ML-Aided Collision Recovery for UHF-RFID Systems

Talha Akyıldız*, Raymond Ku*, Nicholas Harder*, Najme Ebrahimi†, Hessam Mahdavifar*
* Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA
† Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32603, USA
Email: {akyildiz, rayku, nharder}@umich.edu, najme@ece.ufl.edu, hessam@umich.edu

***Abstract*—We propose a collision recovery algorithm with the aid of machine learning (ML-aided) for passive Ultra High Frequency (UHF) Radio Frequency Identification (RFID) systems. The proposed method aims at recovering the tags under collision to improve the system performance. We first estimate the number of tags from the collided signal by utilizing machine learning tools and show that the number of colliding tags can be estimated with high accuracy. Second, we employ a simple yet effective deep learning model to find the experienced channel coefficients. The proposed method allows the reader to separate each tag's signal from the received one by applying maximum likelihood decoding. We perform simulations to illustrate that the use of deep learning is highly beneficial and demonstrate that the proposed approach boosts the throughput performance of the standard framed slotted ALOHA (FSA) protocol from 0.368 to 1.756, where the receiver is equipped with a single antenna and capable of decoding up to 4 tags.**

I. Introduction

Passive Ultra High Frequency (UHF) Radio Frequency Identification (RFID) is a wireless communication system. UHF-RFID systems have a wide range of application area, e.g., logistics & supply chain, item inventory tracking and materials management to mention a few. In UHF-RFID systems, the communication is conducted between the reader and arbitrary number of passive tags with backscatter modulation. In our work, we will focus on passive UHF-RFID systems where the tags can only operate by absorbing energy from the reader through the radio waves and have a limited operation capability.

In UHF-RFID systems, the reader can only communicate with at most one tag, and whenever two or more tags try to communicate with the reader a collision occurs which results in loss of information. To avoid that, EPC Gen2 standards employ framed slotted ALOHA (FSA) protocol to randomize the transmission procedure of passive tags [1]. In FSA, frames are formed by slots and each tag chooses its slot uniformly over the frame where its size is determined by the reader and also known by the tags [2]. Even though FSA is employed as an anti-collision protocol, collisions still occur. This situation leads to the application of collision recovery algorithms with an aim of extracting or possibly recovering information from the tags under collision.

In [3], the authors study joint-decoding of multiple tags by investigating the signal constellation points and parameter estimation methods for low frequency RFID systems. Another work [4] shows that the number of tags up to four can be estimated from the collided signal, and this information later can be used to recover the signals of colliding tags. The utilization of fixed beam-forming for different sub-spaces to detect collisions which can be combined with other anti-collision algorithms is presented in [5]. In [6], it is shown that using antenna array which is coupled with blind source separation techniques can reduce the collision probability and also remove the interference signals. Another related work [7] formulates a maximum likelihood estimator to find out the number of tags in collision for each slot and also compare its performance against the other well known estimators.

The authors in [8] first propose a different type of single antenna collision recovery receivers, i.e., zero-forcing and ordered successive cancellation, to improve the performance of FSA by estimating channel coefficients for at most two colliding tags, and then, they generalize their approach for multiple antenna setting in [9]. Another related work [10] shows that it is possible to recover more than two colliding tags by exploiting the additional diversity (increasing the number of receiver antennas) with an assumption of perfect channel knowledge at the reader. Later, they extend their work by also proposing a channel estimation technique using the post-preamble symbols in [11]. In [12], the number of tag estimation is coupled with a channel estimation technique for collision recovery and [13] studies code designs to recover tag collisions given the number of colliding tags.

However, none of the above works study the application of deep learning tools for the collision recovery in UHF-RFID systems. With this motivation, we first estimate the number of tags in collision by considering Gaussian mixture models (GMM) which clusters the received signal points. We then replace GMM model with two different neural networks architectures, i.e., feed-forward and convolutional. We observe that the deep learning architectures can estimate the number of tags with high accuracy by out-performing GMM. The estimated number of tags is then utilized to estimate the channel coefficients of tags with the aid of additional symbols via deep learning. As a last step, we decode the backscattered modulated signals using minimum-distance decoding.

The remainder of this paper is organized as follows. In Section II, we present the system model for both framed slotted ALOHA and communication channel between the reader and passive tags. Section III describes the machine learning based proposed model for collision recovery. Section IV illustrates our numerical results in comparison with the other approaches. Finally, Section V concludes the paper.

