УДК 004.94:681.518
https://www.doi.org/10.47813/mip.7.2026.20.3021
EDN CWRDOM

Цифровой двойник распределённой сенсорной сети Smart Dust

А.А. Грейс[1*], И.В. Ковалев[1,2]

[1]Сибирский федеральный университет, пр. Свободный, 79, Красноярск, 660041, Россия
[2]Красноярский государственный аграрный университет, пр. Мира, 90, Красноярск, 660049, Россия

[*]E-mail: alena.yabl@yandex.ru

Аннотация. В статье рассматривается архитектура цифрового двойника распределённой сети автономных сенсорных узлов Smart Dust, в которой данные сети используются совместно с информацией о состоянии её узлов для определения параметров контролируемой физической системы. Предлагаемый подход основан на совместном учёте состояния сенсорной инфраструктуры и параметров контролируемой физической системы. В отличие от традиционных схем, в которых данные узлов используются только для обновления цифрового представления объекта, в предлагаемой архитектуре учитываются характеристики параметров измерительной информации (в частности, пространственное положение узла, остаточный энергоресурс, состояние калибровки, качество канала связи, уровень сетевой связности, временная актуальность измерений, согласованность показаний с данными соседних узлов и интегральная оценка достоверности формируемых данных). Предлагаемая архитектура состоит из двух взаимосвязанных контуров: цифрового представления сети и цифрового представления наблюдаемой среды. Для формализации процессов обработки информации введена пятислойная модель данных. Модель содержит уровень первичных сенсорных измерений, уровень верификации и предобработки данных, уровень формализованного представления состояния сенсорной инфраструктуры, уровень пространственно-временного восстановления состояния контролируемой физической системы и уровень интеграции состояния сенсорной инфраструктуры и состояния контролируемой физической системы в единую модель цифрового двойника. Для каждого слоя определены состав данных, состояние, операторы обработки и функции, взаимодействие между уровнями описано посредством операторов межслойного преобразования. Приведены модели оценки достоверности сенсорных данных, энергоресурса и сетевой связности узлов, пространственно-временного объединения измерений, реконструкции состояния наблюдаемой среды и синхронизации физического и цифрового представлений. Предложенный подход позволяет учитывать текущее состояние сенсорной инфраструктуры при формировании цифрового представления контролируемой физической системы и тем самым уменьшать влияние отказов узлов, деградации энергоресурса, нарушений калибровки, ухудшения характеристик каналов связи, снижения сетевой связности и временного устаревания сенсорных данных.

Ключевые слова: цифровой двойник, Smart Dust, архитектура.

Для цитирования: Грейс, А. А. Цифровой двойник распределённой сенсорной сети Smart Dust / А. А. Грейс, И. В. Ковалев // Модернизация, Инновации, Прогресс (MIP ENGINEERING-VII-2026): материалы VII Всероссийской (национальной) научной конференции (Красноярск, 11–12 сентября 2026 г.). – Красноярск: Красноярский краевой Дом науки и техники Российского Союза научных и инженерных общественных объединений, 2026. – Т. 20. – С. 170-188. – https://doi.org/10.47813/mip.7.2026.20.3021. – EDN: CWRDOM.

Digital twin of a distributed Smart Dust sensor network

A.A. Grace[1*], I.V. Kovalev[1,2]

[1]Siberian Federal University, 79 Svobodny pr., Krasnoyarsk, 660041, Russia
[2]Krasnoyarsk State Agrarian University, 90 Mira Avenue, Krasnoyarsk, 660049, Russia

*E-mail: alena.yabl@yandex.ru

Abstract. This paper presents a digital twin architecture for a distributed network of autonomous Smart Dust sensor nodes that integrates sensor measurements with information about the operational state of the network to estimate the state of a monitored physical system. The proposed approach accounts for both the condition of the sensing infrastructure and the parameters of the physical system under observation. Unlike conventional approaches that use sensor measurements solely to update the digital representation of a monitored system, the proposed architecture explicitly considers factors affecting the reliability of measurement data. These include sensor node position, remaining energy reserves, calibration status, communication link quality, network connectivity, measurement timeliness, consistency with readings from neighboring nodes, and an aggregate assessment of data reliability. The architecture comprises two interconnected digital representations: one describing the sensor network and the other representing the monitored

environment. A five-layer data model is introduced to formalize how information is processed and integrated across these representations. The model encompasses raw sensor measurements, data validation and preprocessing, a formalized representation of the sensing infrastructure's operational state, spatiotemporal reconstruction of the monitored physical system, and integration of the infrastructure and physical system states into a unified digital twin. Each layer is characterized by its data structures, state variables, processing operators, and associated functions, while interactions between layers are defined through interlayer transformation operators. The proposed framework includes mathematical models for assessing sensor data reliability, node energy reserves, and network connectivity, as well as for spatiotemporal data fusion, reconstruction of the monitored environment, and synchronization between the physical system and its digital representation. By incorporating the current operational state of the sensing infrastructure into the digital twin, the proposed approach accounts for changes and limitations within the sensor network when reconstructing the state of the monitored physical system. This helps mitigate the effects of node failures, energy depletion, calibration drift, deteriorating communication link quality, reduced network connectivity, and outdated sensor measurements on the resulting digital representation.

Keywords: digital twin, Smart Dust, architecture.

