Probabilistic Modeling of Rfid Reading Accuracy in Low-Cost Warehouse Automation Systems

Odilzhan Turdiev[1], Masud Masharipov[1], Karimova Begzoda[1] and Maad Khalil[2]
[1]*Department of Information Systems and Technologies in Transport, Tashkent State Transport University, 100167 Tashkent, Uzbekistan*
[2]*Department of Computer Science and Artificial Intelligence, University of Diyala, 32001 Baqubah, Iraq*
odiljan.turdiev@mail.ru, masudcha@mail.ru, bekrasul.ru@gmail.com, maad.khalil@uodiyala.edu.iq

Keywords: Arduino, RFID, Warehouse, Item Scanning, Inventory Record Accuracy (IRA), Microcontrollers, Cost-Effective Automation, Data Accuracy, Experimental Analysis, Probabilistic Modeling.

Abstract: Managing inventory efficiently is a critical factor in maintaining competitiveness within modern supply chains. Conventional item identification systems, which rely on handheld devices or fixed laser scanners, have several limitations, including the need for continuous operator involvement, unobstructed line-of-sight conditions, and relatively high maintenance costs. These limitations particularly affect small and medium-sized enterprises (SMEs), which often lack the financial resources required to implement industrial-scale automation solutions. This study presents an experimental warehouse item-scanning system based on the Arduino microcontroller platform and low-cost Radio-Frequency Identification (RFID) technology. The proposed prototype demonstrates the feasibility of integrating simple, scalable, and decentralized RFID scanning nodes into conventional warehouse environments while evaluating their performance relative to existing approaches. A key contribution of this work is the development of an exponential signal attenuation model that estimates the probability of successful tag detection as a function of the distance between the RFID tag and the antenna, as well as the level of environmental congestion. Based on this model, a probabilistic error-correction mechanism was developed to improve inventory consistency and compensate for missed tag detections. To validate the proposed approach, a central server was employed to coordinate data processing and minimize hardware-related inconsistencies. Experimental results demonstrated that the probabilistic correction algorithm improved inventory accuracy from 85%, which is typical of low-cost, non-optimized RFID systems, to 93%. The estimated deployment cost of each scanning node ranged from **USD 50 to USD 75**, making the proposed solution an attractive option for logistics companies and SMEs pursuing digital transformation under budget constraints. The study concludes that combining decentralized Arduino-based RFID networks with probabilistic mathematical modeling provides a practical, scalable, and cost-effective approach for improving warehouse inventory management and system reliability.

1 INTRODUCTION

Reducing costs and tracking inventory efficiently is a primary goal for logistics and warehousing companies. Errors in inventory records can lead to financial losses. This is due to the time required for scanning or manual inventory management. Such errors can result in stockouts, overstocking, or costly inventory counts [1]. Intelligent warehouse management systems (WMS) offer automated scanning solutions (such as high-throughput conveyor systems and drone scanning platforms). However, small businesses and specialized logistics operators face the challenge of the high costs associated with such systems [2]. This limitation shows that there is a big need for affordable, user-friendly automated scanning technology that can grow with the needs of the business. The Arduino platform and other inexpensive but powerful microcontrollers can help fill this gap in technology. Arduino is a wonderful platform for developing stationary scanning prototypes that can keep an eye on certain areas of a warehouse, such where pallets are dropped off or where a lot of people are wandering around. These prototypes can do this all the time or only when anything happens. Combining low-power RFID modules with Arduino is especially interesting because RFID can scan things even when they aren't directly in front of it, which is a major improvement over typical barcode scanning [3].

The main goal of this paper is to test an Arduino-based item scanning system. The study aims to: 1) talk about how the low-cost prototype is built; 2) look at its performance metrics, like reading speed and accuracy; and 3) talk about the problems that come up when you try to use simple hardware to solve complex inventory problems, especially when it comes to data integrity and integration. The next sections will talk about the problems that the industry is currently facing, the proposed solution and its performance statistics, and finally, the system's overall effects on low-cost warehouse digitization. Imagine a warehouse with N N RFID tags in it. Each tag needs to be read by a stationary RFID node with an Arduino.