II. System Model

In this section, we describe a mathematical framework for UHF-RFID systems. First, we briefly mention the properties of FSA and also compute the analytical throughput performance of FSA with collision recovery. We also represent the communication model between the reader and passive tags. Finally, we depict the constellation points of the received signal at the reader in in-phase (I) and quadrature (Q) plane (I/Q plane).

A. Framed Slotted ALOHA

EPC Gen2 standards employ FSA protocol where frames are formed by K slots and population of N tags select their slots randomly over the frame. In FSA, slots can be in three different states: 1) Empty slot: None of the tags transmit, 2) Singleton Slot: Only one specific tag transmits, 3) Collision Slot: Two or more tags try to communicate. In conventional UHF-RFID systems, the reader can only identify singleton slots and collision slots are discarded. The number of slots with R colliding tags is a random variable denoted by $\mathcal{X}_R$ where its expected value can be written as

$$\mathbb{E}\{\mathcal{X}_R\} = K\binom{N}{R}\left(\frac{1}{K}\right)^R\left(1-\frac{1}{K}\right)^{N-R}. \quad (1)$$

The throughput of FSA is defined as the number of decoded users per slot which corresponds to $\frac{\mathbb{E}\{\mathcal{X}_1\}}{K}$ and it is maximized when $K = N$ with a value of 0.368.

We consider the scenario where the receiver is capable of decoding tags when collision occurs and derive the throughput under this scenario. Let J denote the number of tags can be decoded and M denote the number of tags can be resolved by the receiver with a condition $R \leq M$. If $J = 1$, the receiver can decode one of the tags out of at most M colliding ones. In this case, the throughput can be computed as

$$\frac{1}{K}\sum_{R=1}^{M}\{\mathcal{X}_R\} = \sum_{R=1}^{M}\binom{N}{R}\left(\frac{1}{K}\right)^R\left(1-\frac{1}{K}\right)^{N-R}. \quad (2)$$

We illustrate the throughput values of FSA over different values of frame size to tag ratio (K/N) for M up to 4 in Fig. 1. It can be seen that as M increases, the throughput performance improves, and the maximum throughput can be achieved for lower frame sizes because the collisions are now recoverable. The throughput can be increased to 0.817 $(M = 4)$ which is a substantial gain compared to the throughput 0.368 of FSA.

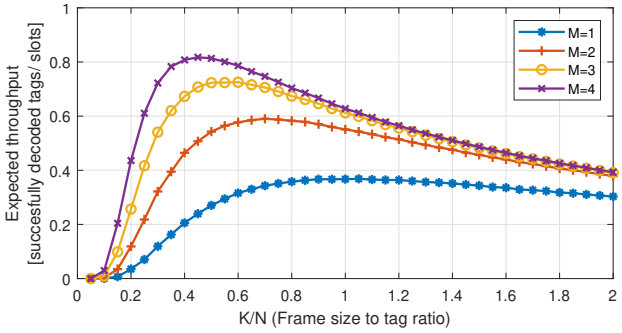

Fig. 1: Expected throughput comparison for different values of M over different values of K/N when $J = 1$.

We also consider the scenario where the receiver can decode more than one tag $(J \geq 1)$ out of up to M colliding tags. In this case, the throughput value can be written as

$$\frac{1}{K}\left(\sum_{R=1}^{J} R\{\mathcal{X}_R\} + \sum_{R=J+1}^{M} J\{\mathcal{X}_R\}\right).$$

We depict the maximum throughput values of FSA for different values of M and J in Fig. 2 by calculating the optimal frame size to tag ratio values. It can be seen that recovering more than one tag is significantly beneficial and can boost the system performance further. The throughput value can be up to nearly 2 which is 0.817 when the receiver is only capable of decoding a single tag. The throughput analysis of FSA with collision recovery shows its potential and how it can improve the performance of conventional FSA.

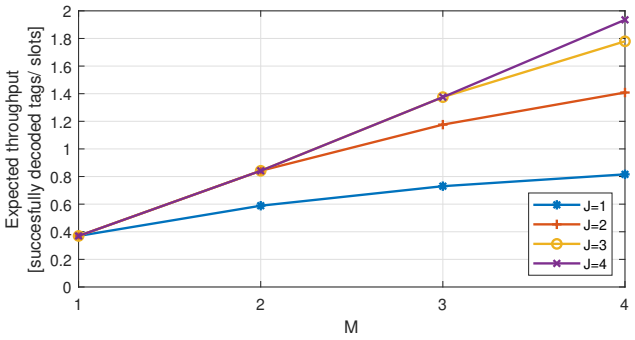

Fig. 2: Expected throughput comparison for different values of M and J with optimal frame size to tag ratio (K/N).

