Machine Learning Enabled Food Contamination Detection Using RFID and Internet of Things System

Abubakar Sharif 1,2,*, Qammer H. Abbasi 1, Kamran Arshad 3, Shuja Ansari 1, Muhammad Zulfiqar Ali 1, Jaspreet Kaur 1, Hasan T. Abbas 1 and Muhammad Ali Imran 1,*

1 James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, UK; sharifuestc@gmail.com (A.S.); qammer.abbasi@glasgow.ac.uk (Q.H.A.); Shuja.Ansari@glasgow.ac.uk (S.A.); muhammad.ali@glasgow.ac.uk (M.Z.A.); 2585552K@student.gla.ac.uk (J.K.); Hasan.Abbas@glasgow.ac.uk (H.T.A.)
2 School of Electronic Science and Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China
3 College of Engineering and IT, Ajman University, Ajman AE 346, United Arab Emirates; k.arshad@ajman.ac.ae
* Correspondence: muhammad.imran@glasgow.ac.uk

Abstract: This paper presents an approach based on radio frequency identification (RFID) and machine learning for contamination sensing of food items and drinks such as soft drinks, alcohol, baby formula milk, etc. We employ sticker-type inkjet printed ultra-high-frequency (UHF) RFID tags for contamination sensing experimentation. The RFID tag antenna was mounted on pure as well as contaminated food products with known contaminant quantity. The received signal strength indicator (RSSI), as well as the phase of the backscattered signal from the RFID tag mounted on the food item, are measured using the Tagformance Pro setup. We used a machine-learning algorithm XGBoost for further training of the model and improving the accuracy of sensing, which is about 90%. Therefore, this research study paves a way for ubiquitous contamination/content sensing using RFID and machine learning technologies that can enlighten their users about the health concerns and safety of their food.

Keywords: ultra-high-frequency (UHF); radio frequency identification (RFID); Internet of Things (IoT); machine learning; food contamination sensing

1. Introduction

The Internet of Things (IoT) and machine learning (ML) are reshaping our lives by providing numerous emerging applications ranging from healthcare, smart environments, smart sensing, etc. [1–7]. Moreover, short-range IoT technologies such as RFID are considered to be last-mile solutions in many applications such as inventory management, supply chain tracking, healthcare, waste management, and so forth [8–14]. The UHF RFID technology provides sensing benefits due to its inherent capability of noticing impedance variations with respect to the permittivity of background environments [15–19]. Moreover, the passive UHF RFID tag also provides a relatively long read range as compared to other competitors such as low frequency (LF) RFID and high frequency (HF) RFID. Additionally, the passive UHF RFID tags pose easily printable sticker-type structures, which helps their low-cost and bulk manufacturing [20,21].

Food contamination is one of the biggest issues among public health problems. Moreover, the spoilage and deterioration of food quality during storage is another challenge for both the food industry and environmental perspectives [22,23]. According to the world health organization (WHO) fact sheet, every year almost 600 million people fall ill after eating contaminated food. Similarly, almost 0.42 million people die after eating contaminated food [24]. An RFID sensor was proposed for detecting the quality of food [25]. The quality
and contamination of food were detected by measuring the read range due to variation in permittivity of background food packets. However, this technique requires a pair of tags to be mounted at a fixed distance. A remote patient monitoring system has been proposed in [26] using RFID and ML for early detection of suicidal behavior in mental health facilities. A range of machine learning algorithms was tested and found that the Decision tree algorithm provides a better result as compared to random forest and XGBoost in this scenario. In [27], RFID and ML based techniques were used to detect the human presence and daily human activities. The proposed algorithms successfully demonstrated the accuracy of 96.7% in recognizing 24 different daily activities. The dielectric properties of organic aqueous liquids were tested with a range of different permittivity values. The label-type RFID tag antenna was mounted to either a clear borosilicate glass bottle or Petri plate. Different solutions were tested solutions consist of high-relative permittivity (such as water) along with low permittivity, lossy liquids (such as xylene) having distinctive frequency characteristics with a read range of up to 7 m for each type of container. The proposed sensor was also able to detect ‘unknown’ solutions and determine the dielectric properties by utilizing standard curve analysis with an accuracy of ±0.834 relative permittivity and ±0.050 S · m⁻¹ conductivity.

