Safety in Distributed Sensor Networks: A Literature Review

Tobias Altenburg\textsuperscript{a}, Sascha Bosse\textsuperscript{b} and Klaus Turowski

Very Large Business Applications Lab, Faculty of Computer Science Otto-von-Guericke University Magdeburg, Germany
\{tobias.altenburg, sascha.bosse, klaus.turowski\}@ovgu.de

Keywords: Dependability, Sensor Network, Internet of Things, Fault Tolerance, Fault Prevention.

Abstract: The connectivity megatrend dominates the current social change. The number of networked devices and the resulting amount of data is constantly increasing worldwide. For this reason, the dependability of computer systems is becoming increasingly relevant. Especially in the context of civil infrastructures, the constant availability of computer systems is of great importance. This paper provides a structured overview of the current literary status of safety in distributed sensor networks. Most approaches from the literature focus on the design phase. By the following connection with the existing dependability theory, the potential for the optimization of dependability could be proven.

1 INTRODUCTION

The visionary article "The Computer for the 21st Century" by Marc Weiser (1991) is the birth of the integrative technology paradigm later known by the term “Internet of Things”. For the first time he takes up existing technological developments and brings them into a relationship. Driven by the laws of Robert Metcalf (Gilder, 1993) and Gordon Moore (1965) the steady progress of microelectronics, communication and information technology continues. The number of networked devices is estimated to be about 75 billion by 2025 (Lucero, 2016) and the resulting global data volume will increase to 175 Zettabyte (Reinsel, Gantz, and Rydning, 2018).

The background to this rapid increase is the new way in which sensor networks can obtain data (Borgia, 2014; Akyildiz, Su, Sankarasubramaniam, and Cayirc, 2002). These networks use sensory acquisition to transfer the dynamic real world into the digital world in real time. In contrast, conventional information systems have manual data entry. Borgia (2014) illustrates these essential architectural differences of general information systems and sensor networks. The Internet of Things (IoT) has become a key technology for forward-looking scenarios. The applications are manifold. The use in civil infrastructure facilities in particular has a particularly high social relevance and impact (BMI, 2009).

These critical infrastructures are organizations or institutions of importance to the community. Their systems and services, such as the supply of water or electricity, are increasingly dependent on highly available and functioning information technology. A malfunction, impairment or even failure can lead to significant disturbances of public safety or other dramatic consequences (BSI, 2018). The resulting dependence of progressive society on complex information systems is constantly increasing and digital systems are consequently entrusted with ever more critical tasks that require a correspondingly high degree of dependability (Nelson, 1990; BSI, 2015). The categorical statement that “anything that can go wrong will go wrong”, known as Murphy's Law, already describes the importance of dependability in information and communication technology. Domains such as Smart Cities and Smart Grids are becoming increasingly important as a result of global population growth and urbanization. The existence of such a large network entails an enormous risk (Andrea, Chrysostomou, and Hadjichristofi, 2015). The conventional dependability procedures for safety-critical information systems are not universally applicable in the context of IoT (Boano et al., 2016). The challenges here are the resource-limited hardware and the basically complex IoT environment, e.g. the harsh environmental conditions or the high dynamics in the communication network.

\textsuperscript{a} https://orcid.org/0000-0002-2490-363X
\textsuperscript{b} https://orcid.org/0000-0002-1433-4912
itself. Since the stability of information systems has become a national or worldwide social concern of the highest priority, methods must be researched which significantly improve the dependability of future information systems, especially in critical infrastructures (Avizienis, Laprié, and Randell, 2001).

The field of dependability research was shaped by Jean-Claude Laprié. He established a standard framework and general terminology for reliable and fault-tolerant systems (Laprié, 1995). A distinction is made between dependability methods and dependability validation (Avizienis, Laprié, Randell, and Landwehr, 2004). Fault prevention and fault tolerance are part of dependability procurement and form the focus of our future investigations. We assume that these two methods promise the greatest potential in terms of optimizing dependability, while ensuring simple and cost-effective implementation.

Error prevention refers to methods which are intended to prevent the occurrence of an error condition or the introduction of an error cause into the system (see also Laprié, 1995; Avizienis et al., 2004). These occur during the design phase or during the runtime of the system. In contrast, in fault tolerance methods are developed which are intended to avoid a failure even though a cause of the fault already exists in the system (Laprié, 1995; Avizienis et al., 2004). Two basic concepts are relevant in the environment of distributed sensor networks - Security and Safety (see also Avizienis et al., 2004). The modelling of errors can include technological, human and organizational factors. The focus in science and practice in most cases lies in the development of approaches in the context of security. However, safety is often not considered as another important system property (Kufás, 2002). In order to be able to make a scientific contribution in this area, our investigation focuses on safety, i.e. protection against random failures in computer systems (Idso, and Jakobsen, 2000).