B. The Communication Model

We now present the communication model between the reader and passive tags as depicted in Fig. 3. In our model, we only consider a single receiver antenna at the reader. The communication can be divided into two sub-parts: 1) Forward channel: The reader supports energy and data to the passive tags with continuous carrier wave transmission, 2) Backward channel: Passive tags absorb the energy from the reader and reflects it back to the reader. Each tag modulates its 16 bit random number (RN16) with an on-off keying: 0 (OFF) and 1 (ON) corresponding the absorbing and reflecting states. Let $a_i(t)$ denote the modulated RN16 signal for tag i which is

$$a_i(t) = \sum_k a_i[k]p(t - kT_i - \tau_i), \quad (3)$$

where $a_i[k]$ denotes the transmitted symbols (± 1) and $p(t)$ is a rectangular pulse of the modulated signal. T_i and τ_i correspond to the symbol period and modulation delay, respectively.

The backscattered modulated signal for tag i is transmitted via carrier frequency f_c and denoted by $s_{tag,i}$, i.e.,

$$s_{tag,i} = |h_i^f|\sqrt{|\Delta\sigma_i|}a_i(t)\sin(2\pi f_c t + \varphi_i^f + \varphi_i^{\Delta\sigma}), \quad (4)$$

where $|h_i^f|$ is the magnitude of forward channel channel coefficient, $|\Delta\sigma_i|$ is the normalized differential radar cross section (RCS) coefficient of tag i [14]. φ_i^f and $\varphi_i^{\Delta\sigma}$ are phase shifts of the forward channel and the modulation.

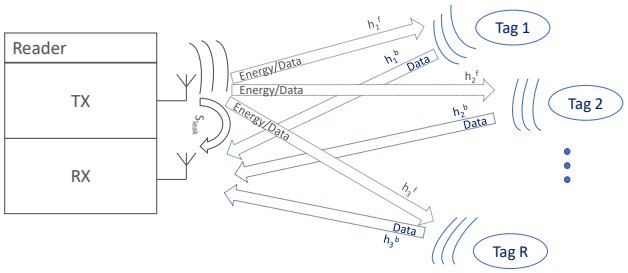

Fig. 3: The communication model between the reader and passive tags of number R.

During the transmission from the reader to tags, a leakage occurs between the transmitter and receiver antenna, i.e.,

$$s_{leak}(t) = |L| \sin(2\pi f_c t + \varphi^{leak}), \tag{5}$$

where $|L|$ is leakage magnitude and φ^{leak} is the phase shift.

After deriving all necessary components, we finally compute the received signal at the reader denoted as $s(t)$ and given by

$$s(t) = \sum_{i=1}^{R} |h_i^b||h_i^f|\sqrt{|\Delta\sigma_i|}a_i(t)\sin(2\pi f_c t + \varphi_i^f + \varphi_i^{\Delta\sigma} + \varphi_i^b) + s_{leak}(t) + n(t),$$

where R is the number of colliding tags in a certain slot. The backward channel is denoted by h_i^b and φ_i^b is the corresponding phase shift for tag i. $n(t)$ is the additive white Gaussian noise (AWGN) with power spectral density N_0 [1].

It is possible to down-convert the received signal into the baseband using I/Q demodulators since all signal components have the same carrier frequency. Utilizing that, we can write the complex-valued baseband signal at the reader as

$$s^b(t) = \sum_{i=1}^{R} h_i^b h_i^f \sqrt{\Delta\sigma_i}a_i(t) + L + n^b(t), \tag{6}$$

where $h_i^f = |h_i^f|e^{\varphi_i^f}$ and $h_i^b = |h_i^b|e^{\varphi_i^b}$ are complex valued channel coefficients. The normalized RCS, leakage and noise can be written as $\sqrt{\Delta\sigma_i} = \sqrt{|\Delta\sigma_i|}e^{\varphi_i^{\Delta\sigma}}$, $L = |L|e^{\varphi^{leak}}$, and $n^b(t) = n(t)e^{j2\pi f_c t}$, respectively.

