HADMÉRNÖK

XIII. Évfolyam 2. szám – 2018. június

A SURVEY ON DATA MANAGEMENT TECHNIQUES IN MILITARY WIRELESS SENSOR NETWORKS

KATONAI VEZETÉK NÉLKÜLI SZENZORHÁLÓZATOK ADATKEZELÉSI MEGOLDÁSAI

BOGNÁR Eszter Katalin

(ORCID: 0000-0002-3697-7871)

bognarek@uni-nke.hu

Abstract

Wireless sensor networks have huge potential in many areas of life. The emergence of the Internet of Things has made research into these networks of great importance. They also have proven applications in military and defense, with sensors being placed on soldiers' bodies and in battlefield monitoring systems.

The ultimate goal of sensor networks is the collection and storage of data, and to make that data available on demand. For this reason, it is critical to understand the various processes through which these functions are achieved. This article aims to elucidate and analyze current research into wireless sensor network data management techniques.

„Supported by the ÚNKP-17-3-I-NKE-20 New National Excellence Program of the Ministry of Human Capacities”

***Keywords*:** military sensor networks, Internet of Things, data management, data storage

Absztrakt

A vezeték nélküli szenzorhálózatok alkalmazása hatalmas potenciált jelent az élet számos területén, a dolgok internetének (Internet of Things) világában élve a témában folyó kutatások nagy jelentőséggel bírnak. A szenzorhálózatok a honvédelem számos területén is sikeresen alkalmazhatóak, különösen a harctéri érzékelő rendszerek, katonák testén elhelyezett szenzorok esetében. A szenzorhálózatok végső célja az érzékelt adatok összegyűjtése, tárolása, majd igény szerinti lekérdezése, ezért nélkülözhetetlen azon eljárások megismerése, melyek révén ezek a funkciók a lehető leghatékonyabban valósíthatók meg. A cikk célja feltárni a témában folyó kutatások aktuális eredményeit, elemezni, összehasonlítani az egyes megoldásokat.

***Kulcsszavak**: katonai szenzorhálózatok, dolgok internete, adatkezelés, adattárolás*

A kézirat benyújtásának dátuma (Date of the submission): 2018.03.05.
A kézirat elfogadásának dátuma (Date of the acceptance): 2018.06.06.

INTRODUCTION

The term wireless sensor network refers to hundreds or even thousands of linked sensors that can record changes in their physical environment. Acting in a coordinated fashion, these sensor networks are able to continuously monitor the conditions of a particular area or physical object. Recent technological developments make it possible to manufacture especially small and inexpensive sensors that can be installed virtually anywhere. Due to their low cost and practical versatility, these sensors have already become widespread in the civil sphere, with applications in healthcare, manufacturing, environmental monitoring, and so-called smart environments. They also have proven applications in military and defense, with sensors being placed on soldiers' bodies and in battlefield monitoring systems. The true potential of sensor networks in all areas of life is only beginning to be explored. The further development of the Internet of Things relies on research into sensor network technology.

However, there are major restrictions facing the optimal operation of sensor networks. These restrictions stem from limitations in sensor memory, storage, computing capacity, and useful life, as well as limitations on bandwidth. Therefore, efforts are being made to optimize the flow of data traffic, routing between large numbers of nodes, and the storage and retrieval of data. Any attempts at optimization must be carefully tailored according to scarcity of available resources. The conventional data management solutions used in conventional computer networks are somewhat inefficient, and because of this it is necessary to develop novel methods and procedures. Several studies relevant to sensor network optimization have been published domestically. The work of Nagy T. I. [1][2] focuses on routing strategies utilized in the network layer of sensor networks. Haig Zs. [3][4] and Kovács L.'s [5] research explores potential opportunities for use of sensor networks in electronic warfare. The ultimate goal of sensor networks is the collection and storage of data, and to make that data available on demand. Therefore, it is critical to understand the various processes through which these functions are achieved.

Currently, researchers are furthering data management techniques that are better optimized for large amounts of data resulting from sensor networks. Because the quantities of raw data are so large, useful data must be specified and sorted according to their intended application. To accomplish this aim, various data storage and query mechanisms are created. When developing these new mechanisms, the limited storage capacity of the nodes must be taken into consideration. Furthermore, there may be network communications and package delivery fees associated with these mechanisms, which become financially significant. The purpose of this article to suggest solutions for current data storage and query management issues by comparing the advantages and disadvantages of different methods.

DATA MANAGEMENT IN SENSOR NETWORKS

One of the main challenges in planning and utilizing sensor networks is the efficient storage and querying of sensor data, known generally as data management. Management of sensor network data must suit the specific needs of each application. There are three main questions regarding data management:

1. What are the main characteristics and limits of sensor networks?
2. Where can data from sensor networks be stored?
3. How can data storage systems respond to the needs of users, and how are the results of data storage communicated to users?

Since sensors have numerous limits in terms of storage capacity, communication capability, and useful life, the development of novel procedures become necessary. One such procedure leading to further optimization is the distributed processing and storage of data within the network. By taking into consideration a number of specific parameters, we can optimize the operation of the system and select the most effective method for data management. The expected useful life of the system (including short and long-term installation), the nodes' power, computing and storage capacity, and the data requirements of the particular application are all significant elements which must be considered. Furthermore, the nature of the user's queries (single, continuous, event-driven, etc.), the quantity of generated data during the operation, the network size, and potential preprocessing parallelization opportunities are also all relevant. Current research points towards the use of distributed data processing. There is a spatial and temporal correlation between data collected by the sensors, which has substantial potential to overcome many current limitations. The purpose of this chapter is to describe the characteristics of sensor networks and the limitations of data management.