In literature, the different dependability methods are usually considered independently and isolated from other methods. For this reason, the goal is to define a pattern catalogue that can be used as a set of rules for improving safety in distributed sensor networks. This contains generally valid solution schemes for recurring problems in the context of safety, especially in fault prevention and fault tolerance. The defined patterns show the different possibilities for dependability optimization under consideration of the technical as well as logical correlations. This provides an essential decision-making basis for the optimization of safety, especially in the area of IoT. Based on this literature review, we would like to answer the following research question:

"Is the use of classical safety approaches qualified for dependability optimization in distributed sensor networks?"

This article presents the current state of the art for the dependability of distributed sensor networks, based on a structured literature review according to Webster and Watson (2002). Below, the structure of the literature search and the material collection is described first. Afterwards a descriptive analysis is presented in section (3). Section (4) gives an insight into the defined analytical categories and section (5) presents the corresponding results of the material evaluation. Section (6) concludes this contribution by summarizing the paper and outlining the way forward.

2 LITERATURE REVIEW

PROCESS AND MATERIAL COLLECTION

The literature review process derived from Seuring and Müller (2008) and includes four steps:

- Material collection
- Descriptive analysis
- Category selection
- Material evaluation

The material collection identifies the relevant contributions for the research area. The formal aspects of the relevant material are then evaluated in the descriptive analysis in order to be able to present the result set in a ordered form. For the material evaluation, content dimensions are defined in step 3 and then analysed in order to achieve an interpretation of the results. A paper is considered relevant if it provides theoretical foundations for the dependability paradigm or approaches to optimizing dependability in the context of distributed sensor networks. The scientific investigation focuses on fault tolerance and fault avoidance as dependability methods according to Laprié (1995). For the work, essentially software-technical, hardware-technical as well as interaction approaches are classified as relevant. The first step of material collection was a keyword search in the most important publication databases of the IT domain. The following expression was used for the search term: (Reliability OR Dependability OR Availability OR Robustness) AND (fault tolerance OR fault prevention OR fault Avoidance) AND (Wireless sensor networks OR Sensor Network OR distributed sensor Networks OR Internet of Things). About 15,000
articles were found using the keyword search. After these results were filtered from the databases by checking title and abstract for relevance, 190 papers remained. Subsequently, content filtering was carried out by cursory reading of the documents, which resulted in 82 relevant publications.

3 DISSCRIPTIVE ANALYSIS

This section evaluates the formal attributes of the selected publications. For this reason, the year of publication and the type of publication are taken into account for each relevant contribution.

Figure 1: Number of relevant publications per year.

Figure 1 shows the number of publications per year from 1993 and earlier until 2017. Research in the field of dependability methods began before 1993 (Avizienis, 1976; Iver et al., 1991; Avizienis, 1967; Laprié, 1985) and this work forms the basis for subsequent publications, since the generally developed theories on dependability are generally applicable to all information systems (Laprié, 1995). The distributed sensor networks as an architectural concept were first analysed in 1980 (Wesson, and Hayes-Roth, 1980) and later coined in 1999 by the term “Internet of Things” (Ashton, 2009). With the breakthrough of IoT paradigm, the use of complex IT systems, which are also used in critical application domains, is now growing. This development increases the need for resilient ICT and is becoming increasingly important (Boano, and Römer, 2014). In Figure 1, the linear trend line (represented by the dotted line) of the number of papers per year shows that the number of publications is increasing. This illustrates the growing relevance of the topic and the existing research interest. As early as the mid-2000s, most of the contributions addressed possible applications and the safety-related aspects lost their relevance.

Figure 2 shows that the identified contributions were published in four types: Journal articles (25), book chapters (13) as well as conference papers (26) and symposium contributions (18). The high number of journal articles indicates that there are completed research projects already that are known to have validated approaches for optimizing dependability. More than half of the relevant papers are conference and symposium papers, which highlights the intensive discussions on the topic. Book chapters often provide a detailed theoretical basis for the development of relevant approaches and thus form the basis for the field of research.

4 CATEGORY SELECTION

This paper examines approaches for optimizing the dependability of distributed sensor networks by using dependability procedures, i.e. error avoidance and error tolerance (Avizienis et al., 2004). The identified approaches serve as a basis for deriving the categories for material evaluation and are classified accordingly in six categories, cf. Table 1.