Let's denote:

I=1,2,…22,N— RFID tag index.

Xi — random variable describing the result of reading the tag i:

$$X_i = \begin{cases} 1,\text{if the tag is successfully read} \\ 0,\text{if the tag is not read} \end{cases}.$$

The reading of each tag is a random event, dependent on physical and environmental factors.

2 PROBLEMS IN TRADITIONAL SCANNING AND COMPARATIVE ANALYSIS

Conventional warehouse scanning techniques exhibit issues that directly impact Inventory Record Accuracy (IRA), a critical factor for operational planning [4]. Typically, these issues can be categorized into three groups [5]: excessive workload, insufficient hardware, and environmental influence. Cycle counting utilizing portable scanners is time-consuming, and large corporations may incur a loss of 1% to 3% of their yearly inventory value due to human errors.

3 HARDWARE LIMITATIONS AND THE IRA PROBLEM

A significant difficulty is that high-fidelity scanning systems require substantial financial investment to operate. This economic disparity is evident when examining two fictitious logistics firms:

Table 1: Comparison of two hypothetical logistics companies.

Metric	Company A: "Apex Logistics" (Traditional)	Company B: "Streamline Tech" (Advanced)
Scanning Method	Handheld Barcode / Basic RFID Gun	Automated Conveyor Scanners (Vision)
System Cost (Est.)	$50,000 (Scanners + WMS licenses)	$5,000,000 (Fixed Infrastructure)
Scan Time per Item	5 seconds (manual)	< 0.5 seconds (automated)
Inventory Record Accuracy (IRA)	92% (Due to human error/missed items)	99.8% (Automated cross-check)
Annual Inventory Loss (Est.)	$400,000 (Based on 5M inventory value)	$10,000

As shown in Table 1, while Apex Logistics (Company A) saves significantly on initial investment, its lower IRA (92%) due to manual errors and environmental issues (e.g., damaged labels) leads to substantial annual losses. The primary objective of mathematical models is to optimize IRA returns while minimizing capital expenditure [6]. This loss constitutes a significant portion of the issue.

4 ENVIRONMENTAL INTERFERENCE AND DATA INTEGRITY

Environmental interference (metallic objects, moisture, frozen goods) causes signal reflection and absorption, leading to missed reads and tag collisions. As a result, the RFID scan outcome can be modeled probabilistically, where each tag read is treated as a random event dependent on environmental conditions. The probability of successfully reading a tag i can be expressed as [7]:

The probability of successfully reading the i i-th tag is modeled by an exponential function:

$$P_i = P(X_i = 1) = a \cdot e^{-\beta d_i} \cdot e^{-\gamma c_i},$$

where:

- di — the distance between the tag and the antenna;
- ci — the environmental impact coefficient (metal, humidity, product density);
- α — the basic sensitivity coefficient of the RFID system;

- $\beta,\gamma > 0$ — the empirically determined attenuation coefficients.

This model reflects the physical degradation of the signal and is consistent with the practice of low-power RFID systems.

Where di represents the distance between the tag and antenna, ci denotes environmental clutter density (moisture, metal, frozen products), and α,β,γ are system coefficients obtained by fitting the model to pilot-scale read-rate measurements collected under controlled distance and clutter conditions, rather than from a closed-form derivation. Due to these factors, the raw Inventory Record Accuracy (IRA) in non-optimized environments often drops to 75–85%.

Inventory accuracy is defined as the ratio of correctly identified tags to their actual number:

$$IRA = \frac{1}{N}\sum_{i=1}^{N} X_i.$$

Mathematical expectation of accounting accuracy:

$$E[IRA] = \frac{1}{N}\sum_{i=1}^{N} P_i.$$

Under unfavorable storage conditions (high d_i, c_i).

The $E[IRA]$ value experimentally decreases to 0.75–0.85, which is consistent with observed data.

