sensor data collected, including sensor data during the disaster. In Section 5, we show this ideal traffic reduction quantitatively and the corresponding detection performance. The cases without the knowledge of all the sensor data are investigated in the Section 5.3.

The data processing complexity is low and the decision scheme is simple as shown in the manuscript, hence we believe this can be done easily using the state-of-the-art computation tools or cloud computing platforms.

5 Performance evaluation

This section shows the detection results along with the detection performance using the data samples after the traffic reduction.

5.1 Landslide detection from raw data

We first detect the occurrence of the landslide using the raw data and the aforementioned landslide detection mechanism: *If the mean, and local variance of the collected sensor data in any direction are larger than the pre-assigned threshold δ_m, δ_l, an 'unusual' or landslide is claimed to be happening.* The detection results are shown in Fig. 10, The result indicator defines whether there is disaster happening.

[5]Please note that sensors broadcast generated samples every two minutes, which is insufficient in emergency cases. Whether it is possible to use adaptive transmission rate, i.e. samples with adaptive total number to be delivered, is left as the future work. The most difficult part of the adaptive transmission frequency is the limitation of the hardware.

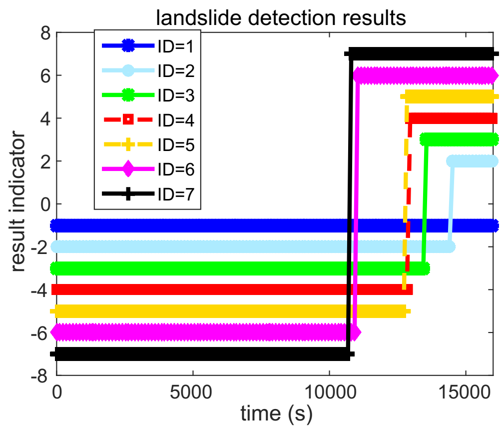


Fig. 10 Illustration of detection performance using the raw data according to the pre-defined rules

For sensor i, $+i$ means the detected result is 'there is a disaster' and $-i$ means everything is normal. From the results, we can observe that sensor functions well and the control center can obtain the knowledge of the scene in reality, which is also consistent with the occurrence of landslide. Figure 4 shows the first image of the landslide and we cannot find "landslide" in the previous photos, where the photo was taken every five minutes.

By comparing with Fig. 6, we notice that there is a period before the claimed 'unusual', but the collected data samples are non-zero, and this is the 'pre-landslide' as discussed. The red circle points out this time period. We also mark these in Fig. 11, in which we can trace the sequence of the occurrence of the landslide in the time domain. The number in red color denotes the sequence of the landslide's happening.

According to the designed management mechanism, the control center will look at the scene immediately if the sensor 1 is close to the track. Otherwise, it will keep quiet until the next trigger. The results of the weight are shown in Table 1. This table shows the increase of the weight as the sensor X detects the landslide, i.e. $ID = X$ means sensor X detects the landslide, and the corresponding weight is shown below. The weight is for the control center to make further decisions.

Table 1 Weight vs the ID

ID	ID=1	ID=2	ID=3	ID=4	ID=5	ID=6
Sum of Weight	3	6	10	15	20	23

5.2 Landslide detection with ideal traffic reduction

This section shows the detection results with the data samples after the traffic reduction according to Alg. 1, and the results are shown in Fig. 12, in which the τ is set to be 0.99. Note that only partially received data (after the data reduction) is used in the disaster detection instead of all the received data.

From the detection results, we can observe the traffic reduction does not affect the detection performances at all. The results are the same as what are shown in Fig. 10, but the data traffic is reduced significantly. The data saved in terms of percentage is shown in Fig. 13.

From Fig. 13, we can observe that the traffic can be reduced greatly without affecting the detection performance. Note that smaller τ may lead to larger possibility of reducing the sensor data traffic. Since the disaster detection is a serious issue, we take a relatively larger τ in the detection. However, if the τ=0.95, we can observe from Fig. 13 that more traffic can be reduced for sensor 3, 4, 5, 6 and 7. By considering less unusual detection parameters, the traffic can be further reduced, which may degrade the detection performance to some extent.