Table 1: Categories and common values for material evaluation.

<table>
<thead>
<tr>
<th>Category</th>
<th>Common Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>architecture layer</td>
<td>utilization layer, transmission layer, collection layer, none</td>
</tr>
<tr>
<td>temporal phase</td>
<td>pre life, peri life, none</td>
</tr>
<tr>
<td>scope</td>
<td>one Node, many, all, none</td>
</tr>
<tr>
<td>error class</td>
<td>design, physical, interaction, none</td>
</tr>
<tr>
<td>error type</td>
<td>fault, error, failure, none</td>
</tr>
<tr>
<td>evaluation type</td>
<td>example, experiment, case study, argumentative, other</td>
</tr>
</tbody>
</table>

The methods for optimizing the dependability of distributed sensor networks can be divided into the three basic architecture layers (see also Borgia, 2014). The collection layer refers to procedures for capturing
the physical environment and for general perception of the real world. The transmission layer, which contains the mechanisms for rendering the collected data to the application level or the attached services, is superordinate. The utilization layer is the actual processing, administration and usage level. This includes the processing of information flows, the bidirectional forwarding of data and general functions such as device management or data aggregation.

In addition to the basic architecture, time \( t \) plays an important role in the dependability domain, because it is directly related to dependability \( R(t) \) (see also Laprié, 1995). According to Dubrova (2012) and Laprié (1995), a distinction is made between the design and operation phases. The overall model is a temporal classification of the methods into two life-cycle phases - the pre-phase and the peri-phase (Byron, and Blake, 2019). In the pre-phase, the focus is on planning and designing the information system at the software and hardware levels. According to Laprié (1995), the design of an IT system has the greatest influence on its dependability. Suitable procedures are to be defined, which also optimize dependability in the operating phase through fault tolerance or fault avoidance. The peri-phase is the actual operating phase of the IT system. Here methods are classified, which can be used exclusively during the performance of the system. The scope category was examined to determine the intensity of the various optimization methods. This analyses the size of the efficiency in relation to the number of nodes in the sensor network. A distinction is made between three orders of magnitude:

- One Node - here procedures are classified which refer to only one single node (Cotroneo, Natella, Pietrantuono, and Russo, 2014; Grey, and Siewiorek, 1991).
- Many - by this we mean dozens to several hundred nodes, which simultaneously achieve an improvement in dependability through the respective method (Asim, Mokhtar, and Merabti, 2008; Johnson, 1996; Cooper, 1985).
- All - all nodes of the entire network segment are influenced by the method used (Denning, 1976; Bishop, 1988; Lyu, 1995).

In summary, it can be concluded that the sphere of influence of the dependability procedures examined represents an important indicator for the prioritization of the methods in the pattern catalogue. Due to the large number and high dynamics in current sensor networks (Tubaishat, Madria, 2003) the effort for the implementation and application of the methods can also increase significantly. Error classes can be derived from the already known error causes (Avizienis et al., 2004). According to Laprié (1995) there are three essential error classes to be distinguished, which we use for our investigation:

- **Design** - this includes all software and hardware errors that occurred during the design phase and affect the actual operating phase.
- **Physical** - contains all errors that affect the hardware of the three architecture layers (Borgia, 2014).
- **Interaction** - if all external and internal input or communication errors occur between the various nodes or components of the information system.

A system is referred to as being correct when it performs as specified for the system. As soon as the system no longer meets these specifications, it switches to a faulty state (Avizienis et al., 2004). Laprié (1995) defines three fundamental error types which form the basis of error theory. The various dependability optimization methods are assigned to the respective fault types and thus provide an appropriate indicator of efficiency. In the area of dependability, the terms “failure” for downtime, “fault” for a cause of failure and “error” for malfunction have gained acceptance (Avizienis et al., 2004). A fault is the responsible or hypothetical cause of a malfunction. An error is the part of the system state that is responsible for a failure. In general, it is a deviation from the required operation of the system or subsystem. Failure occurs when the performance of the system no longer corresponds to the correct or expected performance. It is the transition from a correct to an incorrect system state. The system therefore no longer performs the required functions.

Figure 3 of Avizienis et al. (2001) shows the error process described in a component. If a fault is activated, it can cause an error, which in turn (if it spreads) can cause a failure. This failure is evaluated as a fault by the next component.