In [28], a human activity recognition system was proposed by combining passive RFID tags and a machine learning algorithm. A passive UHF RFID tag-based wall was designed for the activity recognition experiment. Moreover, a machine learning algorithm was implemented using a multivariate Gaussian algorithm for classification and prediction of sampled activities. In this scenario, the multivariate Gaussian algorithm achieved better performance in terms of accuracy as compared with standard algorithms such as random forest, logistic regression, and support vector machine (SVM) classifiers.

An idea regarding food quality sensing using RFID tags is presented in [29]. The authors used RFIDds and USRP N210 software radios for food content sensing. This set uses two frequency excitation techniques for food quality sensing. The first frequency was utilized for delivering the power in industrial, scientific, and medical (ISM) bands. The second frequency was aimed to record the changes in RFID tag’s response mounted on liquid over a wideband (due to dielectric effects). Moreover, an algorithm was implemented in MATLAB for averaging 50 RFID responses for extracting amplitude and phase. In addition to this, another XGBoost algorithm was implemented in python for gradient boosting tree classifiers. This experiment was tested for alcohol tainting and baby formula adulteration with an accuracy of 96%. Although, this experiment provides good accuracy with a difference of 25% approximately 10 g’s addition each time in sample. Therefore, the sample having in between values was not tested. Additionally, this setup is very expensive and can be used for a commercial solution.

Therefore, this paper provides a simple approach that only requires a small handheld RFID reader for measuring backscatter power from tagged food samples in terms of RSSI. The proposed technique employs sticker-type inkjet printed RFID tags and a machine learning algorithm for food contamination sensing and accuracy improvements. The received signal strength indicator (RSSI), as well as phase of the backscattered signal from RFID tag mounted on a food item, are measured using Tagformance Pro setup. The normal spring water was taken as a food sample. A known amount of salt and sugar quantity was deliberately added to water and mixed evenly. The food contamination/contents were sensed with an accuracy of 90%. We used the XGBoost algorithm for further training of the model and improving the accuracy of sensing, which is about 90%. Therefore, this research study paves a way for ubiquitous contamination sensing using RFID and machine learning technologies that can enlighten their users about the health concerns and safety of their food.

2. Proposed Methodology for Sensing Contamination

Figure 1 shows the proposed system for food contamination detection using UHF RFID tags and machine learning. For food contamination sensing proposes, the RFID
reader is placed at a fixed distance ‘R’ from the food item to be sensed. A UHF RFID tag antenna is mounted on each food item such as designed in [30]. The backscattered power from pure food items and contaminated food items will be compared and the data would be given as input to the machine learning algorithm. The machine learning algorithm trains its self and improves food contamination sensing.

Figure 1. Proposed system for food contamination sensing using RFID and machine learning.

Figure 2 illustrates the methodology for food contamination sensing using UHF RFID tags. Let "c" represents the quantity of substance added as a contaminant in a pure substance. Moreover, the known parameters of reader setup such as transmitted power \( P_{\text{transmit}} \) and reader antenna gain \( G_{\text{reader}} \) would help to calculate \( P_{\text{received}} \) by the tag antenna. Accordingly, the equations presented in [20,30,31] can be modified as follows:

\[
P_{\text{received}} = \frac{P_{\text{transmit}} G_{\text{reader}}}{4\pi R^2} \times \frac{\lambda^2}{4\pi} G_{\text{Tag}}[c] \eta_{\text{polarization}}
\]

where \( G_{\text{Tag}}[c] \) is the associated gain of tag antenna with respect to the quantity of contaminant substance contents \( c \). Moreover, \( \eta_{\text{polarization}} \) represents a polarization mismatch between the tag and reader antenna, which will be equal to 1 in our case as both tag and reader antenna are aligned.

Figure 2. Methodology for food contamination sensing using UHF RFID technology.

The power extracted by RFID chip from tag antenna can be expressed as follow:

\[
P_{r, \text{chip}} = \frac{P_{\text{transmit}} G_{\text{reader}}}{4\pi R^2} \times \frac{\lambda^2}{4\pi} G_{\text{Tag}}[c] \eta_{p} \ast \tau[c]
\]
where $\tau[c]$ measures the impedance mismatch between RFID chip and tag, also known as power transmission coefficient:

$$\tau[c] = \frac{4R_{\text{chip}}R_{\text{Tag}}[c]}{|Z_{\text{chip}} + Z_{\text{Tag}}[c]|^2} \tag{3}$$

The tag and chip impedance associated with contaminant quantity are $Z_{\text{Tag}}[c] = R_{\text{Tag}}[c] + jX_{\text{Tag}}[c]$ and $Z_{\text{chip}} = R_{\text{chip}} + jX_{\text{chip}}$, respectively.