The characteristics and limitations of wireless sensor networks

The optimal operation of sensor networks is limited by a number of factors. Most factors can be attributed to the characteristics of sensor network elements and the communication channel. Based on the work of Fan W. et al. [7] and Rawat P. et al. [8] in the case of sensor networks, the following limitations should be considered:

- **Limited memory, storage, power supply, and computing capacity:** For energy-intensive operations (e.g. maintaining routing tables, handling datasets, implementing data security algorithms) the use of cost-effective protocols are necessary in order to reduce the energy consumption of nodes, thus maximizing the useful life of the network.
- **Dynamic network topology, self-organization, peer-to-peer networks:** Sensor networks must support frequent changes in network structure and in quality of data links.
- **Unreliable wireless communication:** To achieve proper operation of sensor networks, the necessary data must be transmitted to the base station. In sensor networks frequent unexpected errors might occur e.g. problems with the connection, wrecked or discharged nodes. To overcome this, the network must be fault tolerant. This can easily be improved through redundancy.
- **A large number of often heterogeneous nodes:** As there can be up to many thousands of nodes in the monitored area, faults in system architecture and network protocols must be corrected. The network should be self-organizing and self-healing. Because giving unique addresses to thousands of nodes would be extremely expensive, data-centric approaches that focus on groups of nodes become significant.

Wireless sensor network data management features

Focusing on data management according to Zhao F. and Guibas J. L. [6] the following features and challenges arise:

- sensor networks dynamically change, hubs may go dead, and the connection between two nodes might be terminated. However, the data management system must hide it from the end-user; it must give the impression of being a robust and stable system;
- besides traditional static data tables continuously generated real-time data streams must be also handled;
- since data transfer is one of the most expensive operations, adequate solutions must be applied to minimize unnecessary communication between nodes of the network. Such solutions include data preprocessing and in-network data aggregation;

- during data transmission, longer delays may occur. It is not enough to optimize a query once, but it must be continuously monitored and revised in order to operate at the most efficient speed and resource utilization;
- since storage capacity of nodes is limited, the possibility of storing historical data must be considered.

THE STORAGE OF SENSOR DATA

The limitations of sensor networks have to be already considered at the physical level of data storage. To find the most efficient storage solution for sensor data, storage capacity and cost of data transmission must be taken into account. The purpose of this chapter is to present the most important storage strategies that are currently used in sensor networks.

External, centralized storage

Centralized data storage is the easiest way to store sensor data. This method is the well-known storage technology of traditional computer networks. Wireless sensor networks are constructed of data-collection facilities, passages linking internal and external components, and externally-located base stations.

The essence of centralized data storage is that all the collected data from sensor network components is immediately transmitted to the external base station. Subsequent storage and processing occurs at the base station, which is connected to the external network. This method has far greater storage capacity as well as computer processing power for processing data.

The advantage of the method is that data is stored in a centralized fashion, making data query simple and convenient. Communication costs apply only to the transmission of data. Therefore, there is no need for additional interaction with the sensor system in order to answer queries from users. The base station has a double function. On the one hand, it provides data storage functions: receive, store, and process data collected by the sensors. On the other hand, the base station also processes and answers queries from users through the sensor network nodes.

There is one disadvantage to this system. Since the base station carries significant external communication with both users and sensor nodes, a bottleneck may occur due to overloading and thus lead to delays. Additionally, because nodes located nearer to the base station have to execute a greater degree of communication, there is a possibility that those batteries will be more quickly discharged, reducing the useful life of the entire system. Data security must also be taken into account; the base station may be flooded by unwanted data requests or face attacks that threaten the integrity of the data. [20]

This solution is most optimal for sensor network of small size (approx. 100 sensors) and for situations in which the data collection is not constant, as the system only detects and transmits data in response to an event. [7][9][11]

In the case of military applications centralized storage can be used in force protection/base protection by instrumenting the perimeter of the base with networked wireless sensors. Since the monitored area is demarcated and relatively small, there isn’t much multihop data transfer between the sensor nodes and the gateway. Usually each node is within the range of the gateway or just few hops away, so it’s easier to transfer each data to the central storage point where it will be stored and processed. A possible set up can be seen on Figure 1 .

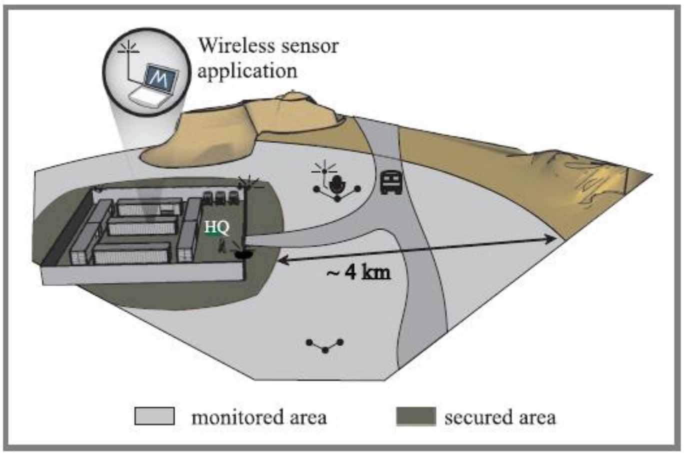

Figure 1 WSN in base protection [21]