We further simplify the equation (6) by denoting $h_i = h_i^b h_i^f \sqrt{\Delta\sigma_i}$ as the general coefficient including both forward and backward channels as well as differential RCS coefficient. By introducing $\mathbf{h}$ with elements h_i, $\mathbf{a}(t)$ with elements $a_i(t)$, we can rewrite the received signal as

$$s^b(t) = \mathbf{h}\mathbf{a}(t) + L + n^b(t), \tag{7}$$

where signal to noise ratio (SNR) is defined as $\mathbb{E}\{|h_i|^2 a_i^2\}/N_0$.

C. Received Signal Constellation Points

The received baseband signal is complex-valued and contains both in-phase and quadrature components. The received signal $s^b(t)$ and its constellation points in I/Q plane for different values of number of tags up to 4 is illustrated in Fig. 4. The first column plots correspond to the amplitudes of both I/Q components of the received signal where second column plots depict the constellation points in I/Q plane. One major observation is that for each number of tag we have different amplitude levels which results in higher number of clusters as we increase the number of tags. Since we only consider ON-OFF keying, the transmitted symbols can only take values ± 1. As a result of that, the cluster centers and signal levels correspond to the distinct combinations of the channel gains, h_i, and therefore, we will have 2^R different levels and clusters when the number of tags in collision is R. If we consider only 1 tag scenario, the received signal levels and the cluster centers are simply $\pm h_1$.

[1]We note that the parameters, i.e., the channel coefficients, phase shifts, differential RCS is constant during the transmission of the modulated backscattered tag signals.

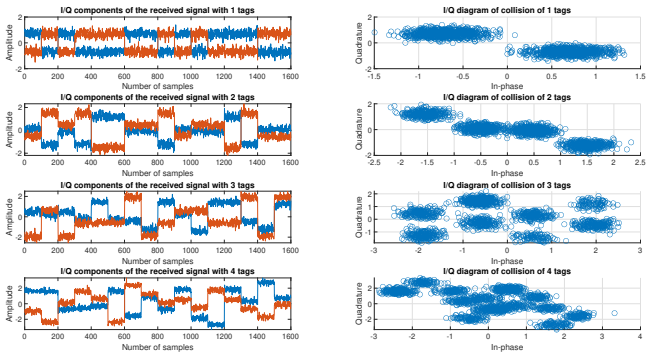

Fig. 4: Illustration of the received signal levels and constellation points for different number of tags up to 4 when SNR is 20 dB.

III. ML-Aided Collision Recovery Algorithm

In this section, we present our proposed ML-aided collision recovery algorithm for number of tags up to 4 and the diagram of our algorithm is shown in Fig. 5. As a first step, we estimate the number of tags in collision by using GMM and neural network architectures, i.e., feed-forward (FNN) and convolutional (CNN), utilizing the received signal constellation points in I/Q plane. Then, 4 different FNN models are trained to estimate the channel coefficients for given number of tags with the aid of 4 additional symbols. After finding the number of tags and the channel gains, we apply minimum distance decoder to separate the transmitted signals of the passive tags.

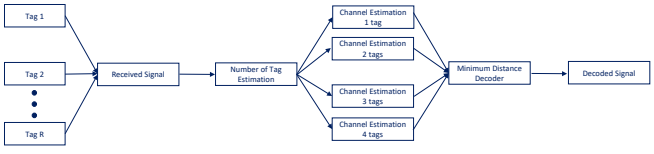

Fig. 5: The workflow diagram of the proposed ML-aided collision recovery algorithm for number of tags up to 4.

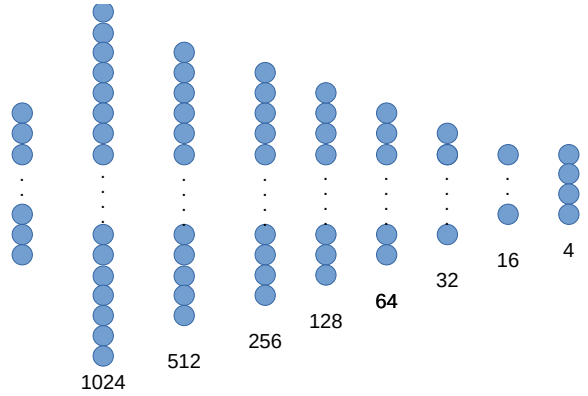

Fig. 6: Overview of the feed-forward network architecture with 8 linear layers where the number of units are given at the bottom of each layer.