Therefore, the backscatter power from the tag can be expressed as:

$$P_{\text{backscatter}} = P_{\text{transmit}} \frac{G_{\text{reader}}}{4\pi} \frac{\lambda^2}{4\pi} G_{\text{Tag}}[c] \eta_{\text{polarization}} \ast \frac{G_{\text{Tag}}[c]}{\Gamma_m[c]} | \ast |^2 \tag{4}$$

where $|\Gamma_m[c]|$ reflection coefficient of tag and is related to power transmission coefficient as $\tau[c] = 1 - |\Gamma_m[c]|^2$.

Accordingly, the received signal strength indicator (RSSI) extracted by the RFID reader from backscatter power is represented as:

$$P_{\text{RSSI}}[c] = \frac{1}{4\pi d^2} \frac{\lambda^2}{4\pi} G_{\text{reader}}^* G_{\text{Tag}}^2 |\Gamma_m[c]|^2 \tag{5}$$

$$P_{\text{RSSI}}[c] = \frac{1}{4\pi d^2} \frac{\lambda^2}{4\pi} G_{\text{reader}}^* G_{\text{Tag}}^2 |\Gamma_m[c]|^2 \tag{6}$$

where $rcs_{\text{Tag}}[c] = \frac{\lambda^2}{4\pi} G_{\text{Tag}}^2 |\Gamma_m[c]|^2$ can be represented as radar cross-section of the tag antenna.

$$P_{\text{RSSI}}[c] = \frac{1}{4\pi} \left( \frac{\lambda}{4\pi R} \right)^2 P_{\text{transmit}} G_{\text{reader}}^2 \frac{G_{\text{Tag}}[c]}{\eta_{\text{polarization}}^2} \tag{7}$$

The quantity of contaminant ‘$c$’ can be sensed by comparing the $P_{\text{RSSI}}[c]$ and pure food item $P_{\text{RSSI}}$. Similarly, the different quantity of contaminant ‘$c_1$’ and also be sensed by comparing $P_{\text{RSSI}}[c]$ and $P_{\text{RSSI}}[c_1]$.

3. Experimental Setup

Figure 3 shows the experimental system includes Tagformance Pro setup from voyantic company (Espoo, Finland) [32,33] and water bottles samples having a different quantity of salt and sugar. The Tagformance Pro includes a transceiver unit, a 6 dBi linearly polarized antenna, and a foam spacer. The transceiver unit was attached to a computer system with a pre-installed Tagformance software setup that helps to record different RFID tag’s performance parameters such as read range, backscatter power, and RSSI. The Tagformance setup uses Frii’s formula as described by (1) for calculating RSSI and theoretical read range. The RSSI or read range was determined for known fixed distance, which is kept fixed by a foam spacer. The water samples are placed 30 cm apart using a foam spacer and the corresponding RSSI was recorded. The tag antenna was mounted on each sample as described in the subset of Figure 3. The Tagformance Pro setup also uses a similar principle as described in the previous section for RSSI measurement.

A frequency sweep was run from 860 to 960 MHz and the corresponding RSSI associated with tag antenna mounted on a particular food sample was recorded by Tagformance Pro software. The Tagformance also uses a similar principle as described in the previous section for RSSI measurement. The 500 mL water sample packed in a PET bottle (with relative permittivity $\epsilon_r = 3.4$) was used for experiment proposes. We prepared 84 samples in total, with half of the samples for salt and half for sugar contamination detection.
4. Results and Discussion

In experimental testing, the Tagformance Pro setup and 500 mL water-filled plastic bottle were used. This experiment can be done by any RFID reader, which has RSSI collecting features. Most handheld RFID readers from other manufacturers also have similar RSSI collecting features. However, for accuracy and benchmarking, the Tagformance setup was used for recording RSSI over a complete RFID band ranging from 860 MHz to 960 MHz. We solved different quantities of salt and sugar contents separately and prepared different samples with the same quantity of contaminant for robustness. The salt contents were mixed evenly in water and each sample was tagged using RFID tag antenna [20]. Figure 4 represents the corresponding RSSI values associated with different quantities of salt added as a contaminant as well as RSSI value associated with pure water. It can be observed, the RSSI value decreases as the quantity of salt increases. This is because increasing the salt content increases the conductivity of water. Hence, the corresponding RSSI value decreases as salt content increases. The value of RSSI at 915 MHz was taken into consideration for comparison proposes. The RSSI value associated with simple water is around $-51 \text{ dBm}$. Moreover, the value of RSSI for 2, 4, 6, 8, and 10 g of salt contents were $-52$, $-53.4$, $-53.7$, $-54.5$, and $-55 \text{ dBm}$, respectively.

