

## How to Identify and Authenticate Users in Massive Unsourced Random Access

Kotaba, Radoslaw; Kalør, Anders Ellersgaard; Popovski, Petar; Leyva-Mayorga, Israel; Soret, Beatriz; Guillaud, Maxime; Luis, Garcia Ordonez

*Published in:*  
I E E E Communications Letters

*DOI (link to publication from Publisher):*  
[10.1109/LCOMM.2021.3118882](https://doi.org/10.1109/LCOMM.2021.3118882)

*Publication date:*  
2021

*Document Version*  
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*  
Kotaba, R., Kalør, A. E., Popovski, P., Leyva-Mayorga, I., Soret, B., Guillaud, M., & Luis, G. O. (2021). How to Identify and Authenticate Users in Massive Unsourced Random Access. *I E E E Communications Letters*, 25(12), 3795 - 3799. Article 9564037. <https://doi.org/10.1109/LCOMM.2021.3118882>

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# How to Identify and Authenticate Users in Massive Unsourced Random Access

Radosław Kotaba, *Member, IEEE*, Anders E. Kalør, *Student Member, IEEE*, Petar Popovski, *Fellow, IEEE*, Israel Leyva-Mayorga, *Member, IEEE*, Beatriz Soret, *Member, IEEE*, Maxime Guillaud, *Senior Member, IEEE*, and Luis G. Ordóñez, *Member, IEEE*

**Abstract**—Identification and authentication are two essential features for traditional random access protocols. In ALOHA-based random access, the packets usually include a field with a unique user address. However, when the number of users is massive and relatively small packets are transmitted, the overhead of including such field becomes restrictive. In unsourced random access (U-RA), the packets do not include any address field for the user, which maximizes the number of useful bits that are transmitted. However, by definition an U-RA protocol does not provide user identification. This paper presents a scheme that builds upon an underlying U-RA protocol and solves the problem of user identification and authentication. In our scheme, the users generate a message authentication code (MAC) that provides these functionalities without violating the main principle of unsourced random access: the selection of codewords from a common codebook is i.i.d. among all users.

**Index Terms**—massive access, unsourced random access

## I. INTRODUCTION

One of the hallmarks of the fifth generation (5G) wireless systems and beyond is massive Internet of Things (IoT) connectivity [1]. A scenario for massive IoT access features a large number of devices (typically in the order of thousands) connected to a Base Station (BS), each being sporadically active and sending short data packets (e.g., a few kilobytes or bytes). This sporadic activation entails that the set of devices trying to access at a given instant is unknown, thereby requiring random access protocols.

In the classical ALOHA model for random access [2], a packet is the smallest, atomic unit of information. The analyses in massive access scenarios are usually performed with an infinite population, where the number of users is  $N \rightarrow \infty$ . However, in order to examine the fundamental performance bounds of massive access protocols, one needs to look into the structure of the packet. This is where the assumption  $N \rightarrow \infty$  leads to a paradox: to make user identification possible, a field with a unique user address of  $\approx \log_2 N$  bits must be included in a packet of finite and relatively short length. To deal with this paradox, two information-theoretic approaches have been introduced. In the many access channel [3] the number of users is given as a function of the codeword length, which allows to preserve identification capabilities even as both tend to infinity.

Differently from this, [4] addresses the problem of  $N \rightarrow \infty$  with finite blocklength (FBL) packets by assuming that a packet does not contain the address of the sender. This makes the access scheme *unsourced*, and leads to the case in which all users share the same codebook. While U-RA was initially proposed as a theoretically elegant scheme, it can also be justified by the desire to simplify the receiver and reduce the communication overhead. This is particularly important for short IoT packets where the address field can constitute a large portion of the packet [5].

The unsourced, uncoordinated nature of the problem and the FBL effects have implications in the design of practical low-complexity coding schemes, which has been the focus of several works. Bounds of the performance of finite-length codes were derived in the initial paper by Polyanskiy [4], and later generalized to the quasi-static fading channel [6]. The basic unsourced random access was extended to the case with a large number of antennas in [7], and the impact of correlated activations was studied in [8].