A. Number of Tag Estimation

We first explain the preferred architectures for the tag estimation problem. Then, we describe the training dataset and procedure. Finally, we compare the performance results of each approach in detail.

We first consider a well-known clustering method Gaussian mixture model (GMM) as our initial approach. GMM is a probabilistic model-based clustering technique where the samples are drawn from the mixture of normal distributions, i.e., $p(\mathbf{x_n}) = \sum_{l=1}^{L} \pi_l N(\mathbf{x_n} \mid \mu_l, \Sigma_l)$ where $\mathbf{x_n}$ is a set of received signal constellation points, π_l is mixture probabilities, i.e., $0 \leq \pi_l \leq 1$ for all l and $\sum_{l=1}^{L} = 1$, $N(\mathbf{x_n} \mid \mu_l, \Sigma_l)$ is a normal distribution with a mean μ_l and a covariance Σ_l where L is the number of components in the mixture which is equal to 2^R. GMM technique suits well for our problem since the received signal is a mixture of normally distributed constellation points and they are centered at the combinations of the channel coefficients (mean) and deviates around the center points with an AWGN noise (covariance)[2]..

We use expectation-maximization (EM) algorithm as an iterative solution to find a well-performing parameters (π_l, μ_l, Σ_l) of GMM [15] since the optimal parameter values cannot be obtained in a closed-form solution.

We extend our approach for number of tag estimation by using FNN and the architecture is depicted in Fig. 6. The received signal is fed into the model. The FNN model comprises of 7 hidden layers and an output layer with the number of units starting from 1024 and gradually decreasing through the layers where the output layer has 4 units corresponding to the number of tags to be estimated. All linear layers applies a matrix multiplication with weight parameters and also adds bias terms. After linear operations, a ReLU (rectified linear unit) activation function is applied at the hidden layers. Differently, at the output layer, soft-max activation function is used to get probabilities for different number of tags.

We also consider CNN with an aim of improving the performance of FNN since CNN might be able to extract the features from the constellation points through the convolutional layers. We use similar architecture for CNN as in FNN which is illustrated in Fig 7. Different from the FNN, we feed the received signal to the 3 convolutional layers with a number of channels 16, 64 and 32. The kernel size is selected as 5 and ReLU activation function is used for each convolutional layer. The convolutional layers is then followed by 5 linear layers with a number of units 512, 256, 64, 32, and 4 respectively where each layer is activated by ReLU function.

[2]For GMM, we pick the number of tags by utilizing the estimated number of clusters. We select the number of tags as x when the number of clusters lies between 2^{x-1} and 2^x.

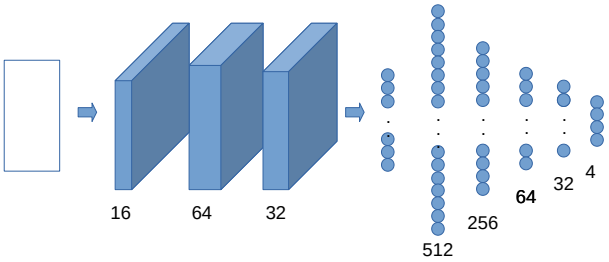

Fig. 7: Overview of the convolutional network architecture with 3 convolutional and 5 linear layers where the number of units are given at the bottom of each layer.

We describe the training dataset used for the training of the GMM, FNN and CNN. We generate 10^5 received signal samples for each number of tags in a total of 4×10^5 samples. Each received signal sample is formed as follows: 1) Each tag modulates 16 bit random number (RN16), 2) The Rayleigh channel coefficients are generated as an independent and identically distributed (i.i.d) with zero mean and unit variance complex Gaussian random variables, i.e., $\mathcal{CN}(0, 1)$, 3) The noise realizations are also i.i.d complex Gaussian random variables with zero mean and variance N_0, i.e., $\mathcal{CN}(0, N_0)$.

For the training procedure, we apply Bayesian-based information criterion (BIC) score to avoid over-estimating the number of clusters by applying penalization over the likelihood function for GMM [16]. For FNN and CNN, Adam optimizer [17] is used with a learning rate $\eta = 0.001$. We select the mini-batch size as 1024 and employ cross-entropy loss function.