For citation: Grace, A. A. Digital twin of a distributed Smart Dust sensor network / A. A. Grace, I. V. Kovalev // Modernization, Innovations, Progress (MIP ENGINEERING-VII-2026): Proceedings of the VII All-Russian (National) Scientific Conference (Krasnoyarsk, September 11–12, 2026). – Krasnoyarsk: Krasnoyarsk Regional Science and Technology City Hall of the Russian Union of Scientific and Engineering Public Associations, 2026. – Vol. 20. – P. 170-188. – https://doi.org/10.47813/mip.7.2026.20.3021. – EDN: CWRDOM.

1. Introduction

Sensor technologies make it possible to monitor physical systems and processes directly under operating conditions. Smart Dust offers one approach to distributed monitoring through wireless networks of miniature, autonomous sensor nodes [1]. Each node incorporates sensing elements, onboard data processing capabilities, a wireless communication module, and a power source. Individual nodes can acquire measurements, perform local data processing, and transmit information to neighboring nodes or higher-level monitoring systems.

Smart Dust networks are particularly suitable for monitoring large, spatially distributed, or difficult-to-access systems that require dense sensor coverage, scalability, and long-term autonomous operation. Distributing sensor nodes across a monitored area allows measurements to be collected at multiple locations, providing information about spatial variations in the parameters of interest. The integration of spatially distributed observations into environmental monitoring information systems has also been investigated using GIS technologies and remote sensing data [2]. Processing measurements locally reduces the amount of data transmitted over the network and the associated communication overhead [3].

The reliability of these measurements, however, depends on more than the state of the physical system being monitored. It is also influenced by the condition of the sensing infrastructure. Declining energy reserves, sensor calibration drift, degraded wireless links, disruptions in network connectivity, packet loss, and transmission delays can all affect the quality of the data received [3, 4]. Consequently, constructing an accurate digital representation of a physical system requires information about the operational status, availability, and performance of the sensor network in addition to the measurements themselves. Accounting for these factors helps distinguish changes in the physical system from apparent variations caused by faults or limitations in the measurement infrastructure [5].

Digital twin (DT) technology provides a framework for combining measurements and system information into a coherent digital representation of the monitored physical system [6-7]. In Smart Dust networks, this representation needs to capture not only the physical parameters being measured but also the current state of the sensing infrastructure. The reliability, completeness, and timeliness of incoming measurements determine how accurately the digital model reflects the physical system at any given time.

Existing digital twin approaches for sensor-based systems largely concentrate on the acquisition, processing, storage, and interpretation of measurement data. The operational condition of the sensor network is often treated as supporting information rather than an integral part of the digital model. Such an approach has limitations in distributed Smart Dust networks, where changes in node availability, communication performance, or measurement quality can directly affect the accuracy and timeliness of updates to the digital representation of the monitored system. Related research on autonomous systems has also addressed the resilience of software control through multiversion execution environments [8], although this approach does not directly address the reliability of distributed sensor measurements in digital twin architectures.

A digital twin for a Smart Dust network should therefore incorporate two interconnected representations: one describing the monitored physical system and the other capturing the operational state of the sensing infrastructure that supplies its measurements [9]. The objective of this study is to develop an architecture for a distributed Smart Dust network digital twin in which both components are considered when constructing and updating the digital representation of the physical system.

2. Materials and methods

The proposed architecture adopts a dual-model approach, with separate but interconnected digital twins representing the Smart Dust sensor network and the monitored environment. Data are transferred from the physical domain to the digital models through a series of processing and transformation layers.

The integrated digital twin is defined as follows:

$$DT(t) = \langle DT_N(t), DT_E(t), \mathcal{H}(t), \mathcal{U}(t) \rangle,$$

where:

– $DT_N(t)$ represents the digital state of the sensor network;

– $DT_E(t)$ represents the digital state of the monitored environment;

– $\mathcal{H}(t)$ contains the history of system states and measurements;

– $\mathcal{U}(t)$ defines the set of control actions.

The overall architecture and the relationships between its components are illustrated in Figure 1. At the physical level, the Smart Dust network acquires measurements that are transmitted through the communication subsystem for preprocessing and validation. The processed data feed two interconnected digital twin models, one tracking the operational state of the sensing infrastructure and the other representing the monitored physical system. A dedicated analysis and control subsystem completes the architecture by processing information from both models and generating control actions.