Similarly, the corresponding RSSI values associated with different quantities of sugar added as a contaminant as well as RSSI values associated with pure water are illustrated.
It can be observed, the RSSI value decreases as the number of sugar contents increases. This is because increasing the sugar contents produces a variation in the permittivity of water. Therefore, the corresponding RSSI value decreases as sugar content increases. The value of RSSI at 915 MHz was taken into consideration for comparison proposes. The RSSI value associated with simple water is around $-51$ dBm. Moreover, the value of RSSI for 2, 4, 6, 8, and 10 g of sugar were $-52$ dBm, $-52.25$ dBm, $-52.7$ dBm, $-53$ dBm, and $-53.5$ dBm, respectively. The change in RSSI value associated with salt contents is more obvious as compared with sugar contents.

This paper proposes a simple approach, which requires a small handheld RFID reader for measuring backscatter power from tagged food samples in terms of RSSI. The proposed technique uses sticker-like inkjet printed RFID tags for food contamination sensing. Moreover, this work includes the application of a machine learning algorithm on RFID sensors data for accuracy improvements. The received signal strength indicator (RSSI), as well as the phase of the backscattered signal from the RFID tag mounted on a food item, are measured using Tagformance Pro setup. The normal spring water was taken as a food sample. A known amount of salt and sugar quantity was deliberately added to water and mixed evenly. The food contamination/contents were sensed with an accuracy of 90%. To keep the setup commercially deployable, a handheld UHF RFID reader-based setup connected with smartphone having an android app was used for food contamination sensing as shown in Figure 6. The RFID reader has a size of 135 $\times$ 75 $\times$ 32 mm$^3$ with 10,000 mAh battery that lasts after 16 working hours. The food sample was placed 30 cm apart using the foam spacer. The RFID reader was connected to a smartphone using Bluetooth low energy (BLE) [34,35], that has preinstalled app associated with this reader setup to show the tag’s Electronic Product Code (EPC) [36] value, as well as RSSI of the tag, mounted on the food item.

The RSSI data collected using Tagformance pro setup was exploited for machine learning algorithm in order to better food contamination section accuracy. The python program was used for the implementation of XGBoost algorithm [37]. The reason behind the use of XGBoost algorithm is its scalability feature, which enables it to use less memory by utilizing distributive computing and parallel data. XGBoost uses bagging and boosting technique. For any dataset having n instances and m attributes, the explanatory variables such as sensor RSSI data can be defined as $s_i = (s_{i1}, s_{i2}, s_{i3}, \ldots, s_{im})$. We can also define objective variable as $c_i, i = 1, 2, 3, \ldots, n$. 

![Figure 5. Measured RSSI using Tagformance setup with different quantities of sugar as a contaminant.](image-url)
Now, the output of decision tree can be represented as $\hat{c}^{(t)}_i$. The error $\epsilon^{(1)}_i$ between objective function $\hat{c}^{(t)}_i$ and first decision tree output can be represented as

$$\epsilon^{(1)}_i = \hat{c}^{(1)}_i - c_i$$  \hspace{1cm} (8)

The $k$th decision tree can be described as $f_k$. So, the predicted value after $K$ time boosting can be shown as:

$$\hat{c} = \sum_{k=1}^{K} f_k(s_i)$$  \hspace{1cm} (9)

The final target is to minimize the objective function $L(\phi)$, which is based on loss function ($L(\hat{c}_i, c_i)$).

$$L(\phi) = \sum_{i=1}^{I} L(\hat{c}_i, c_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2$ is penalize function.