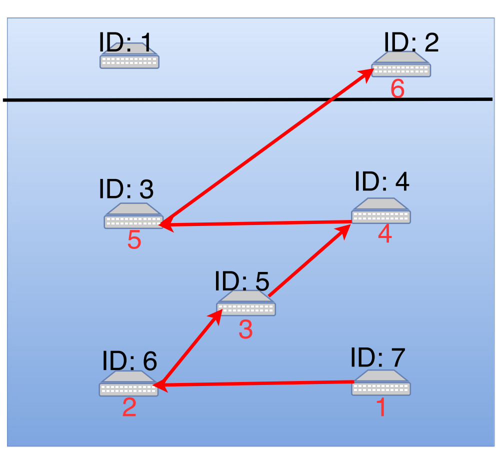


Fig. 11 Illustration of sequence of the occurrence of the landslide detected using the raw data according to the pre-defined rules

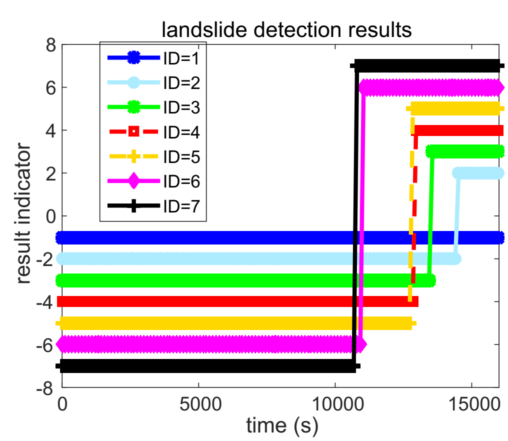


Fig. 12 Illustration of detection performance using the data samples after the traffic reduction

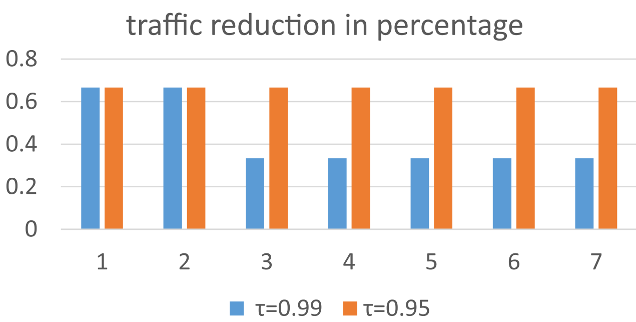


Fig. 13 Illustration of the traffic reduction in terms of data saving in percentage

5.3 Landslide detection with reduced data

Section 5.2 shows the detection performance of the landslide with ideal traffic reduction. In most scenarios, it is hard to obtain the sensor data for calculating the correlation in advance especially the sensor data during the disaster. This section shows the detection performance with naive selection of the sensed data in different direction and shows the corresponding detection performance, this does not need to obtain the sensor data in advance.

Table 2 shows the performance with all the possible options of the traffic reduction. *XYZ* means the acceleration values of the X-direction, Y-direction and Z-direction are used in the detection, and *X* means only the acceleration values of the X-direction are used in the detection. Other notations are defined similarly. The values under *XYZ* stand for the time when the disaster is detected when *XYZ* is used. The data reduction may delay the disaster detection, $+t$ (unit is second) denotes the detection delay when only the corresponding partial data is sent.

From Table 2, we can observe that by sending less data, the disaster notification may be delayed. However, if the sensor is not at an emergent position, this delay is tolerable. Moreover, these time period is marked as the *'pre-landslide'* already, which helps release the negative influence of the delay. The benefit is the battery life extension given the amount of sensor data transmitted is greatly reduced. Meanwhile, by sending combinations of the data such as Y-direction and Z-direction, the same detection performance can be obtained. These show the effectiveness of the traffic reduction. And this is reasonable because the acceleration sensor captures the vibration of the nearby environment in the 3D domain, which causes the acceleration values' simultaneous change in x-direction, y-direction and z-direction.