Distributed Data Store

Centralized storage of data presented in the previous section can be used effectively only for small-scale sensor networks. [11] With networks consisting of thousands of nodes, centralized storage approaches are not practical. Because a large amount of data and communications for the base station are generated, neighboring nodes can become overloaded. Overloading leads to system delays and in the long-term a reduction in the system's useful life. For very large amounts of data, decentralized sensor-data approaches are used instead, being one of the most energy efficient and energy saving forms of data storage. The sensor data model considers the whole network as a distributed database where each sensor node is a basic unit of storage. Decentralization helps in saving energy; however, in contrast to other approaches, only local information is available at each node. The system's complex algorithms are designed for decentralization, thereby placing a limit on their performance. Overall, decentralized design leads to greater energy efficiency, but the complexity of the data management system increases significantly. [11] Regarding security, there is some danger in individual nodes becoming compromised. [20]

In military applications decentralized data storage are beneficial in the case of battlefield surveillance where thousands of small and cheap sensors are deployed on the battlefield to provide wide range of surveillance and to support the soldier with real time situational awareness and joint operability among forces. Typically, sensor nodes are deployed randomly (e.g., via aerial deployment), and are expected to self-organize to form a multi-hop network. Because of the large territorial coverage and the high density of sensors the communication between the nodes increasing significantly. To keep in mind that saving energy thus improving the lifetime of the sensor nodes is crucial, it is more beneficial to choose the distributed data storage strategy. In that case the costly data transfer to the gateway can be avoided and the sensor nodes may also perform data aggregation/compression/fusion to reduce the communication overhead in the network.

The following presents two major theories about distributed data storage.

Local Storage

The simplest distributed storage solution is to store all generated data its place of origin; that is, in the memory of the sensors. The primary advantage of this system is that since the data is stored locally, there is no communication cost associated with delivering data to storage. Of course, this storage strategy is not perfect. The primary disadvantage is the cost of query processing. Since no global distribution center is available, each node must take an active part in the polling process. Nodes need to parse and process queries and as well as be involved in the transmission of query results. That represents significant energy consumption at each node.

As an actively researched area, a number of solutions have been proposed which can optimize the transmission routes between the nodes and reduce communication costs in the network. During the processing of queries, a significant role is considered for the optimum operation of the network layer and efficient routing algorithms. One such algorithm is known as directed diffusion. To summarize, the initial request for data from the base station goes to all nodes, flooding the network. Afterwards, the nodes generate gradients in order to identify where the request is coming from and which neighboring nodes have the same data. When this type of request is repeated, the network has already formed the most optimal path for processing it. If queries are periodic, efficiency can be greatly increased. However, in the case of a new query, it is necessary to start over and flood the network again. [7]

Data-centric storage

Between centralized and local storage, there is a middle course specifying that data must be stored within network on predefined nodes according to the type of measurement and the place of its origin. The central premise of the theory is to store the same data type on the same network location. After that, any query that relates to a particular type of data is directly addressable to that data's storage, thereby avoiding queries associated with high communication costs. The request can therefore be processed more efficiently, decreasing the energy and latency of processing. Data storage in this system consists of two steps: If a sensor detects an event, it will label it using a hash function; following this, the data will be directed to a node according to the label. This process, GHT (geographic hash table), is a hash function that provides solutions for mapping between the name and storage location of the event. Data centric-storage creates a distributed storage structure, where events are named and grouped spatially. Names can be regarded as hash function arbitrary keys and categorize basic units. This method is also advantageous since the same data is stored in the same location. Through this, efficient aggregation can be achieved. Of course, there is a critical question regarding this storage strategy: How do we select the containing nodes? Additionally, for this strategy to function effectively, the nodes must have increased storage capacity and resources. [13]

In order to make queries more effective, a hierarchy between storage nodes is formed.

The stores nearer to the base station have greater spatial and temporal coverage but are more cumulative and contain compressed data. In contrast, storage in the lower levels is local and contains more detailed data. Queries start from the base station and then reach the accumulated data, only proceeding to lower levels of the hierarchy if greater accuracy is required. [10] Because of the finite capacity of the storage nodes, algorithms are needed that can manage data with the ability to delete or compress older values. Figure 2 describes the three storage strategies:

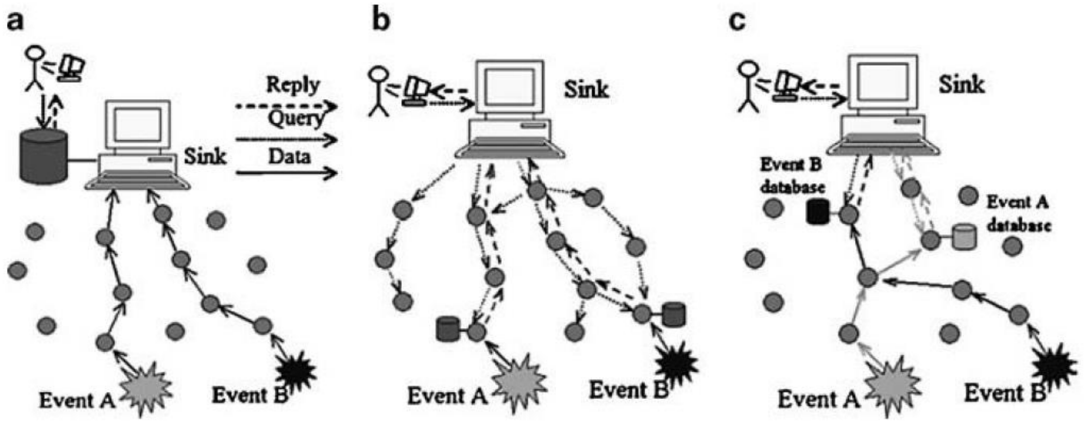

Figure 2 Sensor networks for storing data strategies. [12]