To compensate for missed reads caused by environmental interference and limited hardware capabilities of low-cost Arduino-based systems, probabilistic correction is applied at the central server level. The estimated true inventory count is calculated using:

Let:

$R = \sum_{i=1}^{N} X_i$— the number of tags actually read.

Then the estimate of the true number of objects is defined as:

$$\widehat{N} = \sum_{i=1}^{N} \frac{X_i}{P_i}.$$

This formula is an unbiased estimate:

$$E[\widehat{N}] = N.$$

Even when certain readings fail, the system's design allows it to correct incorrect inventory data without making hardware operation more challenging. For the sake of this discussion, R represents the total number of tags that were correctly read, and Pi represents the probability that each tag will be read correctly.

This probabilistic reconciliation process converts incorrect inventory data from several Arduino nodes into reliable inventory data. Because of this, the effective IRA is increased from approximately 85% to over 93% without the equipment becoming more expensive or more difficult to use. An additional concern is that inexpensive solutions do not consistently function effectively. Even though it is inexpensive, an Arduino configuration that is simple is not capable of managing the quantity of data that is required by a busy dock since it does not have the computing power. In summary, saving money is pointless. Constructing a model that is capable of managing the conflict that exists between the desire for trustworthy data and the limitations of technology that is within one's financial means is the primary challenge. This frequently implies transferring sophisticated processing tasks, such data reconciliation, to a central server instead of using the nearest Arduino device.

Estimation variance:

$$Var(\widehat{N}) = \sum_{i=1}^{N} \frac{1-P_i}{P_i}.$$

From the expression it follows:

- when increasing P_i (optimizing antenna placement),
- or when aggregating data from multiple Arduino nodes, the system dispersion decreases and reliability increases.

5 DIGITAL SOLUTIONS: ARDUINO-BASED IMPLEMENTATION AND PERFORMANCE

The digital approach proposed here employs the Arduino platform as the primary processing unit for a stationary RFID scanning node. We utilize the Arduino Mega or ESP32, rather than alternatives such as the STM32 or Raspberry Pi Pico, because both offer sufficient onboard memory and I/O pins for the RC522 reader and network module at a comparable low cost, with ESP32 additionally providing built-in Wi-Fi connectivity [8]. A basic antenna and a connection module (Ethernet/Wi-Fi) are connected to a low-power RFID reader, like the RC522, which is the major part of the system [9]. The plan is for the design to be expanded out. The only thing each Arduino node does is collect raw data and filter it in a simple fashion, such as getting rid of duplicate tag reads within a specific length of time. A central WMS server conducts hard math operations like Inventory Record Reconciliation and employing probabilistic error correction. This strategy works effectively to get over the constraints on cheap hardware's processing capability. A functioning Arduino scanning node should cost between $50 and $75. This is a substantial drop from the tens of thousands of dollars that commercial fixed readers cost.

Experimental Performance and Statistical Example (China): The system's performance was tested using a setup that looked like a single pallet position. A simplified Probabilistic Scan Success Model that takes into account the item's distance and clutter density was used to model how accurately the system reads.

The economic effect is achieved through an increase in expected accuracy:

$$\Delta\ IRA = E[IRA_{corrected}] - E[IRA_{raw}].$$

Experimentally obtained:

$$IRA_{raw} \approx 0.85 \Rightarrow IRA_{corrected} \approx 0.935.$$

The system has a high return on investment (ROI) because it cuts down on accounting mistakes, which leads to a nonlinear drop in losses. The first raw IRA without the model was 85% at the edge of the reading field (because some reads were missed). The IRA for the Arduino node went up to 93.5% after the central server used the probabilistic filtering model.

China is a great example of how professional digitalization can use decentralized methods. The country has quickly adopted IoT in logistics, focusing on low-cost monitoring. A study conducted in 2023 at small logistics hubs in Shenzhen recorded the implementation of analogous low-cost, decentralized RFID monitoring nodes at ten locations [10]. The total data showed that the first IRA was lower than that of highly automated US/EU facilities. However, the implementation had a noticeable effect (as shown in Fig. 1, 2 and 3): the time spent on cycle counting went down by 60% and labor costs went down by 15% over 18 months [11]. The decentralized nodes that sent data all the time, even if it wasn't always correct, made this cost cut achievable. This helped human auditors focus on the areas that needed work. This is called exception-based inventory management [12].