Table 2 Detection performance with different selected data

	X	Y	Z	XY	XZ	YZ	XYZ
ID=1	NA	NA	NA	NA	NA	NA	NA
ID=2	+0	+0	+0	+0	+0	+0	14401.85
ID=3	+0	+120	+0	+0	+0	+0	13559.85
ID=4	+0	+240	+0	+0	+0	+0	12958.85
ID=5	+0	+120	+0	+0	+0	+0	12840.85
ID=6	+360	+0	+240	+0	+240	+0	10920.85
ID=7	+600	+0	+480	+480	+0	+0	10608.85

6 Conclusions

This paper introduces a real-time landslide monitoring system using the Wi-Sun acceleration sensors. The landslide can be noticed by the control center according to the designed detection mechanism from the sensor data. By investigating the intra-correlations inside the data samples generated by each sensor, a data-driven traffic deduction mechanism is proposed by exploring the sensor data's intra-correlations, which can reduce the data volume significantly and guarantee the detection performance. A field test is conducted and the detection performance shows the effectiveness of the proposed system and traffic reduction method.

In the future, we plan to adopt this method in the large scale landslide detection system and using compressive sensing, etc. to further reduce the data transmission to save energy consumption.

Acknowledgments The research results have been achieved by "Research and Development on Fundamental and Utilization Technologies for Social Big Data," the Commissioned Research of National Institute of Information and Communications Technology (NICT), Japan.

References

1. Saha S, Matsumoto M (2007) A framework for disaster management system and wsn protocol for rescue operation. In: TENCON 2007-2007 IEEE Region 10 Conference. IEEE, pp 1–4
2. Chen D, Liu Z, Wang L, Dou M, Chen J, Li H (2013) Natural disaster monitoring with wireless sensor networks: a case study of data-intensive applications upon low-cost scalable systems. Mob Netw Appl 18(5):651–663
3. Tuna G, Mumcu TV, Gulez K, Gungor V, Erturk H (2012) Unmanned aerial vehicle-aided wireless sensor network deployment system for post-disaster monitoring. In: International Conference on Intelligent Computing. Springer, pp 298–305
4. Ali K, Nguyen HX, Vien Q-T, Shah P (2015) Disaster management communication networks: Challenges and architecture design. In: IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), 2015. IEEE, pp 537–542
5. Yick J, Mukherjee B, Ghosal D (2008) Wireless sensor network survey. Comput Netw 52(12):2292–2330