The accuracy values of GMM, FNN and CNN are given in Table I for tag numbers up to 4 at training SNR value of 20 dB. While GMM performs well for 1 and 2 tags, the accuracy value suffers for higher number of tags. On the other hand, FNN and CNN perform quite well for 3 and 4 tags, especially CNN. It can be seen that CNN estimate the number of tags almost perfectly with an overall accuracy value of 0.972.

TABLE I: Comparison of number of tags accuracy values for GMM, FNN and CNN at training SNR value of 20 dB.

Number of tags	GMM	FNN	CNN
1 tag	0.9999	1	1
2 tags	0.9377	0.969	0.998
3 tags	0.8664	0.814	0.972
4 tags	0.7541	0.811	0.919
Overall	0.889	0.899	0.972

B. Channel Estimation

We now explain our proposed approach for channel estimation by utilizing the estimated number of tags. After finding out the number of tags, four different FNN models are trained with the aid of 4 fixed symbols in addition to the RN16. We use fixed additional symbols for this specific problem because if the symbols are formed in a random manner similar to the RN16, the neural network cannot capture the relationship between the received signal and the channel coefficients[3]. We select the number of symbols as 4 since we are interested in estimating the channel coefficients for at most 4 tags. We note that the additional symbols should be different for each tag and orthogonal to each other. The main idea of using FNN model is to extract the channel coefficients from the different combinations by denoising the received signal which is corrupted by the AWGN noise.

The FNN model for channel estimation is depicted in Fig. 8 and it contains 4 hidden layers and an output layer. Similar to the previous model, the received signal with only 4 additional symbols is fed into the model. The hidden layers have equal number of units which is 128 and ReLU activation function is used at each hidden layer. The output layer has no activation function and consists of $2R$ units where multiplier 2 is used for the real and imaginary parts. For the training of each FNN architecture, we produce the training dataset in a similar fashion to the number of tag estimation case. We use Adam optimizer with learning rate $\eta = 0.001$. We select the mini-batch size as 1024 and also employ minimum mean squared error (MMSE) as the loss function.

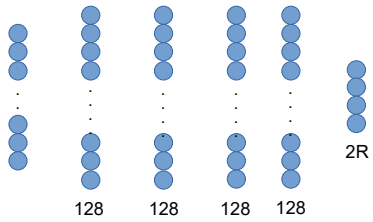

Fig. 8: Overview of the feed-forward network architecture for channel estimation with 5 linear layers where the number of units are given at the bottom of each layer.

C. Minimum Distance Decoder

At the final step in our collision recovery algorithm, we perform minimum distance decoding after finding the channel coefficients for each tag. As described earlier in Section II, the received signal levels or the cluster centers are the combinations of the channel coefficients with different signs. As an example, if we consider 2 tags with respective channel gains h_1 and h_2, we will have 4 different levels (cluster centers) in the received signal denoted by l_1, l_2, l_3, and l_4 which can be computed as

$$l_1 = h_1 + h_2, \;\; l_2 = h_1 - h_2, \;\; l_3 = -h_1 + h_2, \;\; l_4 = -h_1 - h_2.$$

After computing the signal levels which can be done for higher number of tags in a similar manner, the maximum likelihood decision rule is equivalent to a minimum distance decoding for the given observation (received signal $s^b(t)$)[4], i.e.,

$$\hat{i} = \underset{i \in \{1,2,\cdots,2^R\}}{\operatorname{argmin}} \; || \, l_i - s^b(t) \, || \tag{8}$$

As a final step, after finding the optimal signal level for each RN16 symbols, we can derive the transmitted bits/symbols of passive tags for the given signal level. Specifically, if $\hat{i} = 1$, both tags transmit 1 in their reflecting states and if $\hat{i} = 2$, first tag and second tag transmit 1 and -1, respectively.

IV. Numerical Results

In this section, we illustrate the throughput performance of the proposed collision recovery algorithm. We first consider the scenario where the receiver can only decode one tag from the collision ($J = 1$), and then, extend our examples to the scenario where the receiver is capable of decoding up to 4 tags ($J = 4$). For both scenarios, we assume that the receiver is equipped with single antenna and the channel coefficients are i.i.d with zero mean and unit variance complex Gaussian random variables, i.e., $\mathcal{CN}(0, 1)$. We perform our simulations over different SNR values under the optimal frame to tag ratio (K/N) as derived in Section II.