Despite its benefits in terms of efficiency, U-RA keeps the question of user identification (and, consequently, user authentication) open. In this paper, we aim to answer the following: *assuming that a given protocol for unsourced random access is available as a black box, how can it be extended to support user identification and authentication?* Rather than deferring this question to the higher layers or additional transmissions, in this contribution we present a scheme that enables those functionalities at the lower layers, in a way that is consistent with the paradigm of U-RA, i.e., when users share the same codebook. In that sense, the main contribution of our scheme is that it enables the identification and authentication of users over U-RA; the potential performance gains compared to sourced random access is of secondary importance.

The key idea is to generate and append a message authentication code (MAC)<sup>1</sup> to the packets (rather than an explicit address), which enables the identification and authentication of the users while complying with the main assumptions of U-RA. For this, we employ a two-step procedure as illustrated in Fig. 1. First, the BS broadcasts a beacon with a *nonce* to the users prior to data transmission. A nonce is an arbitrary number generated periodically by the BS that is allowed to be used only once by each node to prevent replay of messages. Then, each active user generates a MAC based on the nonce,

R. Kotaba, A. E. Kalør, P. Popovski, I. Leyva-Mayorga, and B. Soret are with the Department of Electronic systems, Aalborg University, Aalborg 9220, Denmark (e-mail: {rak, aek, petarp, ilm, bsa}@es.aau.dk).

M. Guillaud, L. G. Ordóñez are with Huawei Technologies, Paris, France (e-mail: maxime.guillaud@huawei.com, luis.ordonez@huawei.com).

<sup>1</sup>To avoid confusion between this term and the widely-used acronym for medium access control, the latter is avoided throughout the paper.

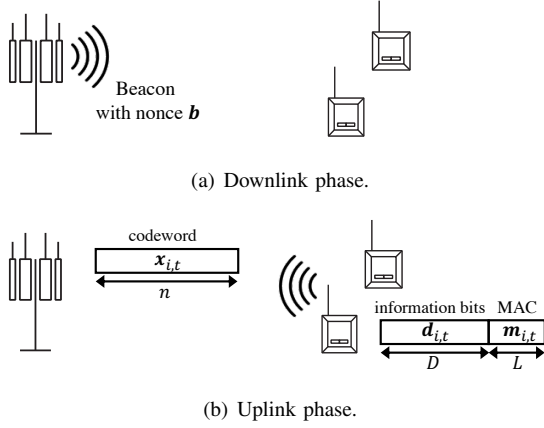


Fig. 1. The two-step procedure.

a secret key known only by the user and the BS (e.g., pre-shared using Universal Subscriber Identity Module (USIM) as in LTE [9]), and the data to be sent; this field is appended to the packet and transmitted as shown in Fig. 1(b).

## II. SYSTEM MODEL

We study the massive random access scenario as described by Polyanskiy [4], where  $N \rightarrow \infty$  users communicate through a time-slotted channel with a single BS. Although the proposed scheme works without modifications with a (potentially massive) MIMO BS, we assume a single antenna BS to simplify the presentation. At each time slot,  $K$  out of the  $N$  users are active and send messages  $\mathcal{W} = \{W_1, W_2, \dots, W_K\}$  in the uplink, where  $W_i$  is drawn independently and uniformly at random from the message set  $\mathcal{M} = \{1, 2, \dots, M\}$ . For typical massive IoT scenarios,  $K$  will be in the range of 50 to a few hundreds [10]. All users share the same encoder  $f : [M] \rightarrow \mathcal{X}^n$ , and use it to construct the codewords  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K$  as  $\mathbf{x}_i = f(W_i)$ , which are subject to the constraint  $\|\mathbf{x}_i\|_2^2 \leq nP$ , where  $P$  is the average energy per symbol. The codewords are transmitted over a permutation-invariant and memoryless multiple access channel  $P_{Y|X^K} : \mathcal{X}^{n \times K} \rightarrow \mathcal{Y}^n$ , i.e., it satisfies  $P_{Y|X^K}(\mathbf{y}|\mathbf{x}_1, \dots, \mathbf{x}_K) = P_{Y|X^K}(\mathbf{y}|\mathbf{x}_{\pi(1)}, \dots, \mathbf{x}_{\pi(K)})$  for any  $\mathbf{y} \in \mathcal{Y}^n$  and  $\mathbf{x}_1, \dots, \mathbf{x}_K \in \mathcal{X}^{n \times K}$ , and any permutation  $\pi$ .