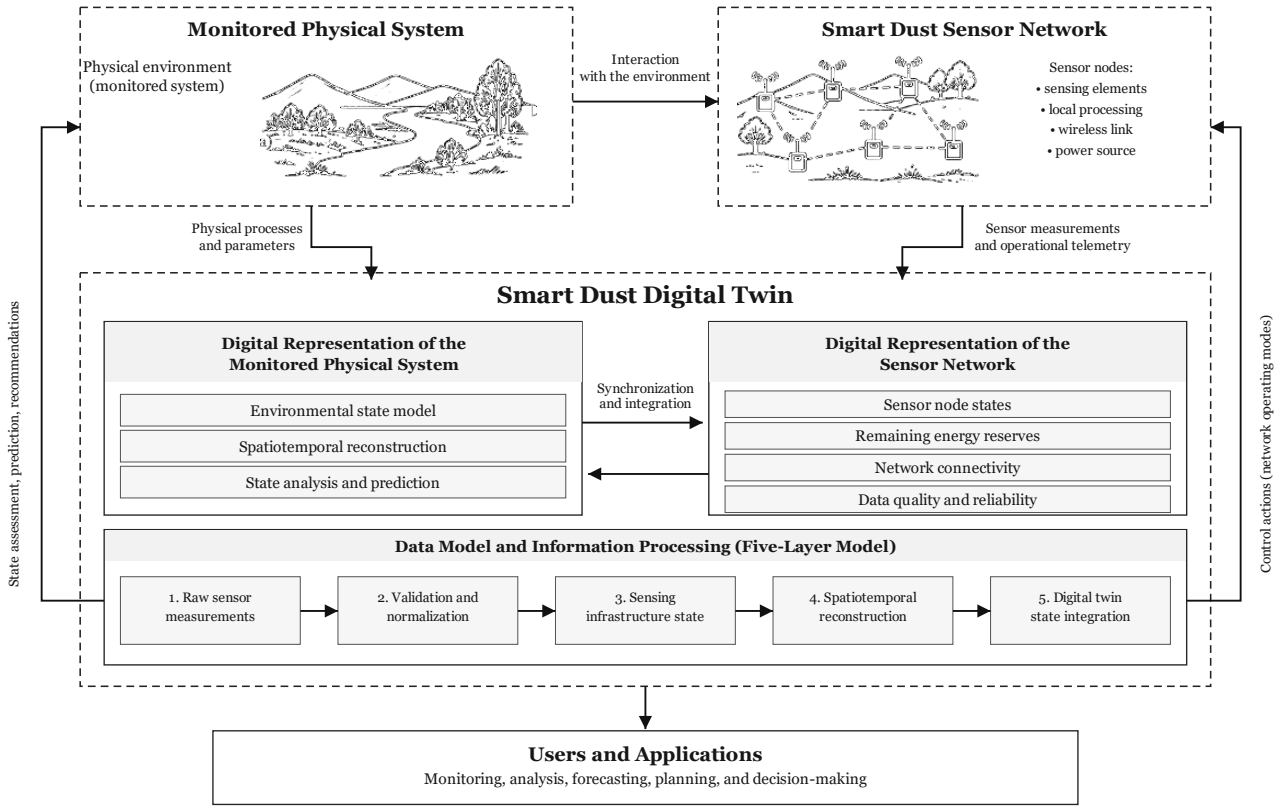

Figure 1. Structural and functional architecture of the Smart Dust digital twin.

The physical Smart Dust sensing environment consists of autonomous sensor nodes deployed throughout the monitored area. Each node interacts with its surroundings and measures a predefined set of physical parameters. Alongside these measurements, the nodes generate telemetry describing their current operating conditions, including spatial coordinates, remaining energy reserves, sensor status, and wireless communication parameters.

The communication subsystem handles data exchange between sensor nodes and transfers information to higher layers of the architecture. Depending on the network configuration, messages may be sent directly to a gateway or forwarded through intermediate nodes along multihop routes. Both sensor measurements and operational telemetry are transmitted through this subsystem. The

telemetry provides information about network connectivity, communication link quality, node availability, and routing conditions. These characteristics are subsequently used to assess the reliability of incoming measurements.

The raw data layer receives information from the physical sensor nodes before it is interpreted by the digital twin. The incoming data include measured values, source identifiers, timestamps, node coordinates, remaining energy reserves, communication parameters, and other operational telemetry. Messages are retained in their original structure at this stage, keeping data acquisition separate from subsequent processing and quality assessment.

The validation and normalization layer prepares the acquired data for use in the digital models. Processing at this stage includes checking message structure and completeness, synchronizing measurement timestamps, applying calibration corrections, identifying invalid or anomalous readings, and converting data into a consistent format. Measurement timeliness is also assessed, as the interval between data acquisition and subsequent use can affect the accuracy of digital twin updates. The output is a validated and standardized dataset that serves as the basis for further analysis.

The sensor network digital twin maintains a representation of the sensing infrastructure's current operational state. Its model incorporates network topology, node availability and operating status, remaining energy reserves, calibration status, communication link quality, network connectivity, and the timeliness of transmitted measurements. These characteristics are used to derive a reliability estimate for data received from each source. The network digital twin thus provides information about the condition of the sensing infrastructure and the extent to which its measurements can be trusted.

The physical system digital twin reconstructs the spatial and temporal state of the monitored environment. It uses validated sensor measurements together with information supplied by the network digital twin. Measurement fusion takes account of sensor locations, acquisition timestamps, and source-specific reliability estimates. The resulting model combines discrete observations into a consistent representation of the monitored physical parameters across the specified spatial domain at a given point in time.

The analysis and control subsystem draws on both digital twins to assess the current system state and determine appropriate control actions. Its functions include identifying anomalies, detecting sensor node failures and performance degradation, evaluating discrepancies between the physical system and its digital representation, and assessing the quality of the reconstructed model. The analysis may also inform adjustments to network operation, including measurement frequency, active node selection, data routing, and energy consumption settings.

Together, these components establish an integrated data processing and transformation pipeline, with the sequence of operations defined as follows:

$$\mathcal{P} \to L_0 \to L_1 \to L_2 \to L_3 \to L_4 \to \mathcal{U},$$

where:

– $\mathcal{P}$ denotes the physical environment and the physical sensing infrastructure;

– $L_0, \dots, L_4$ denotes the information layers of the digital twin;

– $\mathcal{U}$ denotes the control actions.

The feedback loop is defined by the mapping:

$$\mathcal{U}: L_4 \to \mathcal{P}.$$

In the proposed architecture, the digital twin is organized into two interconnected components. The first represents the current state of the Smart Dust sensing infrastructure and is used to assess node operability, communication conditions, and the reliability of the data being generated. The second represents the state of the monitored physical system and is constructed from validated measurements while accounting for the characteristics of the data sources.