In the first simulation setup, we consider the case where the receiver can only decode one tag from the collision up to 4 tags ($M = 4, J = 1$). Fig. 9 illustrates the throughput performance of GMM, FNN and CNN over various SNR values. For comparison purposes, we also include the throughput results of the ideal scenario where the number of tags and channel coefficients are known perfectly by the receiver and conventional FSA algorithm. The dashed lines indicate the theoretical throughput of conventional FSA ($M = 1, J = 1$) and FSA with collision recovery ($M = 4, J = 1$). We observe from Fig. 9 that CNN outperforms both GMM and FNN since it estimates the number of tags with higher accuracy. Indeed, CNN provides very close throughput values to the ideal scenario and the loss is almost negligible. It can also be seen that the throughput improves as SNR increases and approaches to the theoretical limit.

In the second simulation setup, we consider the case where the receiver decode up to 4 tags out of 4 colliding tags ($M = 4, J = 4$). We depict the throughput values of GMM, FNN and CNN in Fig. 10 along with the ideal scenario and conventional FSA. The throughput gain compared to the previous setup is easy to observe and throughput value nearly approaches to 2. The observations are similar to the previous setup and CNN still outperforming the other alternatives.

Finally, we compare our results with the existing works in the literature with single antenna receiver and present the comparison of maximum throughput values in Table II. The work [9] considers only 2 colliding tags and recovery of a single tag ($M = 2, J = 1$) with zero-forcing receiver, thus,

[3]It is also worth to mention that preamble symbols cannot be used for this problem because they are same for all tags.

[4]We note that the decision rule we consider is identical to the maximum likelihood decoding of the Pulse Amplitude Modulation (PAM) signals.

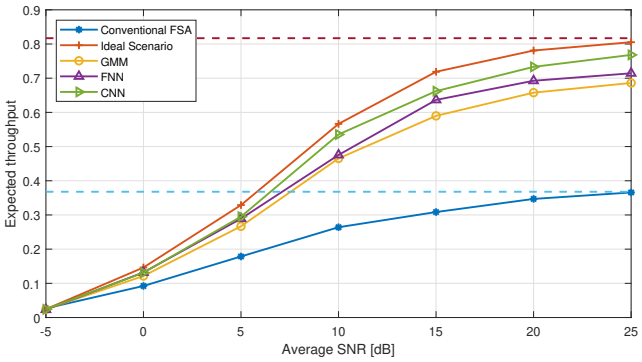

Fig. 9: Throughput performance comparison of GMM, FNN and CNN along with the ideal scenario with the perfect knowledge of number of tags and channel coefficients ($M = 4, J = 1$) and conventional FSA.

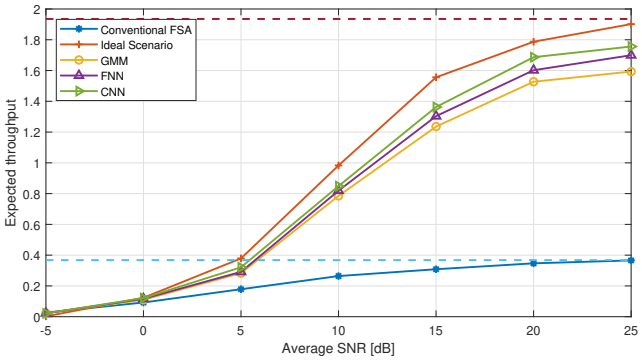

Fig. 10: Throughput performance comparison of GMM, FNN and CNN along with the ideal scenario with the perfect knowledge of number of tags and channel coefficients ($M = 4, J = 4$) and conventional FSA.

the throughput performance is lower than the other works. The authors in [11] employs a MMSE receiver which is capable of recovering 2 colliding tags ($M = 2, J = 2$) and obtain a throughput value of 0.841. In [12], voltage clustering is applied for the constellation points with a throughput value of 0.85 under the simulation setup $M = 3, J = 2$. With the proposed approach, we show that it is possible to decode 4 tags out of 4 colliding ones ($M = 4, J = 4$) by improving throughput value to the 1.756.