Sensor networks and cloud technology integration: sensor-cloud

Most applications use real-time data when possible. This enormous amount of data must be processed in some general way. As sensor networks have limited capacity, cloud technologies offer a significant opportunity to meet storage and computing capacity needs. Cloud-based Internet technology is the emerging new paradigm of software systems. Sensor networks and cloud technology become integrated as sensor cloud technology. The sensor cloud allows users to easily collect, store, analyze, visualize and share data using various web applications. The web applications utilize computing and storage capacity in the cloud. [8]

In military case the usage of sensor cloud is beneficial when there is a need for a scalable processing and storage infrastructure that can be used to perform the analysis of online as well as offline data streams provided by the sensors. Beside the analytics capabilities, the possibility of easier sharing of information among other defense agencies and coalition partner can significantly improve situation awareness. Figure 3 illustrates a sensor-cloud structure.

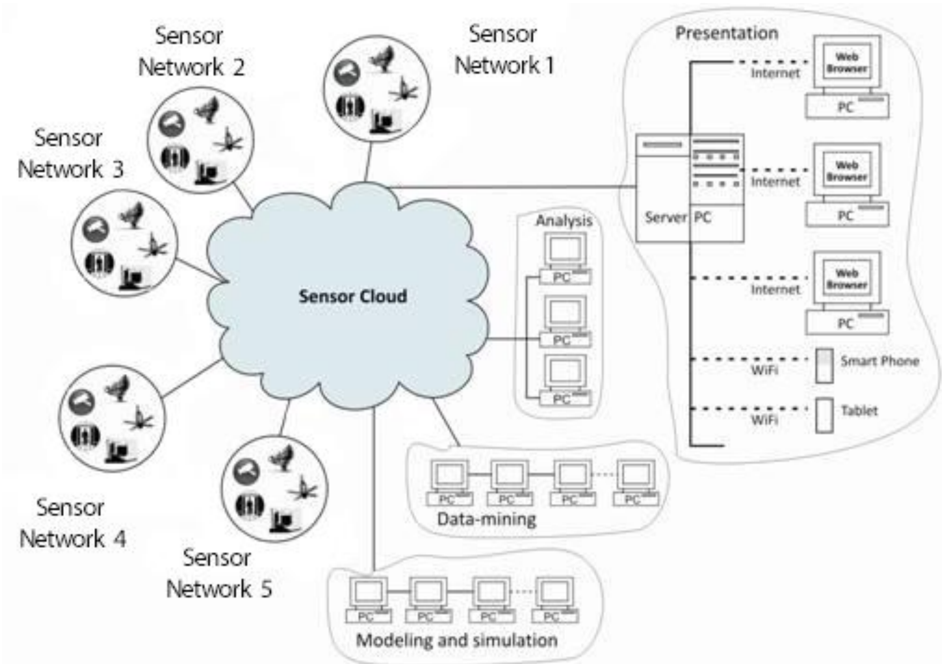

Figure 3 Sensor architecture cloud (edited by the author based on [17])

PROCESSING AND QUERYING OF SENSOR DATA

Besides storage strategies, another important data management topic is the way the collected data is transmitted to the end-user. How can users access the collected data of sensor networks? There is a need for solutions that both manage information collected by sensors and answer the queries of users or applications. Sensor networks can be viewed as complex data collection systems whose goal is to convert measured data as quickly as possible by automatically forwarding or replying to a query, according to the format provided to the end-user. The main purpose of data handling in sensor networks is to separate the logical view (name, access, operations) from physical data. It is not necessary for the user to pay attention to the operational details of the sensor network. The only significant element for the user should be the logical structure of queries. Considerable attention should be given to the question of query processing. For this purpose, the effectiveness of protocols depend on how promptly information is sent to the user. Several studies deal with the issue of managing large amounts of data collected from sensor networks. The primary goal is to give the users access to data gathered by the sensor network while using the least amount of resources possible, thereby increasing the useful life of the network. This chapter will introduce different query interfaces which allow for access to sensor network data, and techniques that optimize their efficiency.

Common types of queries

In order to optimize power consumption, we need to know the physical data: where, when, how often and what type of queries are needed. The most common query types are as follows: [9]

- **Queries on historical data**: queries relating to aggregated historical information stored in the database system. (" e.g. average temperature over the last month at the observed area")
- **"Snapshot" queries:** Queries relating to data collected on a given present or future date. (" Current body temperature of a soldier at the given time")
- **Long time continuously running queries:** Queries that collect information during a specified time interval. ("heart rate of a soldier measured at five minutes intervals during military training")
- **Event related queries:** Queries that are triggered by an event. ("If the motion sensor detects some movement at the border, check the pictures of the optical and infrared camera.")
- **Queries on multi-dimensional ranges:** Queries that involve multiple data attributes and define the desired search range. ("list the position of the sensors at the observed area in which the level of electromagnetic emission and the recorded seismic activity reach a pre-defined threshold.")

Query processing strategies

Depending on the location of the data storage, retrieval and data processing take place in different ways. The following section describes various strategies for accessing data in sensor networks.

Centralized query processing

The easiest and most commonly used solution is to store data in the base station, as was described previously with centralized data storage strategy. Sensors on the monitored area detect status changes in the physical environment and periodically transmit the measured values to the base station. In this case, the collection and processing of queries are separated. Query processing is done on the central server using conventional database management applications and algorithms.

