Deep Learning Architectures for Activity Detection in GF-SCMA: From Extraction to Iteration

Metasebia D. Gemeda * , Minsig Han, Ameha T. Abebe, and Chung G. Kang *
School of Electrical Engineering, Korea University, Seoul, Republic of Korea *
Samsung Research, Seoul, Republic of Korea
Emails:[met4, als4585, ameha_tsegaye, ccgkang]@korea.ac.kr

***Abstract*—Massive machine-type communication (mMTC) requires robust and low-overhead active user detection (AUD) to cope with collisions in grant-free sparse code multiple access. Conventional compressive sensing-based AUD degrades under higher activity, while deep learning (DL) detectors offer higher accuracy but remain challenged by many-to-one preamble–codebook mappings. An extraction-based AUD network has shown that data features can be successfully exploited under collisions, but its one-shot design restricts trade-off control and robustness. We extend this idea to an iterative AUD network with progressive refinement and dual thresholds, enabling tunable false alarm and missed detection rates. Comparative evaluation shows that the iterative method achieves superior receiver operating characteristic (ROC) performance over previous work and other state-of-the-art DL solutions, validating feasibility and advancing toward scalable and reliable AUD for mMTC.**

***Index Terms*—Grant-free random access, active user detection, deep learning**

I. Introduction

Massive machine-type communication (mMTC) requires scalable and low-overhead random access to support a vast number of sporadically active devices. Grant-free sparse code multiple access (GF-SCMA) has emerged as a promising solution, where contention transmission units (CTUs) combine time, frequency, preamble, and codebook resources to enable uplink access without explicit scheduling. To provide sufficient contention opportunities within limited resources, each codebook is typically associated with multiple preambles, leading to a many-to-one preamble–codebook mapping. While this increases flexibility, it also introduces a critical bottleneck: collisions and non-orthogonality amplify the difficulty of active user detection (AUD), and errors at this stage directly degrade both channel estimation and multi-user decoding.

Traditional compressive sensing (CS)-based AUD [1] exploits sparsity but degrades sharply with higher activity and incurs excessive complexity when iterative interaction with multi-user detection is required. Deep learning (DL)-based detectors have been proposed as alternatives, with preamble-based (P-AUDN) [2] offering a simple baseline but lacking data-awareness, and conventional data-aided AUD (D-AUDN) [3], [4] proving effective only under one-to-one mapping. Directly combining preamble and data features in a single network enlarges the feature space, making training inefficient and limiting potential gains. To address this, an extraction method network [5] was introduced to process the superposed data stream into latent priors for an LSTM-based detector, showing for the first time that data-aided AUD can yield gains even with codebook collisions. However, its reliance on a single threshold limited explicit control over the balance between false alarms and missed detections.

Building on this feasibility result, we propose an iterative D-AUD architecture that restructures detection into three progressive blocks: a coarse preamble-based detector maximizing recall, a contextual refinement stage integrating preamble and data features, and a fine detector delivering high-precision final decisions. By employing dual thresholds tuned via the target performance score, this architecture reduces error propagation, enables flexible trade-offs between false alarms and missed detections, and achieves robustness under fading channels and higher system loading. Thus, while the extraction-based design established the feasibility of leveraging data features under many-to-one mapping, our iterative method extends this into a scalable and tunable solution for future GF-SCMA systems.

II. System Model and AUD Architectures

We consider a grant-free SCMA uplink where M potential devices share J codebooks (CBs). Each CB $\mathcal{C}_j = \{\mathbf{c}_m^{(j)}\}_{m=0}^{M-1}$ consists of sparse codewords $\mathbf{c}_m^{(j)} \in \mathbb{C}^K$ with $N < K$ nonzeros. A CTU is defined as a pair of a Zadoff–Chu (ZC) preamble sequence and a CB, such that multiple preambles are associated with the same CB, enlarging the contention space. In each slot, only a small fraction of users are active ($N_A \ll M$). An active user n selects a CTU, i.e., preamble $\mathbf{p}_{\ell(n)}^{(\nu(n))}$ and CB $\nu(n)$, and transmits both a short preamble and a sequence of SCMA codewords $\{\mathbf{w}_i^{\nu(n)}\}$ over flat Rayleigh fading channels h_n. The received signals at the base station are modeled as

$$\mathbf{y}^{(p)} = \sum_{n=1}^{M} \delta_n h_n \mathbf{p}_{\ell(n)}^{(\nu(n))} + \mathbf{n}^{(p)}, \quad \mathbf{n}^{(p)} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}), \quad (1)$$

$$\mathbf{y}_i^{(d)} = \sum_{n=1}^{M} \delta_n h_n \mathbf{w}_i^{\nu(n)} + \mathbf{n}_i^{(d)}, \quad \mathbf{n}_i^{(d)} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}), \quad (2)$$

where $\delta_n \in \{0,1\}$ denotes the activity of user n, and $i = 0, \ldots, N_d - 1$. The detection task is to estimate the activity vector $\hat{\delta}$ from $\mathbf{y}^{(p)}$ and $\{\mathbf{y}_i^{(d)}\}$. On this common system model, we compare two DL-based AUD architectures as seen in Fig. 1: (i) an *extraction-based* detector that maps the received signals to an activity vector via a feature extractor followed by a recurrent decision stage with a single threshold, and (ii) an