We assume that the BS periodically broadcasts a beacon in the downlink as depicted in Fig. 1. The beacon includes the necessary information for the users to synchronize, to obtain the main configuration parameters, and to estimate and invert the channel. To keep the presentation simple and aligned with [4], we assume that channel inversion is perfect, so that fading can be neglected<sup>2</sup> and the uplink transmissions are only affected by additive white Gaussian noise, denoted by  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ . Consequently, the resulting Gaussian multiple access channel model at a given time slot is

$$\mathbf{y} = \sum_{i=1}^K \mathbf{x}_i + \mathbf{z} \quad (1)$$

<sup>2</sup>We note that the users who cannot perform inversion due to poor channel conditions can simply remain inactive, which leads to the problem that is structurally the same.

At the BS, the decoder  $g : \mathcal{Y}^n \rightarrow [M]^K$  outputs an unordered list of  $K$  messages from  $\mathcal{M}$ . In line with the U-RA literature [4], we assume that  $K$  is fixed and known to the decoder. We note that this assumption allows the codebook to be designed based on  $K$ , which does not reflect a true random access scenario. In practice, the codebook would have to be designed based on the expected maximum (or average) number of active users instead. Similarly, in practical implementations the decoder, rather than outputting a fixed number of messages, might rely on separate activity detection [10].

An error occurs whenever the  $g(\mathbf{y})$  does not contain a transmitted message, or if multiple users transmit the same message. More specifically, an error for user  $i$  is defined as  $E_i = \{W_i \notin g(\mathbf{y})\} \cup \{W_i = W_j \text{ for some } j \neq i\}$ . Note that since we assume the decoder always outputs  $K$  messages, it implies that for each error  $E_i$ , the list  $g(\mathbf{y})$  must contain a message which was not transmitted by any of the devices. We shall refer to this set  $g(\mathbf{y}) \setminus \mathcal{W}$  as decoder false positives. Denoting by  $k_{\text{TP}}$  the number of genuine (true positive) messages and by  $k_{\text{FP}}$  the number of false positives in the set  $g(\mathbf{y})$ , we have that  $k_{\text{TP}} + k_{\text{FP}} = K$ .

## III. IDENTIFICATION AND AUTHENTICATION IN UNSOURCED RANDOM ACCESS PROTOCOLS

The key idea behind the proposed scheme is to generate MAC that enables identification and authentication of the users and that can be applied to U-RA protocols. The MAC  $\mathbf{m}_i \in \{0, 1\}^L$  is generated by user  $i$  based on its data  $\mathbf{d}_i \in \{0, 1\}^D$  of size  $D$ , its secret key  $\mathbf{k}_i$ , and a nonce  $\mathbf{b}$ . The secret key is fixed and only known by the corresponding user and the BS, e.g., pre-shared using USIM as in LTE [9]. The MAC length  $L$  is fixed and independent from the other parameters.

Our scheme is divided into phases as shown in Fig. 1. At the beginning of each round, the BS generates a nonce and broadcasts it to all the devices. The nonce is a sequence or pseudo-random number that changes in each round but is otherwise public. Once the nonce is received, a given user  $i$  generates the MAC  $\mathbf{m}_i$  based on the data bits it wants to transmit  $\mathbf{d}_i$ , its secret key  $\mathbf{k}_i$ , and the nonce  $\mathbf{b}$ , i.e.  $\mathbf{m}_i = h(\mathbf{d}_i, \mathbf{k}_i, \mathbf{b})$ , where  $h(\cdot)$  is designed to be computationally hard to invert and have low collision probability (i.e., the output is approximately uniform for any input distribution). The user appends the MAC to the data to create a packet and transmits it, as shown in Fig. 1(b). At the BS, the packets are first decoded to extract  $[\hat{\mathbf{d}}_i, \hat{\mathbf{m}}_i]$  tuples. For each, the message authenticity can be verified and the identity of the sender determined by computing the MACs of the data part  $h(\hat{\mathbf{d}}_i, \mathbf{k}_j, \mathbf{b})$  with different secret keys  $\mathbf{k}_j$  and comparing them with the MAC in the received packet  $\hat{\mathbf{m}}_i$ . If a match is found, the authenticator declares the user with the matching key to be the potential transmitter. The full scheme is depicted in Fig. 2.