The digital twin of the sensor network is defined as:

$$DT_N(t) = \langle G(t), X^N(t), R(t) \rangle,$$

where:

– $G(t)$ denotes the dynamic network connectivity graph;

– $X^N(t)$ denotes the set of sensor node states;

– $R(t)$ denotes the vector of node reliability estimates.

The digital twin of the monitored environment is defined as:

$$DT_E(t) = \langle \hat{S}(t), \Theta_E, H_E(t) \rangle,$$

where:

– $\hat{S}(t)$ denotes the reconstructed spatiotemporal state of the environment;

– Θ_E denotes the parameters of the reconstruction model;

– $H_E(t)$ denotes the history of environmental states.

The two digital twins are coupled through the following state update equation:

$$DT_E(t + \Delta t) = F_E\left(DT_E(t), \Phi\left(D(t); Q_N\left(X^N(t)\right)\right)\right),$$

where:

– $D(t)$ denotes the set of incoming measurements;

– $Q_N\left(X^N(t)\right)$ is a function that evaluates the operational state and reliability of data sources within the sensor network digital twin DT_N;

– Φ denotes the operator responsible for selecting, weighting, and fusing sensor measurements based on these evaluations;

– F_E denotes the operator that updates the state of the monitored environment.

The state of the sensor network digital twin is updated independently:

$$DT_N(t + \Delta t) = F_N\big(DT_N(t), D_N(t)\big),$$

where $D_N(t)$ denotes telemetry data collected from the sensor nodes and the network.

The state of DT_E is reconstructed from validated sensor measurements, with their processing informed by the current state of DT_N. Raw measurements are not transferred directly into the environmental digital twin. Instead, the network digital twin provides information on the operational condition and reliability of individual data sources, which is used to select and weight measurements during reconstruction of the monitored environment. The use of network state information in environmental reconstruction is a defining feature of the proposed architecture.

3. Results and discussion

3.1. Formal model of the data layers

The architecture is represented by an ordered set of data layers:

$$\mathbb{L} = \{L_0, L_1, L_2, L_3, L_4\}.$$

Each layer L_k is defined as:

$$L_k = \langle D_k, X_k, F_k, Q_k \rangle,$$

where:

– D_k denotes the set of data associated with the layer;

– X_k denotes the internal state of the layer;

– F_k denotes the operator responsible for processing data within the layer;

– Q_k denotes the set of functions used to assess the quality of the layer's state and data.

For each k the transformation between successive layers is specified by the operator:

$$T_{k,k+1}: D_k \times X_k \times Q_k \times \Theta_{k,k+1} \to D_{k+1},$$

which maps the data, internal state, quality assessment, and transformation parameters of a given layer to the input dataset of the subsequent layer.

In the general case, the interlayer transformation is expressed as

$$D_{k+1}(t) = T_{k,k+1}\big(D_k(t), X_k(t), Q_k(t), \Theta_{k,k+1}\big),$$

where $\Theta_{k,k+1}$ – denotes the parameters governing the corresponding transformation.

The upper layer incorporates not only the reconstructed data D_3, but also the network state D_2 and the digital twin's historical data $H(t)$. Consequently, the end-to-end transformation cannot be

represented as a simple composition of single-input operators. A generalized end-to-end transformation operator is therefore introduced:

$$\mathcal{T}_{0,4}: D_0 \times X^N \times H \to D_4.$$

Its expanded form consists of the following sequence of transformations:

$$\begin{aligned} D_1(t) &= T_{0,1}\big(D_0(t)\big), \\ D_2(t) &= T_{1,2}\big(D_1(t), X^N(t)\big), \\ D_3(t) &= T_{2,3}\big(D_2(t)\big), \\ D_4(t) &= T_{3,4}\big(D_3(t), D_2(t), H(t)\big). \end{aligned}$$

The resulting end-to-end relationship is expressed as:

$$D_4(t) = \mathcal{T}_{0,4}\big(D_0(t), X^N(t), H(t)\big).$$

The state of the upper layer is thus obtained through successive transformations of the raw sensor data, with the state of the sensing infrastructure and the digital twin's historical data incorporated during the final integration stage.

3.2. Admissibility conditions for interlayer transitions

For each layer, an admissible data domain Ω_k is specified. Data are passed to the next layer only if they satisfy the following condition:

$$D_k^*(t) = \{d \in D_k(t) \mid C_k\big(d, X_k(t)\big) = 1\},$$

where C_k is the admissibility predicate.

The interlayer transformation is then applied to the admissible dataset:

$$D_{k+1}(t) = T_{k,k+1}\big(D_k^*(t)\big).$$

For measurements that remain admissible but require reduced confidence, a weighting factor is introduced:

$$0 \le q_k(d,t) \le 1.$$

The corresponding transformation takes the form:

$$D_{k+1}(t) = T_{k,k+1}\big(\{\big(d, q_k(d,t)\big): d \in D_k^*(t)\}\big).$$

This formulation supports both the strict exclusion of inadmissible data and confidence-based weighting of measurements that satisfy the admissibility criteria but differ in reliability.

3.3. Layer L_0

Layer L_0 contains raw data received directly from the physical Smart Dust sensor nodes and is defined as:

$$L_0 = \langle D_0, X_0, F_0, Q_0 \rangle.$$

For each sensor node v_i the raw data record has the form:

$$d_i^{(0)}(t) = \langle id_i, t_i, p_i(t), z_i(t), e_i(t), \ell_i(t) \rangle,$$

where:

– id_i is the unique identifier of the sensor node;

– t_i is the measurement timestamp;

– $p_i(t)$ specifies the spatial coordinates of the node;

– $z_i(t)$ represents an individual measurement or a vector of measurements;

– $e_i(t)$ is the remaining energy reserve;

– $\ell_i(t)$ characterizes the communication link through an aggregate metric.