Although this solution does not require methods special to sensor networks, the previously described disadvantages must be considered: significant delays might occur due to the increased load on the base station and on the surrounding nodes. Furthermore, due to continuous transmission of data, energy consumption is increased, and the useful life of the system is decreased.

Distributed Query Processing: Sensor Database

The sensors manufactured today are no longer just simple measuring devices, but also intelligent devices; although limited in scope, they have their own data storage and computing capabilities. To exploit this capability, new software solutions are needed that allow for pre-processing of queries using sensor resources.

There needs to be a query processing layer between the application and network layers whose function is to provide a user interface and to manage resources, especially with regards to power consumption. [9] In this case, the sensor network can be considered a distributed database where the nodes are data sources. In the most extreme cases, each node is running a database system for query processing. Tsiftes N. and Dunkels described such a solution called Antelope [19]. However, the more common option is for query interpretation, optimization, and transfer to the data sources to remain on the client-side: users send SQL queries to the server, which describe what kind of data the user interested in. Furthermore, users can specify how they want data to be combined, transformed, summarized, and ultimately how they would like to receive the query results. Using optimized routing strategies, the request is then transmitted to the sensors. At the sensor level, some data preprocessing may happen: filtering undesirable values and aggregating measurements. After this preprocessing, only the relevant information will be transmitted to the base station, significantly reducing communication cost. The most widely used implementation of the sensor database was developed on the TinyOS operating system and is called TinyDB (Berkeley University). In the next chapter, the main features of sensor database systems will be introduced according to the functions of TinyDB. [9][10][14]

TinyDB sensor database architecture and data model

In terms of architecture, it is necessary to separate operations on the server side from that of the sensors during query processing. In close collaboration, the two sides work together to ensure the efficient execution of queries. With the help of base station applications, we can formulate queries and understand SQL code interpretation. The goal is to optimize transmission of data towards the sensors using the most efficient routing protocol. For the sensors, there is a query processing subsystem running for each device on the TinyOS operating system. It receives queries on the nodes, optimizes their runtime, and maintains a catalog consisting of different metadata. The system offers the opportunity for data aggregation, energy efficient transmission, and cooperation between the different layers, which will be introduced in the following section based on the works of Madden, S. et al. [14][15] The structure of this system is shown in Figure 4 .

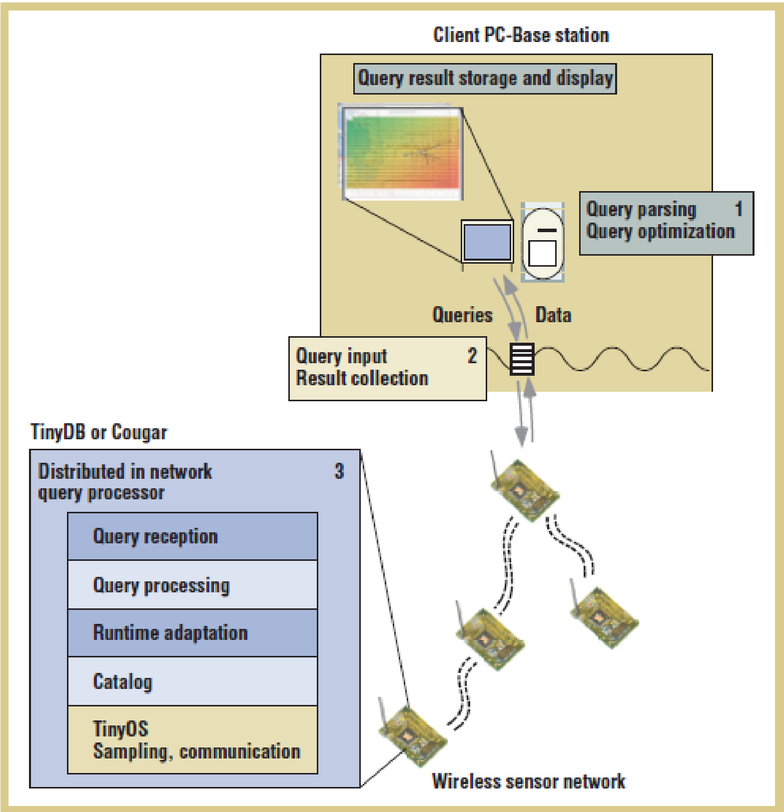

Figure 4 Sensors table (by the author)

The database-oriented approach has two basic objectives: making wording similar to that of conventional database systems thereby allowing simple SQL query commands, and to minimize the energy consumption of the network during the data collection process. [8]
In order to make simple SQL queries, it is useful to imagine a logic level virtual relational table above the physical distributed system in which the columns contain attributes (time stamp, node ID, values etc.) and the rows represent individual measurements.

Sensors table				
Timestamp	Node ID	Temperature (°C)	Location	…
2017-05-05 13:55	1	24	47.251894;21.544690	…
2017-05-05 13:55	2	26	47.250446;21.543759	…
2017-05-05 13:55	3	30	47.249903;21.539952	…
2017-05-05 13:56	1	25	47.251894; 21.544690	…
…	…	…	…	…

Table 1. Sensors table (by the author)

As mentioned previously, in addition to the traditional query types, sensor networks must also handle queries on data streams. The already known SELECT column1, column2, sensors FROM table [WHERE condition] [GROUP BY column1 [column2]] [HAVING group condition] command is therefore extended to the [PERIOD SAMPLE FOR duration] or [LIFETIME period] clause, which specifies the data collection frequency and duration, allowing for continuous queries of data streams. Also, the [STOP ON EVENT (parameter) and WHERE (parameter)] clauses can be used to allow for event-based suspension of queries, and the [OUTPUT ACTION (event)] clause can be used to specify an event that is executed if the statement conditions are true. Of course, various aggregation functions are also supported: (AVG, MIN, MAX, COUNT, SUM).