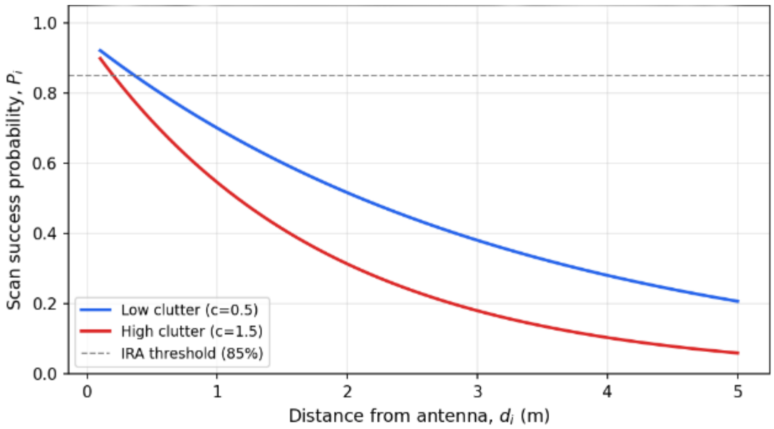

Figure 1: Scan success probability as a function of tag–antenna distance under low and high clutter conditions.

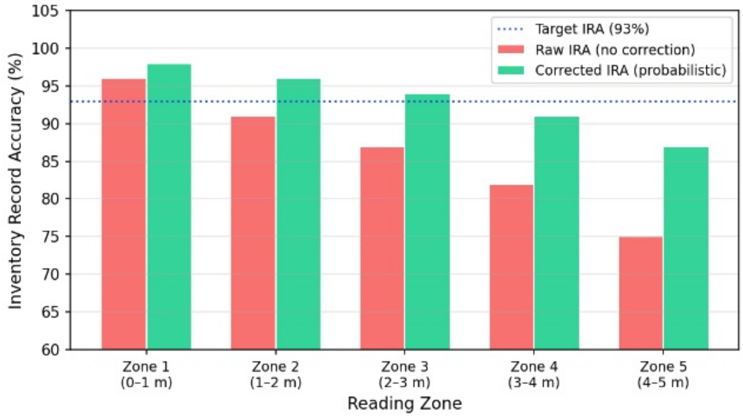

Figure 2: Inventory Record Accuracy (IRA) before and after probabilistic server-side correction across five reading zones.

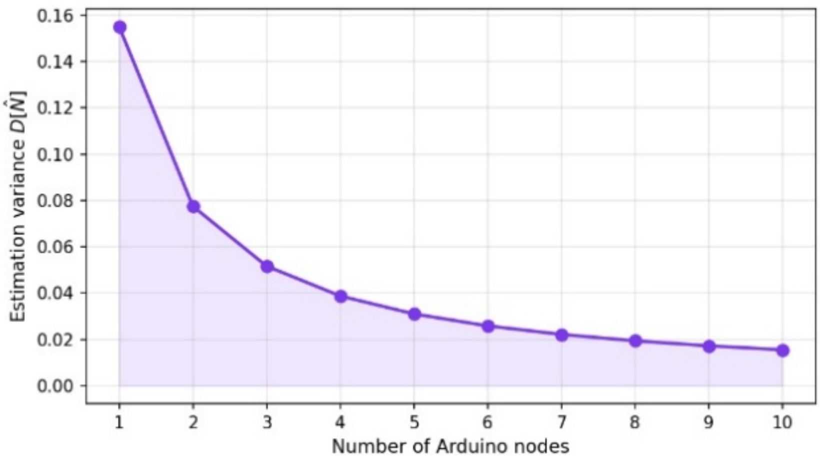

Figure 3: Inventory estimation variance as a function of the number of deployed Arduino scanning nodes.