The complete dataset at this layer is given by:

$$D_0(t) = \{d_1^{(0)}(t), d_2^{(0)}(t), \dots, d_N^{(0)}(t)\}.$$

The measurement process is modeled as:

$$z_i(t) = h_i\big(s(p_i, t)\big) + \varepsilon_i(t),$$

where:

– $s(p_i, t)$ represents the true state of the environment at the sensor location;

– h_i is the sensor measurement function;

– $\varepsilon_i(t)$ accounts for measurement error.

The quality assessment function Q_0 evaluates packet completeness, timestamp validity, measurement range admissibility, and the integrity of the transmitted data.

3.4. Layer L_1

Layer L_1 contains sensor measurements that have undergone validation, temporal synchronization, and calibration correction. The layer is formally specified as:

$$L_1 = \langle D_1, X_1, F_1, Q_1 \rangle.$$

The transformation from the raw data layer is expressed as:

$$D_1(t) = T_{0,1}\big(D_0(t)\big).$$

Calibration-corrected measurements are obtained from:

$$\tilde{z}_i(t) = C_i\big(z_i(t), \theta_i^C\big),$$

where:

– C_i is the calibration correction function;

– θ_i^C represents the corresponding calibration parameters.

When normalization is required, the j-th component of the corrected measurement vector is scaled according to:

$$z_{ij}^{*}(t) = \frac{z_{ij}(t) - z_j^{\min}}{z_j^{\max} - z_j^{\min}}.$$

Measurement age is calculated as the time elapsed since data acquisition:

$$\Delta t_i = t - t_i,$$

The temporal relevance of each measurement is quantified using an exponential decay function:

$$r_i^T(t) = \exp(-\lambda_T \Delta t_i).$$

The dataset associated with layer L_1 is then expressed as follows:

$$D_1(t) = \{\tilde{d}_i(t) : C_0\left(d_i^{(0)}\right) = 1\}.$$

3.5. Layer L_2

Layer L_2 represents the operational state of the sensing infrastructure within the digital twin and is formally specified as:

$$L_2 = \langle D_2, X^N, F_N, Q_N \rangle.$$

The state of each sensor node is described by the vector:

$$x_i^N(t) = [p_i(t), e_i(t), c_i(t), \ell_i(t), k_i(t), q_i(t)]^T,$$

where:

– c_i represents the sensor calibration status;

– k_i characterizes local network connectivity;

– q_i is the local measurement quality estimate.

The network state is formed by the collection of individual node states:

$$X^N(t) = \{x_1^N(t), x_2^N(t), \dots, x_N^N(t)\}.$$

Network topology is represented by a time-varying graph:

$$G(t) = \big(V(t), E(t)\big).$$

The adjacency matrix is constructed by comparing the quality of each communication link with a predefined threshold:

$$a_{ij}(t) = \begin{cases} 1, & \ell_{ij}(t) \geq \ell_{\min}, \\ 0, & \ell_{ij}(t) < \ell_{\min}. \end{cases}$$

For each node, local connectivity is calculated as the proportion of other nodes to which it is connected:

$$k_i(t) = \frac{\sum_{j=1}^{N} a_{ij}(t)}{N-1}.$$

The normalized remaining energy reserve is calculated relative to the maximum energy capacity of the node:

$$r_i^E(t) = \frac{e_i(t)}{e_i^{\max}}.$$

Communication link quality is normalized to the interval [0, 1] using:

$$r_i^L(t) = \mathrm{clip}\left(\frac{\ell_i(t) - \ell_{\min}}{\ell_{\max} - \ell_{\min}}, 0{,}1\right).$$

The initial node reliability score is calculated from its operational characteristics and the temporal relevance of its measurements, before spatial consistency is taken into account:

$$R_i^{(0)}(t) = \beta_E r_i^E(t) + \beta_C r_i^C(t) + \beta_L r_i^L(t) + \beta_K r_i^K(t) + \beta_T r_i^T(t),$$

subject to the constraints

$$\sum_{\alpha \in \{E,C,L,K,T\}} \beta_\alpha = 1, \quad \beta_\alpha \geq 0.$$

The final reliability estimate $R_i(t)$ is obtained by evaluating the spatial consistency of the measurements, with $R_i^{(0)}(t)$ serving as the initial estimate. Calculating reliability in this sequence prevents circular dependencies between node reliability and the local measurement quality assessment.

Following this refinement, the dataset associated with layer L₂ is expressed as:

$$D_2(t) = T_{1,2}\big(D_1(t), X^N(t)\big) = \{\big(\tilde{d}_i(t), R_i(t)\big)\}.$$

If $R_i(t) < R_{\mathrm{crit}}$, the corresponding measurement is either excluded from subsequent reconstruction or assigned zero weight.