OPPORTUNITIES OF QUERY OPTIMIZATION

In sensor networks most of the energy consumption results from communication.
The primary task of the query optimizer system is to decrease energy consumption. The most effective way to achieve this is to reduce the amount of data transmitted between the base station and the sensors. It aims at finding routes which are able to send the data to the right place. While temporary storage (cache) can be used as a basic solution, more complex options are available. The following presents the most significant optimization opportunities in terms of executing queries.

Data aggregation within the network

The primary task of data aggregation is to reduce the amount of transmitted data moving across the network, thus reducing energy consumption. Through aggregation, expensive communications are replaced by cheaper computing operations. Nodes can be grouped in predefined ways, extending the useful life of sensor networks. Aggregation works either by summarizing the values generated during a specified period of time, or by summarizing the data from particular nodes. Query optimization can be achieved through proper selection of different aggregating operator sequences. These sequences exploit all the available possibilities of distributed processing. Aggregation functions must be broken down to simpler functions, partial aggregation operations must be executed in parallel on the nodes that in turn contribute to the calculation of the overall aggregation. Some nodes summarize the values obtained from nodes in lower levels, so that the base station only receives aggregated values.

Cooperation between layers

One extensively researched method of optimizing queries is by exploiting cooperation between different layers and coordinating their operation. For instance, knowledge about the current physical status of the data helps facilitate efficient channel assignment between layers. In the case of processing queries, network data layer operation as well as coordination of queries may be considered. Such solutions can be found in the TinyDB system. In this system, answering and sending query results to the base station is made along a predefined path. The devices have information about their closest neighbor and about the network structure. In order to facilitate intelligent routing decisions, each node is aware of its relationship with nearby nodes. Knowing the needs of the data, aggregation of queries can be enabled according to optimal routing solutions. On a deeper level, knowing the routing strategy makes it possible to assign the channels according to communication needs. Through this it is possible to switch the nodes to standby or off-line mode, thereby reducing their energy usage. Figure 5 illustrates a distributed implementation in a simple counting operation. Nodes with various distances from the gateway (hop count) transmit in specified time slots according to the channel. On the lowest level of the routing tree, the farthest nodes, in this case, H, J and I transmit first. At the next level up, the nodes G and F perform partial aggregation and transmit those aggregate values in the subsequent time slot. Nodes E and D at the following level in the communication topology aggregate those values again and transmit to nodes B and C. At nodes B and C, further partial aggregation occurs, with the final result then forwarded to the gateway.

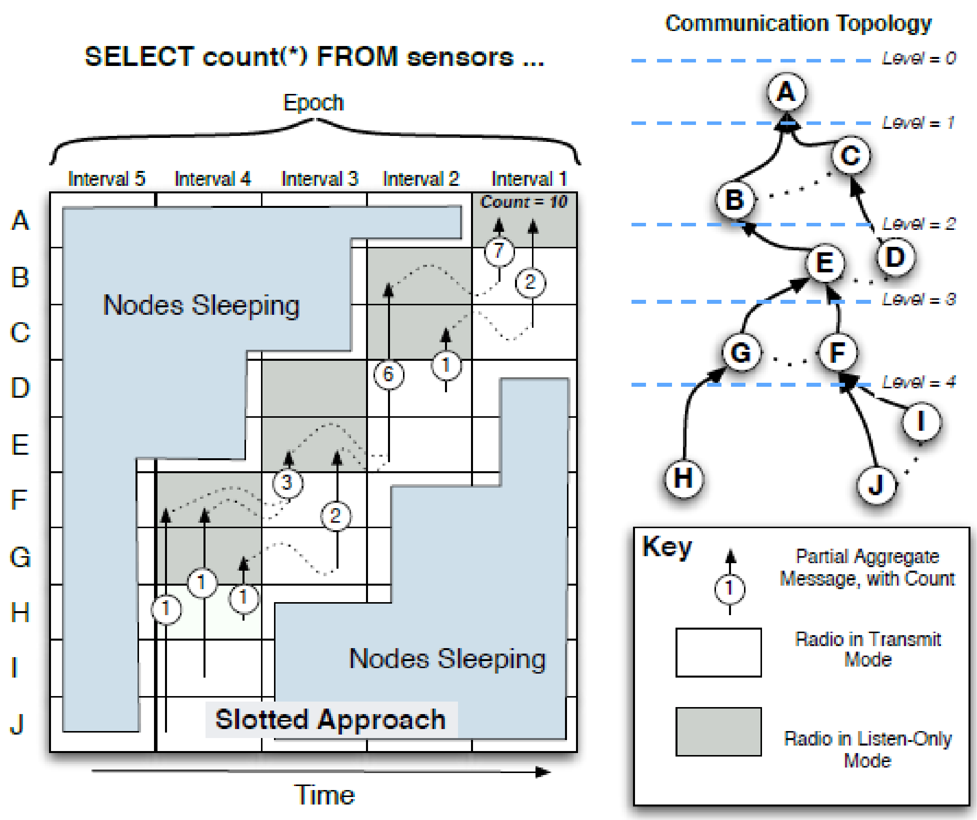

Figure 5 TinyDB aggregation within the network cooperation between layers [15]