First of all, we prepared the samples by adding different amounts of salt concentrations. The first set was pure water without any addition of extra salt. The rest of the samples contain 2, 4, 6, 8, and 10 g of salt. For data collection, we utilized 500 mL water bottles and 7 samples were used for each salt concentration to validate the robustness of the solution. Therefore, 42 samples were used in total with salt contaminating detection. Figure 7a illustrates the results as confusion a matrix for detection of salt contamination. The different columns represent the predicted samples successfully recognized by the algorithm, while the rows represent the actual samples. It can be seen from Figure 7a, the proposed system can classify the different concentrations of salt contents with an average accuracy of 92% due to error in adjacent classes. Therefore, the proposed can be used to classify the contamination of salt contents. Similarly, we prepared the 42 samples by adding different concentration of sugar in water. The first class was pure water without adding additional sugar, while the rest of the samples contain 2, 4, 6, 8, and 10 g of sugar contamination. The result of sugar contamination detection is shown as a confusion matrix in Figure 7b. The rows and columns represent the actual and predicted sample, respectively. Therefore, the proposed system can also successfully classify different concentrations of sugar contaminants. However, the average accuracy of sugar detection is about 90% with more false detection in adjacent samples. So, the proposed system shows a potential towards the classification of the food contents as well as contamination detection.
Figure 7. Confusion matrix for detection of food contamination detection (different columns represent the predicted samples while the rows represent the actual samples). (a) Salt as contaminant; (b) sugar as contaminant.

5. Conclusions

This paper provides a simple approach that only requires a small handheld RFID reader for measuring backscatter power from tagged food samples in terms of RSSI. The proposed technique employs sticker-type inkjet printed RFID tags and a machine learning algorithm for food contamination sensing and accuracy improvements. The received signal strength indicator (RSSI), as well as the phase of the backscattered signal from the RFID tag mounted on the food item, are measured using the Tagformance Pro setup. The normal spring water was taken as a food sample. A known amount of salt and sugar quantity was deliberately added to water and mixed evenly. The food contamination/contents were sensed with an accuracy of 90%.

We used the machine learning XGBoost algorithm that was implemented in python for further training of the model and improving the accuracy of sensing, which is about 90%. Therefore, this research study paves a way for ubiquitous contamination sensing using RFID and machine learning technologies that can enlighten their users about the health concerns and safety of their food. Moreover, this research also provides sufficient information regarding food spoilage and saves a lot of food waste.

Author Contributions: Conceptualization, A.S. and Q.H.A.; methodology, A.S., Q.H.A., S.A., M.A.I. and K.A.; software, A.S., S.A. and M.Z.A.; validation, A.S., Q.H.A. and H.T.A.; formal analysis, J.K. and M.A.I.; writing—original draft preparation, A.S. and Q.H.A.; writing—review and editing, Q.H.A. and M.A.I.; supervision, M.A.I. and K.A.; project administration, M.A.I.; funding acquisition, M.A.I. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by EPSRC IAA award is EP/R511705/1.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Osifeko, M.O.; Hancke, G.P.; Abu-Mahfouz, A.M. Artificial intelligence techniques for cognitive sensing in future IoT: State-of-the-Art, potentials, and challenges. J. Senses. Actuator Netw. 2020, 9, 21. [CrossRef]
2. Sharif, A.; Ouyang, J.; Yang, F.; Chattha, H.T.; Imran, M.A.; Alomainy, A.; Abbasi, Q.H. Low-cost inkjet-printed UHF RFID tag-based system for internet of things applications using characteristic modes. IEEE Internet Things J. 2019, 6, 3962–3975. [CrossRef]
Nikitin, P.; Rao, K.V. Theory and measurement of backscattering from RFID tags. *IEEE Antennas Propag. Mag.* 2006, *48*, 212–218. [CrossRef]

Tagformance Pro | Voyantic. Available online: https://voyantic.com/products/tagformance-pro (accessed on 17 August 2021).

Sohrab, A.P.; Huang, Y.; Hussein, M.N.; Carter, P. A Hybrid UHF RFID Tag Robust to Host Material. *IEEE J. Radio Freq. Identif.* 2017, *1*, 163–169. [CrossRef]

Bluetooth Vs. Bluetooth Low Energy (LE): What’s the Difference? Available online: https://www.link-labs.com/blog/bluetooth-vs-bluetooth-low-energy (accessed on 11 October 2021).

GS1. EPC/RFID | GS1. 2015. Available online: https://www.gs1.org/standards/epc-rfid (accessed on 11 October 2021).

Intro to Bluetooth Low Energy | Bluetooth® Technology Website. Available online: https://www.bluetooth.com/bluetooth-resources/intro-to-bluetooth-low-energy/ (accessed on 11 October 2021).

XGBoost Python Package—Xgboost 1.5.0-dev Documentation. Available online: https://xgboost.readthedocs.io/en/latest/python/index.html (accessed on 17 August 2021).