3.6. Layer L_3

Layer L_3constructs the digital representation of the monitored physical environment and is formally defined as:

$$L_3 = \langle D_3, X^E, F_E, Q_E \rangle.$$

Let $\Omega \subset \mathbb{R}^d$ be the monitoring domain. At any location $p \in \Omega$, the monitored parameter is estimated by combining sensor measurements weighted according to source reliability, spatial proximity, and measurement age:

$$\hat{s}(p,t) = \frac{\sum_{i=1}^{N} R_i(t) K_S(p, p_i) K_T(t, t_i) \tilde{z}_i(t_i)}{\sum_{i=1}^{N} R_i(t) K_S(p, p_i) K_T(t, t_i)}.$$

Spatial weighting is performed using a Gaussian kernel:

$$K_S(p, p_i) = \exp\left(-\frac{\| p - p_i \|^2}{2\sigma_S^2}\right),$$

while the temporal kernel accounts for the age of each measurement:

$$K_T(t, t_i) = \exp\big(-\lambda_T(t - t_i)\big).$$

The reconstructed environmental state is represented as a spatial field:

$$X^E(t) = \hat{S}(t) = \{\hat{s}(p, t) \mid p \in \Omega\}.$$

The transformation from the sensor network state to the reconstructed environmental state is expressed as:

$$D_3(t) = T_{2,3}\big(D_2(t)\big) = \Phi\big(D_1(t), R(t), P(t)\big),$$

where $P(t) = \{p_1(t), \dots, p_N(t)\}$ is the set of sensor node positions.

Spatial reconstruction therefore depends not only on the measured values and their proximity to the location of interest, but also on the assessed reliability of the sensor nodes supplying those measurements.

3.7. Layer L_4

Layer L_4 integrates the two digital twin models, combining the state of the sensing infrastructure with that of the monitored physical environment:

$$L_4 = \langle D_4, X^{DT}, F_{DT}, Q_{DT} \rangle.$$

The integrated state is represented by the combined state vector:

$$X^{DT}(t) = \begin{bmatrix} X^N(t) \\ X^E(t) \end{bmatrix}.$$

The complete digital twin representation incorporates network topology, the states of the sensing infrastructure and monitored environment, source reliability estimates, and historical data:

$$D_4(t) = \langle G(t), X^N(t), X^E(t), R(t), H(t) \rangle.$$

The transition to the integrated representation is defined by:

$$D_4(t) = T_{3,4}\big(D_3(t), D_2(t), H(t)\big).$$

3.8. End-to-End Model

The end-to-end data transformation is governed by the following system of equations:

$$\begin{aligned} D_1(t) &= T_{0,1}\big(D_0(t)\big), \\ D_2(t) &= T_{1,2}\big(D_1(t), X^N(t)\big), \\ D_3(t) &= T_{2,3}\big(D_2(t)\big), \\ D_4(t) &= T_{3,4}\big(D_3(t), D_2(t), H(t)\big). \end{aligned}$$

The integrated digital twin state is obtained by combining the end-to-end transformation with the historical state information:

$$X^{DT}(t) = \mathcal{F}\left(\mathcal{T}_{0,4}\left(D_0(t), X^N(t), H(t)\right), H(t)\right).$$

The coupling between the sensing infrastructure and the environmental digital twin is captured by:

$$X^E(t + \Delta t) = F_E\left(X^E(t), D_1(t), Q_N\left(X^N(t)\right)\right).$$

The reconstructed state of the monitored environment therefore depends on both the sensor measurements and the operational state of the sensing infrastructure. This relationship allows changes in sensor network conditions to be taken into account when updating the environmental digital twin.

3.9. Formalization of the Smart Dust network model

Consider a physical Smart Dust network comprising a set of sensor nodes:

$$V(t) = \{v_1, v_2, \dots, v_N\}.$$

Each node is characterized by a state vector:

$$x_i(t) = [p_i(t), e_i(t), c_i(t), \ell_i(t), k_i(t), q_i(t)]^T.$$

The state of the entire network is represented by the collection of individual node states:

$$X^N(t) = \{x_1(t), x_2(t), \dots, x_N(t)\}.$$

The evolution of each node's state is governed by:

$$x_i(t + \Delta t) = f_i\left(x_i(t), u_i(t), \eta_i(t)\right),$$

where $u_i(t)$ represents the control input applied to the node, while $\eta_i(t)$ accounts for external disturbances and uncertainties.

At the network level, the collective state dynamics are described by:

$$X^N(t + \Delta t) = F_N\left(X^N(t), U_N(t), \eta_N(t)\right).$$

3.10. Measurement reliability and spatial consistency model

The overall reliability score $R_i(t)$ accounts for both the operational condition of a sensor node and the consistency of its measurements with those obtained from nearby nodes.

The spatial neighborhood of node *i* is defined as:

$$N_i(t) = \{j: \| p_j(t) - p_i(t) \| \leq r_N\}.$$

The expected local measurement is estimated from neighboring sensor readings, weighted by their initial reliability scores:

$$\overline{z}_i(t) = \frac{\sum_{j\in N_i} R_j^{(0)}(t)\tilde{z}_j(t)}{\sum_{j\in N_i} R_j^{(0)}(t)}.$$

The deviation of the observed measurement from this estimate is calculated as:

$$\delta_i(t) = |\tilde{z}_i(t) - \overline{z}_i(t)|.$$

Spatial consistency is then quantified using an exponential weighting function:

$$r_i^Q(t) = \exp\left(-\frac{\delta_i^2(t)}{2\sigma_Q^2}\right).$$

The final reliability score combines the initial assessment of node condition with the spatial consistency measure:

$$R_i(t) = \left(1 - w_Q\right)R_i^{(0)}(t) + w_Q r_i^Q(t), \quad 0 \le w_Q \le 1.$$

The initial score $R_i^{(0)}(t)$ is calculated independently of neighboring nodes' final reliability scores $R_j(t)$. The spatial consistency measure $r_i^Q(t)$ is evaluated subsequently and incorporated into $R_i(t)$, thereby avoiding circular dependencies in the reliability calculation.