Data collection with mobile robots

The above data management optimization solutions can be used effectively to a certain degree, but only if the installation area of the system remains small. With present technology used in data collection and transmission, there is still too much inefficient movement of data across the system. The longer the data route, the higher the communication costs, energy consumption, unreliability of data transmission and the probability of data loss. In addition, maintaining large routing tables is costly in itself. However, mobile robots offer a possible solution to the above problems. [8][18] Imagine some system where mobile robots acting as data mules collect data generated by the sensors. The big advantage is that robots have the ability to get close to the collection sites and can upload data with relatively low communication costs. They can save the data transmission system from multi-hop route difficulties. Short-range communication means less chance for data loss. Once each node collects the data, they send it to an online database. With this approach, the nodes can save considerable amount of energy that would otherwise be used to transmit data, thus prolonging the useful life of the network. Last but not least, it is easier to charge the batteries of robots than replace those the of nodes. However, for such an installation to be successful, new problems need to be tackled, such as the design of the robot path according to the needs of data collection. Figure 6 illustrates the structure of a system including mobile robots. [8]

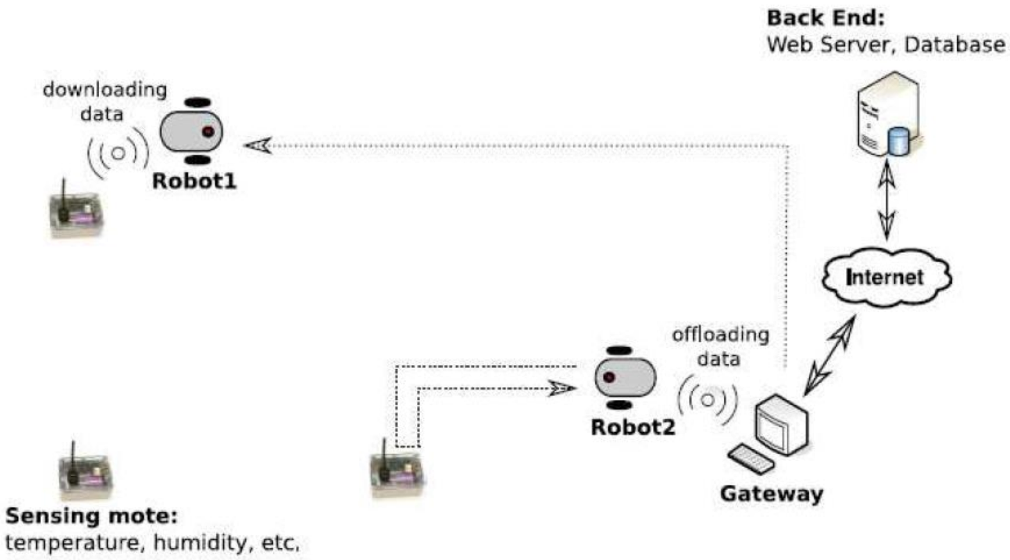

Figure 6 The use of mobile robots in the collection [8]

NOSQL based query processing systems

As mentioned above, enabling SQL based query-processing in sensor databases represents an especially effective solution. Research in this area continues to advance: (e.g. Jan Sipka, Van Der Veen et al. [16]). It’s predicted that NoSQL based query processing systems will be added to traditional relational databases. They are better adapted to dynamic environments, are based on a simple data model, and yield faster query processing. They allow the use of NoSQL database technologies (e.g. MongoDB[1], Apache Cassandra[2], HBase[3] etc.). Considering the data needs of sensor networks, it is more optimal to use NoSQL databases in several respects. Collection is done by modifying data needs, as opposed to more complex queries that are part of relational database management. Therefore, NoSQL databases also most ideal regarding speed of query processing, in particular for continuously generated data streams in large amounts.

SUMMARY

In selecting a suitable data management strategy, the issues of storage and query processing go hand in hand. Deep knowledge of the different storage models and related query processing solutions is needed. Addressing data management strategy is essential for further developments in the field. Table 2. summarizes the strategies and relative merits of the various data storage systems.

	Centralized data storage	**Distributed data storage**	**Storage in cloud**
Communication between nodes and the base station	High	Low	High
Communication between nodes processing queries	Low	High	Low
Scalability	Slightly scalable	Scalable to some extent	High scalability
Optimal size of the network	Small	Medium	Large
Optimal quantity of data being treated	Small	Medium	Large
Optimal data type	Mostly homogeneous	Mostly homogeneous	Homogeneous and heterogeneous
Security Challenges	Compromising gateway	Compromising gateway, Node vulnerability, Data manipulation	Compromising gateway, Less control over data
Query Management	Relational and traditional databases	Sensor distributed databases, SQL-based query processing (e.g. TinyDB)	Distributed databases in the cloud
Optimal query types	Event-driven, Ad hoc queries on past data	Continuously running queries for extended times, Multidimensional queries for domain	Past data for queries; "Snapshot," Continuously running, Event-driven and/or

[1] https://www.mongodb.com/
[2] http://cassandra.apache.org/
[3] https://hbase.apache.org/

			Multidimensional queries for domain
Processing resource constraints	Server	Server and nodes	Theoretically unlimited
Processing and analysis of data	Limited by capacity of local server	Limited by capacity of local server and nodes	Unlimited computational and analytical capacity, Advanced Big Data technologies
Users	Users of the given application	Users of the given application	More sensor database users
Potential military application	force protection/base protection	battlefield surveillance using unattended ground sensors	battlefield surveillance with enhanced data analytics and information sharing capabilities