![Figure 3: Fault-Error-Failure chain.](image)

Thus, it can be deduced that the analysed methods can only be subdivided into the error classes fault and error, since according to Avizienis et al. (2001) causal chain a failure can only occur by the propagation of an error. A further differentiation of the evaluated papers is provided by the individual type of evaluation. In this analysis, the evaluations are classi-
fied according to the following schema:

- An exemplary assessment was performed when a hypothetical scenario was modelled to demonstrate the basic applicability of the developed model.
- An experimental evaluation was performed when a realistic scenario was modelled. This simulation offers a realistic representation.
- A case study was performed when a real scenario was modelled and analysed. This proves the applicability of the model to real problems.
- An argumentative evaluation was carried out when a conclusive argumentation was used to prove the methods.
- All papers that do not contain any of the previous valuation types are classified under miscellaneous.

5 MATERIAL EVALUATION

Once the categories have been identified, the quantitative analysis of the 82 selected papers can be carried out. The design of a system for high dependability implies the need to identify and consider the various possible causes of failure. Figure 4 shows the classification of the approaches into the three essential error classes according to Laprié (1995).

![Figure 4: Number of contributions per error class.](image)

All approaches to optimizing system dependability can be classified into the three error classes. Eleven publications deal with the general theoretical basics and thus form the basis for the investigation. Most methods are assigned to the error class Design. This proves the high importance of the construction phase of information systems, because already at this point measures can be taken to prevent errors or the methods for error tolerance can be implemented and synchronized at an early stage. Nevertheless, the other two fault classes have a similar potential to significantly increase dependability in distributed sensor networks. Errors in hardware components or the interaction between the various system components must be equally avoided or tolerated with regard to dependability (Laprié, 1995). The previous analysis of error classes applied showed that the design phase plays a crucial role in optimizing the dependability of information systems (Byron, and Blake, 2019). The majority of the methods examined relate to software engineering (Birolini, 2017). Various measures for increasing dependability are possible along the entire development process. Holzmann (2006) offers an exemplary set of rules for the software development of safety-critical systems. In order to guarantee the correctness of the result in the algorithm created, the basic concept of design diversity (Bishop, 1988; Xie, Sun, Sahuja, 2008) or the already established method of software testing (Lyu, 2007; Luo, 2001) is used. Quality control techniques must also be used in the installed hardware in order to be able to avoid hardware errors in a targeted manner (Avizienis et al., 2001). The approaches mentioned help to avoid errors in software and hardware already in the design phase, but also fault tolerance, as an important dependability procedure, should guarantee the correct performance in case of active faults (see also Avizienis et al., 2001). Fault tolerance is usually a design goal, which is realized from the requirements of dependability and availability (Auerswald, Herrmann, Kowalewski, Schulte-Coerne, 2002). The error tolerance can basically be divided into reactive and proactive error tolerance. Reactive fault tolerance techniques are used to reduce the impact of failures on a system when the failures have already occurred. In comparison, the proactive fault tolerance techniques include preventive measures to avoid faults during the actual operating time of the IT system (Amin, Sethi, Singh, 2015).

![Figure 5: Number of contributions per scope.](image)

For the evaluation of the individual dependability methods, the scope can be used as a criterion. The aim is to be able to design as many components of the distributed sensor network as
possible more reliably using these methods. In this way, the complexity of the implementation or the difficulty of using the various methods can be approximately determined. Figure 5 shows the scope of the methods found in the Design error class. More than 75 % of the papers examined describe procedures that can be applied to several components of the information system. Figure 5 shows once again that there is great potential for optimizing dependability in the design phase of information systems. The use of high-quality components or systematic design techniques often does not sufficiently reduce the probability of system failures and means must be provided to tolerate errors in the system (see also Nelson, 1990). Accordingly, the use of a dualism between fault prevention and fault tolerance is essential to increase dependability. The following Figure 6 shows the time phase when the methods examined can be used appropriately in the event of an error. The classification is based on the active life cycle of a computer system.

![Figure 6: Number of contributions per temporal phase.](image)

More than 2/3 of the dependability procedures are assigned to the pre-stage, i.e. the design phase. In contrast, about a quarter of the papers examined describe methods that are applied during the operating phase. These results of Figure 6 illustrate Lapré's hypothesis (see also Lapré, 1995). According to Auerswald et al. (2002) and Lapré (1995) the complete avoidance of errors is practically impossible or too costly. After all the improvement of dependability in distributed sensor networks should be in an acceptable relation to the economic costs. This economic point of view has to be considered for all views concerning the optimization of dependability. A sensor network can be considered as a sum of components. Dependability is influenced by the interaction of the various system components (Avizienis et al., 2001). The threats to distributed sensor networks are characterized by faults, errors, and failures. These three types of errors according to Lapré (1995) and Avizienis et al. (2001) are good indicators for assessing the efficiency of the methods analysed. The causal chain of Avizienis et al. (2001) shows the logical course of errors in an information system. Basically, failure is to be avoided completely, whereby we have assigned the methods identified in Figure 7 to the two error types fault and error.