To distinguish genuine local physical events from apparent anomalies caused by sensor faults, a neighborhood agreement coefficient is introduced:

$$C_i(t) = \frac{1}{|N_i|}\sum_{j\in N_i} \mathbf{1}\left(\left|\tilde{z}_j(t) - \tilde{z}_i(t)\right| < \varepsilon\right).$$

If $C_i(t) \ge C_{\min}$, the observed local variation is considered to be corroborated by multiple sensor nodes. In this case, the spatial consistency score r_i^Q is not reduced solely because the measurements deviate from the long-term mean.

3.11. Energy consumption model for sensor nodes

Limited energy availability is a major constraint on the operation of Smart Dust networks. For sensor node i the remaining energy after one operating cycle is calculated as:

$$e_i(t + \Delta t) = e_i(t) - E_i^{sens} - E_i^{proc} - E_i^{tx} - E_i^{rx}.$$

The energy consumed during sensing operations is given by:

$$E_i^{sens} = n_i^{sens}\varepsilon_i^{sens}.$$

For the transmission of a b_i-bit packet over a distance d_i, the energy requirement is modeled as:

$$E_i^{tx} = b_i\left(\varepsilon_{el} + \varepsilon_{amp}d_i^{\alpha}\right).$$

The energy required to receive a packet is calculated as:

$$E_i^{rx} = b_i\varepsilon_{rx}.$$

The remaining energy reserve contributes directly to the node reliability assessment through the normalized factor $r_i^E(t)$. It also informs the selection of sensing and communication operating modes, allowing energy availability to be considered alongside measurement reliability when managing the sensor network.

3.12. Spatiotemporal data fusion

Each sensor measurement is assigned a composite weight that accounts for source reliability, spatial proximity, and temporal relevance:

$$W_i(p,t) = R_i(t)K_S(p,p_i)K_T(t,t_i).$$

The environmental state at location p is then estimated as a weighted combination of the available measurements:

$$\hat{s}(p,t) = \frac{\sum_i W_i(p,t)\tilde{z}_i(t_i)}{\sum_i W_i(p,t)}.$$

For a vector-valued environmental state $s(p,t) = [s_1(p,t), \dots, s_m(p,t)]^T$ reconstruction can be performed either component wise or jointly, with the latter approach accounting for covariance between the measured parameters.

Let $\Sigma_i(t)$ be the measurement error covariance matrix associated with sensor node i. For multivariate data fusion, the corresponding weight matrix is defined as:

$$W_i(t) = R_i(t)\Sigma_i^{-1}(t).$$

The fused state estimate is obtained from:

$$\hat{s}(t) = \left(\sum_i W_i(t)\right)^{-1}\left(\sum_i W_i(t)\tilde{z}_i(t)\right).$$

This formulation accounts for differences in measurement uncertainty across sensor types, allowing more precise measurements to contribute more strongly to the reconstructed state.

3.13. Dynamic model of the digital twin

The combined system state at discrete time step k is represented by the vector:

$$X_k = \begin{bmatrix} X_k^N \\ X_k^E \end{bmatrix}.$$

The evolution of the sensor network state is governed by:

$$X_{k+1}^N = F_N(X_k^N, U_k^N, \xi_k^N).$$

The monitored environment evolves according to:

$$X_{k+1}^E = F_E(X_k^E, Z_k, R_k, \xi_k^E).$$

To explicitly account for the dependence of measurement reliability on the operational state of the sensor network, the environmental state update is reformulated as:

$$X_{k+1}^{E} = F_E(X_k^E, Z_k, Q_N(X_k^N), \xi_k^E).$$

This coupling distinguishes the proposed architecture from conventional models that treat all sensor measurements as equally reliable inputs to the digital representation. By incorporating network-dependent reliability estimates into the environmental state update, the model accounts for variations in the operational condition of the sensing infrastructure.

3.14. Synchronization of physical and digital representations

Let $y(t)$ be the vector of reference values obtained from the physical system, and $\hat{y}(t)$ the corresponding values generated by the digital twin. The synchronization error is calculated as:

$$e_{DT}(t) = y(t) - \hat{y}(t).$$

The overall discrepancy between the physical and digital representations is quantified using the root mean square error:

$$E_{DT}(t) = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(y_j(t) - \hat{y}_j(t)\right)^2}.$$

If $E_{DT}(t) > E_{\max}$, then a diagnostic procedure is initiated to identify the source of the discrepancy. The diagnostic function is specified as:

$$\Gamma(t) = \Gamma\left(E_{DT}(t), R(t), G(t), X^N(t)\right).$$

The diagnostic procedure classifies the discrepancy according to its likely cause: a change in the monitored environment, sensor failure or calibration drift, loss of network connectivity, outdated measurements, or insufficient spatial sensor coverage.

3.15. Adaptive network control

The digital twin state is used to adjust the sensing and data transmission schedules of individual nodes. Let $\tau_i(t)$ be the sampling interval of node i. The rate of local variation in the measured parameter is estimated as:

$$g_i(t) = \frac{|\tilde{z}_i(t) - \tilde{z}_i(t - \Delta t)|}{\Delta t}.$$

The node activity score is calculated from the rate of measurement variation, sensor reliability, and remaining energy reserves:

$$A_i(t) = \alpha\frac{g_i(t)}{g_{\max}} + \beta R_i(t) + \gamma\frac{e_i(t)}{e_i^{\max}}, \text{ where } \alpha + \beta + \gamma = 1.$$

The sampling interval is then adjusted according to:

$$\tau_i(t) = \tau_{\max} - (\tau_{\max} - \tau_{\min})\text{clip}(A_i(t), 0,1).$$

Nodes with sufficient energy reserves and high measurement reliability are assigned shorter sampling intervals when local conditions change rapidly. In areas where the monitored parameters remain relatively stable, longer intervals reduce energy consumption and network traffic.