Table 2. Comparison of sensor networks and storage strategies (by the author)

CONCLUSIONS

Due to the low cost and practical versatility sensors have already become widespread in the civil sphere and they also have proven applications in military and defense. The major sensor network data management solutions demonstrate that the range of possibilities is quite broad. Presently, traditional centralized storage is standard; however, as sensor network data becomes more complex and sensor size shrinks, new solutions are required. Ongoing research in this field points towards distributed data management within the network. A major goal of this research is to optimize the energy consumption of nodes. This is best achieved by using efficient routing strategies that enable communication between nodes and greater collaboration between protocol layers. As the role of Big Data and the Internet of Things grows, local resources will become obsolete and eventually replaced by more solutions with a greater degree of specificity and customization. Improvements in hosting and cloud technologies will render these new network solutions nearly unlimited in their scope. Advanced data analysis solutions can provide computing capacity on demand. In the process of selecting a particular technology, it is important to consider the needs of the application or user. It is relevant to consider the type of data being stored, retrieved, analyzed and selected. Furthermore, it is essential to choose the most appropriate and efficient storage and query-processing procedures.

BIBLIOGRAPHY

[1] NAGY T. I.: *Hálózati réteg a szenzorhálózatokban*; Hadmérnök, VII 3 (2012) pp. 123–130.

[2] NAGY T. I.: A felügyelet nélküli szenzorhálózatok és a programozási nyelvek kapcsolata; Hadmérnök, IV 4 (2009) pp. 303–311.

[3] HAIG ZS.: Networked unattended ground sensors for battlefield visualization; AARMS, 3 3 (2004) pp. 387–399.

[4] HAIG ZS.: Network-Centric Warfare and sensor fusion; AARMS, 2 2 (2003) pp. 245–256.

[5] KOVÁCS L.: Az elektronikai felderítés korszerű eszközei, eljárásai és azok alkalmazhatósága a Magyar Honvédségben. Budapest: ZMNE, 2003. (PhD értekezés)

[6] ZHAO, F., GUIBAS, J. L.: Wireless Sensor Networks: An Information Processing Approach; Morgan Kaufmann 2004.

[7] FAN, W., HAO, Z., GUO, Z.: Wireless Sensor Network Data Storage Optimization Strategy In: WANG, X. et al.: Advanced Technologies in Ad Hoc and Sensor Networks, Proceedings of the 7th China Conference on Wireless Sensor Networks; Springer 2014. pp. 345-351.

[8] RAWAT, P. et al.: Wireless sensor networks: a survey on recent developments and potential synergies; The Journal of Supercomputing 68. 1. (2014) pp. 1–48.

[9] WEI-PENG, C., JENNIFER C. H.: Data Gathering and Fusion in Sensor Networks In: IVAN, S.: Handbook of Sensor Networks: Algorithms and Architectures; John Wiley & Sons 2005. pp. 493-526.

[10] TRIGONI, N. et al.: Querying of Sensor Data In: GAMA, J., GABER, M. M. (Eds.): Learning from Data Streams, Processing Techniques in Sensor Networks; Springer 2007. pp. 73-84.

[11] DIALLO, O. et al.: Distributed Database Management Techniques for Wireless Sensor Networks; IEEE Transactions on Parallel and Distributed Systems 26. 2. (2015) pp. 604-620.

[12] JALLAD, A-H., VLADIMIROVA, T.: Data-Centricity in Wireless Sensor Networks In: MISRA, C., S., WOUNGANG, I., MISRA, S.: Guide to Wireless Sensor Networks. Computer Communications and Networks; Springer 2009. pp. 183-204.

[13] GOVINDAN, R.: Data-Centric Routing and Storage in Sensor Networks In: RAGHAVENDRA, C. S. et al.: Wireless Sensor Networks; Springer 2004. pp. 185-205.

[14] GEHRKE, J., MADDEN, S.: Query processing in sensor networks; IEEE Pervasive Computing, 3. 1. (2004) pp. 46-55.

[15] MADDEN, S. et al: TinyDB: an acquisitional query processing system for sensor networks; ACM Transactions on Database Systems, 30. 1. (2005) pp. 122-173.

[16] VAN DER VEEN, J. S. et al.: Sensor Data Storage Performance: SQL or NoSQL, Physical or Virtual; 2012 IEEE Fifth International Conference on Cloud Computing, 2012, pp. 431-438.

[17] MISTRAL SOLUTIONS: The Sensor Cloud and Homeland Security; Online: https://www.mistralsolutions.com/sensor-cloud-homeland-security/ (Letöltve: 2017. 11. 26.)

[18] CAO, H. et al.: Cloud-Assisted UAV Data Collection for Multiple Emerging Events in Distributed WSNs; Sensors, 17. 8. (2017), Online: http://www.mdpi.com/1424-8220/17/8/1818 (Letöltve: 2017. 11. 26.)

[19] TSIFTES, N., DUNKELS, A.: A database in every sensor; SenSys '11, Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems, 2011, pp. 316-332.

[20] LANGSON, J.: Security Implications of Data Dissemination Methods in Wireless Sensor Networks, Online: https://pdfs.semanticscholar.org/9734/15286a035d2a541d7599e7faf06512229324.pdf (Letöltve: 2017. 11. 26.)

[21] WINKLER, M., TUCHS, K. D., HUGHES, K., BARCLAY, G.: Theoretical and Practical aspects of military wireless sensor networks. Journal of Telecommunications and Information Technology, 2 1 (2008), 37–45.