While a nonce is commonly used to prevent replay attacks, in our scheme it has the additional function of randomizing the MAC. That is, without a nonce, a particular piece of data and secret key from a given device would always produce the same MAC, which violates the assumption that all codewords are equally likely. Typical methods to generate the MAC include,

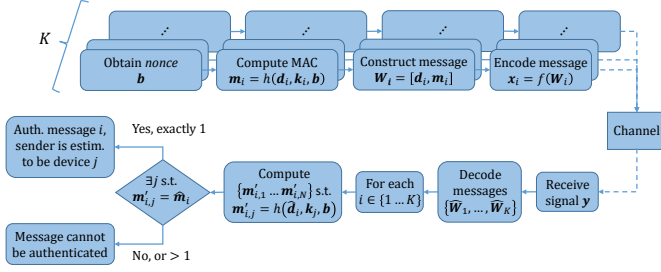


Fig. 2. Block diagram of the proposed scheme including message generation and subsequent decoding, authentication and identification.

e.g., symmetric key cryptography as in AES-CMAC (RFC 4493), used in LoRaWAN, or a HMAC (RFC 2104). Any of these methods can be applied to our scheme, so the MAC is computationally challenging to guess without the secret key.

Note that in our scheme *cryptographic errors*, which we define as any instance where the matching MAC is generated by a key that does not belong to the actual sender, can occur. They are possible since: 1) the generated MAC might not be a unique identifier for the user (unlike the actual address) and 2) the BS must generate many MACs with different secret keys to find the one that matches the one in the received packet.

Therefore, several tradeoffs arise. The first one is between the length of the metadata and the amount of *cryptographic errors*, where in the extreme case with no metadata (i.e. neither MAC nor address) identification and authentication cannot be provided. Meanwhile, longer packets entail higher energy. Another tradeoff involves the computational complexity and probability of cryptographic errors that both increase with the number of devices supported by the system<sup>3</sup>.

#### IV. CRYPTOGRAPHIC ERRORS: COLLISIONS, FALSE POSITIVES AND MISIDENTIFICATIONS

The probability of decoder false positives describes only the physical layer performance of U-RA. The full characterization of the proposed scheme has to take into account also potential cryptographic errors, erroneous acceptance of false positives, and misidentification events. For the purpose of this evaluation, we assume ideal MACs that are uniformly distributed, i.e., the probability that a given  $(data, key, nonce)$  tuple produces a specific MAC of length  $L$  is  $p = 2^{-L}$ .

##### A. Exhaustive search

We first consider authentication using exhaustive search, where all keys are tried on each message. We start by studying the per-user cryptographic error probabilities. A genuine message  $W'$  with data  $d'$  transmitted by user  $u'$  will fail to be authenticated whenever any of the keys from users  $u \neq u'$  produces the same MAC when applied to  $d'$ . We refer to those events as *type 1* errors. Since there are  $N - 1$  other keys, the type 1 event happens with probability

$$p_{t1} = 1 - (1 - p)^{N-1} \quad (2)$$

<sup>3</sup>It could be argued that the scheme is not practical as  $N \rightarrow \infty$ . However, in practice good performance was observed for  $N$  as large as  $10^6$  and  $K > 100$ .

Because we assume that each user transmits at most one message per round, an error occurs also when the key of user  $u'$  produces a valid MAC for any of the other decoded messages in  $g(y) \setminus \{W'\}$ . Given that there are  $K - 1$  other decoded messages, this *type 2* error happens with probability

$$p_{t2} = 1 - (1 - p)^{K-1}. \quad (3)$$

Taking into account both types of errors, the probability that a genuine message is successfully authenticated is

$$p_{s\_auth} = (1 - p)^{N+K-2}. \quad (4)$$

Another type of event is when a false positive message produced by the decoder is erroneously authenticated. While (4) is conditional on the fact that there is at least one key that authenticates the message, here we cannot assume that. Since the keys from the genuine messages cannot be used again without causing type 2 error, there are  $N - k_{TP}$  keys that can potentially decode the false positive message without resulting in a collision. Since each of these keys accepts the message with probability  $p_{s\_auth}$ , the probability of accepting a false positive message from the decoder is

$$\begin{aligned} p_{fp\_auth} &= (N - k_{TP}) p p_{s\_auth} \\ &= (N - k_{TP}) p (1 - p)^{N+K-2}. \end{aligned} \quad (5)$$

Note that the authenticator is generally unable to determine whether a message that *fails* to be authenticated belongs to the set of decoder true positive or decoder false positive messages. The only exception to this is the special case in which no key is able to decode a given message, which can only happen for false positive messages. The probability that this happens for a given false positive message is  $p_{d\_fp} = (1 - p)^N$ .