6. Al-Karaki JN, Kamal AE (2004) Routing techniques in wireless sensor networks: a survey. IEEE Wirel Commun 11(6):6–28
7. Iyengar SS, Brooks RR (2016) Distributed sensor networks: sensor networking and applications. CRC press, Boca Raton
8. Dong M, Ota K, Liu A, Guo M (2016) Joint optimization of lifetime and transport delay under reliability constraint wireless sensor networks. IEEE Trans Parallel Distrib Syst 27(1):225–236
9. Wu F-J, Kao Y-F, Tseng Y-C (2011) From wireless sensor networks towards cyber physical systems. Pervasive Mobile Comput 7(4):397–413
10. Chang K-H, Mason B (2012) The ieee 802.15. 4g standard for smart metering utility networks. In: 2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm). IEEE, pp 476–480
11. Obata K, Mizutani K, Harada H (2016) Carrier frequency offset estimation scheme for ieee 802.15. 4g based wide area wi-sun systems. In: 2016 IEEE International Conference on Communications (ICC). IEEE, pp 1–6
12. Fadel E, Gungor V, Nassef L, Akkari N, Maik MA, Almasri S, Akyildiz IF (2015) A survey on wireless sensor networks for smart grid. Comput Commun 71:22–33
13. Liu X-Y, Zhu Y, Kong L, Liu C, Gu Y, Vasilakos AV, Wu M-Y (2015) Cdc: Compressive data collection for wireless sensor networks. IEEE Trans Parallel Distrib Syst 26(8):2188–2197
14. Gungor V, Hancke GP (2009) Industrial wireless sensor networks: Challenges, design principles, and technical approaches. IEEE Trans Ind Electron 56(10):4258–4265
15. Pottie GJ, Kaiser WJ (2000) Wireless integrated network sensors. Commun ACM 43(5):51–58
16. Alliance Z et al (2006) Zigbee specification
17. Feng J, Liu Z, Ji Y (2014) Wireless channel loss analysis-a case study using wifi-direct. In: Wireless Communications and Mobile Computing Conference (IWCMC), 2014 International. IEEE, pp 244–249
18. Castillo-Effer M, Quintela DH, Moreno W, Jordan R, Westhoff W (2004) Wireless sensor networks for flash-flood alerting. In: 2004. Proceedings of the Fifth IEEE International Caracas Conference on Devices, Circuits and Systems, vol 1. IEEE, pp 142–146
19. Han R, Zhao X, Yu Y, Guan Q, Hu W, Li M (2016) A cyber-physical system for girder hoisting monitoring based on smartphones. Sensors 16(7):1048
20. Standoli CE, Guarneri MR, Perego P, Mazzola M, Mazzola A, Andreoni G A smart wearable sensors system for counter-fighting overweight in teenagers
21. Tan G, Bertier M, Kermarrec A-M (2009) Visibility-graph-based shortest-path geographic routing in sensor networks. In: INFOCOM 2009 IEEE. IEEE, pp 1719–1727
22. Hammoudeh M, Newman R (2015) Adaptive routing in wireless sensor networks: Qos optimisation for enhanced application performance. Inf Fusion 22:3–15
23. Yu F, Pan S, Hu G (2015) Hole plastic scheme for geographic routing in wireless sensor networks. In: 2015 IEEE International Conference on Communications (ICC). IEEE, pp 6444–6449
24. Suriyachai P, Roedig U, Scott A (2012) A survey of mac protocols for mission-critical applications in wireless sensor networks. IEEE Commun Surv Tutorials 14(2):240–264
25. Huang P, Xiao L, Soltani S, Mutka MW, Xi N (2013) The evolution of mac protocols in wireless sensor networks: A survey. IEEE Commun Surv Tutorials 15(1):101–120
26. Rault T, Bouabdallah A, Challal Y (2014) Energy efficiency in wireless sensor networks: A top-down survey. Comput Netw 67:104–122
27. Ren J, Zhang Y, Zhang N, Zhang D, Shen X (2016) Dynamic channel access to improve energy efficiency in cognitive radio sensor networks. IEEE Trans Wirel Commun 15(5):3143–3156
28. Santini S, Romer K (2006) An adaptive strategy for quality-based data reduction in wireless sensor networks. In: Proceedings of the 3rd international conference on networked sensing systems (INSS 2006), pp 29–36
29. Luo H, Luo J, Liu Y, Das SK (2006) Adaptive data fusion for energy efficient routing in wireless sensor networks. IEEE Trans Comput 55(10):1286–1299
30. Zhu Y, Song E, Zhou J, You Z (2005) Optimal dimensionality reduction of sensor data in multisensor estimation fusion. IEEE Trans Signal Process 53(5):1631–1639
31. Msechu EJ, Giannakis GB (2012) Sensor-centric data reduction for estimation with wsns via censoring and quantization. IEEE Trans Signal Process 60(1):400–414
32. Liu Z, Tsuda T, Watanabe H (2015) Traffic deduction exploring sensor data's intra-correlations in train track monitoring wsn. In: SENSORS IEEE, pp 1–4
33. Ryuo S, Hada A, Iwasawa N, Kawamura T, Nozue M, Kawasaki K The early detection method for embankment collapse using acceleration data, pp 517–522
34. Iwasawa N, Hada A, Ryuo S, Kawamura T, Nozue M, Kawasaki K (2015) The applicability for condition monitoring system in railway by using the Wi-SUN network (in Japanese), pp 1–6
35. Abbasi AA, Younis M (2007) A survey on clustering algorithms for wireless sensor networks. Comput Commun 30(14):2826–2841
36. Liu Z, Cheung G, Ji Y (2013) Optimizing distributed source coding for interactive multiview video streaming over lossy networks. IEEE Trans Circ Syst Video Technol 23(10):1781–1794