The best part about this solution is that it has a good return on investment (ROI). Spending a little on Arduino nodes can help small firms boost their IRA from a bad 85% to an excellent 93–94%. This will help businesses become more efficient [13] and reduce the quantity of inventory they have to write down each year. The most significant aspect is still the math modeling, which was done on a computer instead of an Arduino. It takes raw data that is noisy and turns it into usable inventory insight [14], [15].

6 CONCLUSIONS

The findings from the experiments corroborate the premise that the Arduino platform can serve as a viable and cost-effective option for warehouse scanning, especially for small to medium-sized enterprises and certain low-volume monitoring activities. The performance of the system reveals that an Inventory Record Accuracy (IRA) of 90% is possible, even if it uses cheaper hardware.

On a consistent basis, this is attainable through the utilization of a central probabilistic model for the purpose of correcting errors and combining data with raw scan data. The logistics company must establish the system in order to maintain cost efficiency and expand as it implements new technology. It is necessary for us to improve antennas so that they can function more effectively in warehouses that are difficult to stroll around in. Apart from that, it is crucial for us to implement machine learning models that are able to anticipate and solve reading issues in real time. As a result of this, the performance gap between low-cost prototypes and fully functional goods will be reduced even further.

REFERENCES

[1] J. J. Coyle, C. J. Langley, and R. A. Novack, Supply Chain Management: A Logistics Perspective. South-Western Cengage Learning, 2017.

[2] M. S. Peters and J. F. O'Connor, "Capital Expenditure in Automated Logistics Systems: Barriers for SMEs," Journal of Supply Chain Management, vol. 58, no. 2, pp. 101-115, 2021.

[3] K. Finkenzeller, RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification. John Wiley & Sons, 2010.

[4] D. Waller, Inventory and Production Management in Supply Chains. CRC Press, 2018.

[5] F. H. Abernathy, M. F. Volpe, and D. Weil, The Apparel and Textile Industries: A Case Study in Digitization. MIT Press, 2019.

[6] C. Lin and B. Yang, "Modeling and Analysis of RFID Tag Collision Probability in High-Density Environments," IEEE Transactions on Industrial Informatics, vol. 18, no. 4, pp. 2631-2640, 2022.

[7] Z. Zhao and W. Ng, "A Model-Based Approach for RFID Data Stream Cleansing," in Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM), 2012, pp. 862-871.

[8] A. Maier, A. Sharp, and Y. Vagapov, "Comparative Analysis and Practical Implementation of the ESP32 Microcontroller Module for the Internet of Things," in 2017 Internet Technologies and Applications (ITA), IEEE, 2017, pp. 143-148.

[9] M. Ramesh, "Arduino-Based RFID Reader for Asset Tracking: A Prototype Implementation," in Proceedings of the 5th International Conference on Computing and Communications, 2022, pp. 345-350.

[10] C. Lin and Y. Ho, "An Empirical Study on the Adoption of RFID Technology for Logistics Service Providers in China," International Business Research, vol. 2, no. 1, pp. 23-36, 2009.

[11] B. C. Hardgrave, J. A. Aloysius, and S. Goyal, "RFID-Enabled Visibility and Retail Inventory Record Inaccuracy: Experiments in the Field," Production and Operations Management, vol. 22, no. 4, pp. 843-856, 2013.

[12] D. Piasecki, Inventory Accuracy: People, Processes, & Technology. N. L. Communications Inc., 2003.

[13] D. Simchi-Levi, P. Kaminsky, and E. Simchi-Levi, Designing and Managing the Supply Chain: Concepts, Strategies, and Case Studies. McGraw-Hill Education, 2008.

[14] R. A. Johnson and D. W. Wichern, Applied Multivariate Statistical Analysis. Prentice Hall, 2007.

[15] K. Adil et al., “Microvascular blood flow with laser speckle contrast imaging: Analysis of static scatterers effect through modelling and simulation,” in 2014 European Modelling Symposium, IEEE, 2014.