V. Conclusions and Future Work

We have considered the problem of recovering signals in collision for UHF-RFID systems. Different from the existing recovery methods, we propose a machine learning based algorithm. We show that the proposed models can estimate the number of tags and the channel coefficients with suitable design and training. The trained models enable the receiver to recover tag signals from the collided one with the application of minimum distance decoding. We also perform simulations to demonstrate the performance of the proposed approach which provides a significant improvement in throughput.

As a future work, one direction might be to confirm the validity of the proposed method on measurement data obtained from the real time implementations. In our work, we consider a simplified communication model to show the applicability of deep learning for UHF-RFID systems and it is possible to extend our approach for more realistic setups taking into account multi-path fading as well as the impacts of phase and carrier frequency. In addition, we study a system model where the receiver is equipped with only single antenna which can be extended to the multiple antenna setup. Another line of work might be to examine the complexity of the trained models and how they can be coupled with the experimental configuration.

TABLE II: Comparison of our work with the existing works.

Work	Maximum Throughput	Method & Setup
[9]	0.587	Zero-Forcing Receiver ($M = 2, J = 1$)
[11]	0.841	MMSE Receiver ($M = 2, J = 2$)
[12]	0.85	Voltage Clustering ($M = 3, J = 2$)
Our work	**0.768**	**ML-based recovery** ($M = 4, J = 1$)
Our work	**1.756**	**ML-based recovery** ($M = 4, J = 4$)

References

[1] EPCGlobal, "EPC Radio-Frequency Identity protocols Class-1 Generation-2 UHF RFID protocol for communications at 860 Mhz–960 Mhz," *Version*, vol. 1, no. 0, p. 23, 2008.

[2] F. Schoute, "Dynamic frame length ALOHA," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 565–568, 1983.

[3] D. Shen, G. Woo, D. P. Reed, A. B. Lippman, and J. Wang, "Separation of multiple passive RFID signals using software defined radio," in *2009 IEEE International Conference on RFID*, 2009, pp. 139–146.

[4] R. S. Khasgiwale, R. U. Adyanthaya, and D. W. Engels, "Extracting information from tag collisions," in *2009 IEEE international conference on RFID*, 2009, pp. 131–138.

[5] J. Yu, K. H. Liu, X. Huang, and G. Yan, "An anti-collision algorithm based on smart antenna in RFID system," in *2008 International Conference on Microwave and Millimeter Wave Technology*, vol. 3, 2008, pp. 1149–1152.

[6] A. F. Mindikoglu and A.-J. van der Veen, "Separation of overlapping RFID signals by antenna arrays," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 2737–2740.

[7] B. Knerr, M. Holzer, C. Angerer, and M. Rupp, "Slot-wise maximum likelihood estimation of the tag population size in FSA protocols," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 578–585, 2010.

[8] C. Angerer, G. Maier, M. V. Delgado, M. Rupp, and J. V. Alonso, "Single antenna physical layer collision recover receivers for RFID readers," in *2010 IEEE International Conference on Industrial Technology*, 2010, pp. 1406–1411.

[9] C. Angerer, R. Langwieser, and M. Rupp, "RFID reader receivers for physical layer collision recovery," *IEEE Transactions on Communications*, vol. 58, no. 12, pp. 3526–3537, 2010.

[10] J. Kaitovic, R. Langwieser, and M. Rupp, "RFID reader with multi antenna physical layer collision recovery receivers," in *2011 IEEE International Conference on RFID-Technologies and Applications*, 2011, pp. 286–291.

[11] J. Kaitovic, M. Šimko, R. Langwieser, and M. Rupp, "Channel estimation in tag collision scenarios," in *2012 IEEE International Conference on RFID (RFID)*, 2012, pp. 74–80.

[12] X. Tan, H. Wang, L. Fu, J. Wang, H. Min, and D. W. Engels, "Collision detection and signal recovery for UHF RFID systems," *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 1, pp. 239–250, 2016.

[13] H. Mahdavifar and A. Vardy, "Coding for tag collision recovery," in *2015 IEEE International Conference on RFID (RFID)*. IEEE, 2015, pp. 9–16.

[14] P. V. Nikitin, K. Rao, and R. D. Martinez, "Differential RCS of RFID tag," *Electronics Letters*, vol. 43, no. 8, pp. 431–432, 2007.

[15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[16] G. Schwarz, "Estimating the dimension of a model," *The annals of statistics*, pp. 461–464, 1978.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Rep. (ICLR)*, San Diego, USA, May 2015.