The overall network control policy is expressed as:

$$U_k = \Psi\left(X_k^N, X_k^E, E_{DT,k}, H_k\right).$$

3.16. Operational algorithm of the proposed architecture

At each iteration, active sensor nodes generate data packets containing measurements and operational telemetry. The incoming data are validated and normalized before being used to update the sensor network digital twin. The operational state of each node is then assessed to determine source reliability and assign appropriate weights to the measurements. These reliability-weighted data are used to reconstruct the spatiotemporal state of the monitored physical system. The outputs of the two digital twin models are combined into an integrated representation, which is subsequently compared with the physical system to assess synchronization accuracy. Based on this assessment, control actions are determined to adjust the sensing and communication parameters for the next iteration.

4. Conclusion

This study presents a digital twin architecture for a distributed Smart Dust sensor network that accounts for both the operational state of the sensing infrastructure and the state of the monitored physical system. Rather than relying solely on incoming sensor measurements to construct the digital representation, the proposed architecture also incorporates information about the condition and reliability of the data sources.

The architecture comprises two interconnected digital twin models, one representing the sensor network and the other describing the monitored physical system. A five-layer data model specifies how information is transformed from raw sensor measurements into an integrated digital twin state. The framework formalizes data validation and normalization, source reliability assessment, spatiotemporal measurement fusion, and the generation of control actions.

By incorporating the current condition of the sensing infrastructure, the architecture accounts for the effects of network performance on the quality of the digital representation. It also provides a means of distinguishing actual changes in the monitored physical system from apparent changes

caused by sensor failures or network degradation. This establishes a basis for improving the resilience of the digital twin to partial data loss and variations in individual node performance.

The quality of the resulting digital representation nevertheless depends on the spatial coverage of the sensor network, the accuracy of sensor calibration, and the reliability of node state estimates. A substantial reduction in the number of available data sources, disruptions in network connectivity, or accumulated errors in operational telemetry may compromise the accuracy of the reconstructed physical system state. Maintaining a separate digital twin of the sensing infrastructure also introduces additional computational and communication overhead associated with data processing and reliability assessment.

Future work will focus on experimental validation of the proposed architecture, the development of adaptive methods for assessing sensor data reliability, and a quantitative investigation of how Smart Dust network degradation affects the accuracy of the resulting digital representation.

Acknowledgments

This research was supported by the Russian Science Foundation (Grant No. 25-11-20040, https://rscf.ru/project/25-11-20040/) and the Krasnoyarsk Regional Fund of Science and Technology Support.

References

1. Warneke, B. Smart Dust: Communicating with a Cubic-Millimeter Computer / B. Warneke, M. Last, B. Liebowitz, K. S. J. Pister // Computer. – 2001. – Vol. 34, No. 1. – P. 44–51. – https://doi.org/10.1109/2.895117.
2. The concept of creation of information system for environmental monitoring based on modern GIS-technologies and earth remote sensing data / Y. P. Yuronen, E. A. Yuronen, V. V. Ivanov [et al.] // IOP Conference Series: Materials Science and Engineering. – 2015. – Vol. 94. – Article 012023. – https://doi.org/10.1088/1757-899X/94/1/012023.
3. Zhou, W. Research on Wireless Sensor Network Access Control and Load Balancing in the Industrial Digital Twin Scenario / W. Zhou // Journal of Sensors. – 2022. – Vol. 2022. – Article 3929958. – https://doi.org/10.1155/2022/3929958.
4. Jesus, G. A Survey on Data Quality for Dependable Monitoring in Wireless Sensor Networks / G. Jesus, A. Casimiro, A. Oliveira // Sensors. – 2017. – Vol. 17, No. 9. – Article 2010. – https://doi.org/10.3390/s17092010.

5. Getmanskaya, D. V. Research of software subsystems for the tasks of sensory monitoring and ecosystem management / D. V. Getmanskaya // Informatics. Economics. Management. – 2025. – Vol. 4, No. 2. – P. 3032-3037. – https://doi.org/10.47813/2782-5280-2025-4-2-3032-3037.
6. Privacy and Integrity-Preserving Data Aggregation Scheme for Wireless Sensor Networks Digital Twins / Z. Zhang, W. Yang, F. Wu [et al.] // Journal of Cloud Computing. – 2023. – Vol. 12. – Article 140. – https://doi.org/10.1186/s13677-023-00522-7.
7. An Implementation of Sensor Network to Assess Near Real-Time Digital Twin Fidelity / T. A. Jauhar, S. Han, T. H. Khan [et al.] // Cyber-Physical Systems. – 2026. – Vol. 12, No. 3. – P. 366–389. – https://doi.org/10.1080/23335777.2026.2613415.
8. Multiversion environment creation for control algorithm execution by autonomous unmanned objects / I. V. Kovalev, P. V. Zelenkov, V. V. Losev [et al.] // IOP Conference Series: Materials Science and Engineering. – 2017. – Vol. 173. – Article 012025. – https://doi.org/10.1088/1757-899X/173/1/012025.
9. Pyatkov, D. A. Digital platform of the railway transport track sector / D. A. Pyatkov // Modern Innovations, Systems and Technologies. – 2026. – Vol. 6, No. 3. – P. 1017–1040. – https://doi.org/10.47813/2782-2818-2026-6-3-1017-1040.