![Figure 7: Number of contributions per error types.](image)

Almost all identified papers describe methods to avoid or tolerate faults (Xie et al., 2008; Paradis, Han, 2007; Saridakis, 2002; Treaster, 2005). In particular, the software development process is at the forefront. Bishop (1988) and Torres-Pomales (2000) summarize the already identified patterns for fault-tolerant software in their work. These concepts are divided into two groups (see also Lyu, 1995) - Single versions and multi-version software techniques. Single version techniques focus on improving the fault tolerance of a single piece of software by adding mechanisms to the design. In contrast, multi version techniques creates several variants of a software in order to avoid design errors in one version. A third of the papers examined describe techniques that deal with the tolerance or avoidance of errors. According to Hanmer (2007) error processing is characterized by error detection and the subsequent error recovery. A fundamental distinction is made between forward and backward recovery mechanisms (Lapré, 1995; Iver et al., 1991; Saridakis, 2002). Errors can be avoided by implementing redundancy techniques already in the design phase or in the pre-phase by a proactive rejuvenation of the software (Cotroneo et al., 2014).

The analysis of the evaluations carried out in the identified papers shows that 1/3 of the findings are proven by an argumentative explanation, cf. Figure 8. In the other contributions an experimental or exemplary evaluation type is dominating. In only 11 % of the papers was a real case study carried out. These results show that although half of the identified papers are based on empirical research, only a small part was examined in a real environment. Therefore, it can be concluded that there is still a need for realistic approaches in reliability theory for distributed sensor networks.
6 CONCLUSION

Our goal is to maximize the dependability of distributed sensor networks through a pattern catalogue defined by us. The requirements in the real world are manifold. A distributed sensor network has to cope with a large spatial separation of the individual components. It has an open system structure, a high degree of autonomy for the individual system components and a high degree of heterogeneity. Therefore, it is important to be able to achieve standardization by means of a pattern catalogue. For this purpose, a taxonomy is created from the identified dependability procedures of the literature review. On the basis of this classification, characteristically similar methods can be summarized and subsequently standardized. Afterwards, the classified methods are evaluated in the context of distributed sensor networks and included in the pattern catalogue with regard to the optimization of dependability.

In this paper, a structured literature analysis was performed to analyse the current state of dependability theory in distributed sensor networks. Eighty-two relevant publications were identified by keyword search and filter process. After the descriptive analysis pointed out the relevance of the topic, the categories for the content analysis of the examined papers were defined. On this basis the material valuation could be carried out. The methods identified were analysed using the error classes from Laprié (1995). It turned out that most approaches in literature refer to design errors. The design phase thus offers a high potential to optimize the dependability of distributed sensor networks. The following specification of the procedures in the design phase showed that over 75% of the methods can be applied to several components at the same time. This significantly reduces the complexity of implementation into existing or newly emerging sensor networks. Furthermore, more than half of the methods found can be classified in the concept of fault tolerance and one third in error avoidance. According to the hypothesis of Laprié (1995) and Auerswald et al. (2002), this illustrates that errors can never be completely avoided and therefore the focus in literature is on error tolerance techniques. In order to be able to determine the efficiency of the identified dependability procedures, the error types were analysed at the end. Dependability tools are designed to reduce the number of errors. Faults can spread traditionally and cause errors or failures (Avizienis et al., 2001). In this respect, it is important to apply error theory in order to assess the effectiveness of the procedures. Increasing IT penetration and networking are creating potential that a highly developed and industrialized country cannot do without (BSI, 2017). The development of a fully interconnected ecosystem has a major impact on industry and society. At the same time, increasing digitalization is creating new risk situations. Computer failures can have considerable economic and social consequences, which must be responded to quickly and consistently (BSI, 2018). The sources of error are manifold, widely distributed and difficult to locate. It is therefore necessary to pay close attention to dependability and future maintainability when developing sensor networks.

REFERENCES


Kufâs, I, 2002. A framework for information security culture; could it help on solving the insider problem?. Informasjonssikkerhet og innsideproblematikk, NTNU.