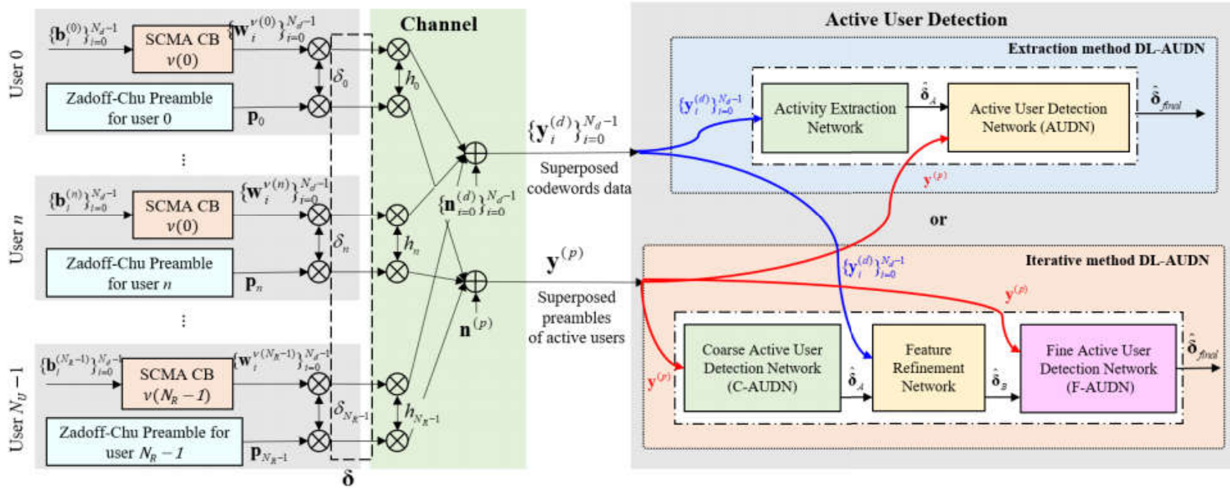

Fig. 1. System model of GF-SCMA transmission and DL based AUD options

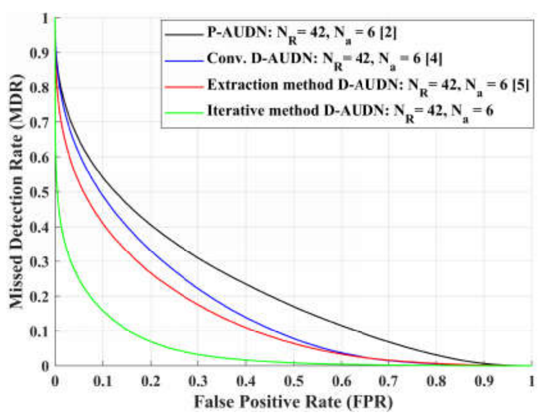

Fig. 2. ROC comparison at SNR = 20 dB, $N_R = 42$, $N_Z = 7$.

iterative method detector that performs coarse preamble-driven screening, collision-aware refinement using both preamble and data, and a final decision stage with dual thresholds.

III. Simulation Results

We evaluate the four AUD architectures under a grant-free SCMA uplink model with many-to-one mapping between preambles and codebooks. The setup assumes $N_a = 6$ active users out of $N_R = 42$ total possible users, each assigned a ZC sequence of length $N_Z = 7$ as the random access preamble. The channel follows i.i.d. Rayleigh flat fading with additive white Gaussian noise. Performance is measured via receiver operating characteristics (ROC), capturing the trade-off between false positive rate (FPR) and missed detection rate (MDR) as the decision threshold varies. Fig. 2 compares four schemes: the preamble-based AUDN (P-AUDN), the conventional data-aided AUDN (D-AUDN), the extraction-based D-AUDN, and the proposed iterative D-AUDN. The results show that both data-aided approaches significantly outperform P-AUDN and conventional D-AUDN across the full FPR range. Moreover, the proposed iterative design achieves the steepest ROC curve, consistently reducing MDR at all operating points. For instance, at an FPR of 0.1, the iterative method achieves a notably lower MDR compared to the extraction-based baseline. This validates that, while the extraction design established the feasibility of leveraging data features under collisions, the iterative refinement further enhances accuracy and tunability with minimal added complexity. Future research will expand evaluation to broader operating regimes and quantify computational overhead to further validate practicality.

IV. Conclusion

This work compared two DL-based AUD designs for GF-SCMA under many-to-one preamble–codebook mapping. The extraction-based AUD demonstrated for the first time the feasibility of using data features to improve detection under collisions. Building on this, the iterative alternative introduces progressive refinement and dual-threshold control, reducing error propagation and enhancing robustness under heavier loading. Simulation results confirmed that proposed method consistently outperforms previous DL-AUD designs in ROC performance, highlighting its potential as a scalable and tunable AUD framework for future mMTC systems.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No.2021-0-00467, Intelligent 6G Wireless Access System).

References

[1] B. K. Jeong, B. Shim, and K. B. Lee, "MAP-Based Active User and Data Detection for Massive Machine-Type Communications," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8481–8494, 2018.

[2] T. Sivalingam, S. Ali, N. H. Mahmood, N. Rajatheva, and M. Latva-Aho, "Deep Learning-Based Active User Detection for Grant-free SCMA Systems," in *2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 635–641, 2021.

[3] W. Kim, Y. Ahn, and B. Shim, "Deep Neural Network-Based Active User Detection for Grant-Free NOMA Systems," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2143–2155, 2020.

[4] Y. Ahn, W. Kim, and B. Shim, "Active User Detection and Channel Estimation for Massive Machine-Type Communication: Deep Learning Approach," *IEEE Internet Things J.*, vol. 9, no. 14, pp. 11904–11917, 2022.

[5] M. Han, M. D. Gemeda, A. T. Abebe, and C. G. Kang, "Deep learning-based data-aided activity detection with extraction network for grant-free sparse code multiple access," in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2025.