##### B. Heuristic search

We now turn our attention to the heuristic search, in which the authenticator tries keys only until it finds a matching key. While more efficient, this approach cannot detect type 1 and type 2 errors defined above, and thus the probability of erroneously authenticating a message increases.

Providing exact analytical expressions for the heuristic case proves to be difficult, due to the dependency on the order in which packets are authenticated, the number of decoder false positives and true positives, and how they are interleaved. To that end, we will provide only approximations, noting that they are very close to the true values. We shall assume without loss of generality that the decoded messages are authenticated in the order  $\hat{W}_1, \hat{W}_2, \dots, \hat{W}_K$ . Furthermore, we will neglect the events where the sender of message  $\hat{W}_j$  becomes incorrectly identified as the sender of one of the previous messages  $\hat{W}_1, \dots, \hat{W}_{j-1}$ , which happens with very low probability<sup>4</sup>.

We first consider the probability of correctly authenticating a genuine message. In the heuristic search case the successful authentication of message  $W_j$  can happen even if there are cryptographic collisions, as long as the correct user happens to be tested first. For a set of  $i$  successfully authenticating keys,

<sup>4</sup>Note that we do not neglect misidentification events in general, but only the case where specific user authenticates a specific message, which is tied to the probability  $p$  and hence very low.

this happens with probability  $1/i$ . By marginalizing over the number of keys *additional* to the genuine key we obtain

$$p_{s\_auth,j} = \sum_{i=0}^{N_j-1} \binom{N_j-1}{i} p^i (1-p)^{N_j-1-i} \left( \frac{1}{1+i} \right), \quad (6)$$

where  $N_j$  is the number of remaining keys which is the total number of keys,  $N$ , minus those that have authenticated any of the previous messages.  $N_j$  is nonincreasing, and  $N_j \geq N - j + 1$  since the authenticator may have been unable to authenticate some of the previous  $j - 1$  messages. As already mentioned, in the heuristic approach the detection of collisions (type 1 and type 2 errors) is not possible, which can result in misidentification, i.e., attributing a genuine message to the wrong user. The probability of misidentifying the  $j$ -th message is the probability that one or more of the  $N_j - 1$  non-genuine keys authenticate the message before the correct one:

$$p_{mis\_id,j} = \sum_{i=1}^{N_j-1} \binom{N_j-1}{i} p^i (1-p)^{N_j-1-i} \left( 1 - \frac{1}{1+i} \right) \\ = 1 - p_{s\_auth,j}. \quad (7)$$

On the other hand, if the message is a false positive, the probability of accepting it is equal to the probability of having at least one key which produces a matching MAC:

$$p_{fp\_auth,j} = 1 - (1-p)^{N_j}. \quad (8)$$

We note that from the point of view of the receiver there is no difference between misidentification and false positive authentication, hence, the total error probability should include both. For a given packet, which is genuine with probability  $p_{TP}$  and a false positive with probability  $p_{FP}$  we obtain

$$p_{mis\_auth,j} = p_{TP} p_{mis\_id,j} + p_{FP} p_{fp\_auth,j}. \quad (9)$$

Lastly, let us remark that when  $N \gg K$ , we have that  $N_j \approx N$ . By making this substitution in (6) - (9), we obtain a rigorous upper bound on the probability of each type of error. Furthermore, they become independent of packet number and allow us to drop the subscript  $j$  which simplifies the comparison between the exhaustive and heuristic approach.

In Fig. 3 we show the probability of successful authentication and probability of mis-authentication as a function of the total number of devices  $N$  for the exhaustive and heuristic search. In addition to the small gain in terms of success probability, the latter method allows to reduce the complexity as, on average, it requires only half of the MAC checks (assuming the probability of transmission is uniform across the devices). This is at the cost of an increased probability of mis-authentication. Since the eq. (6) and (9) used to produce the solid red curves are approximations that neglect some of the effects mentioned earlier, we provide also the results obtained through numerical simulations. Clearly, the differences are very minor making the approximations a viable tool.

### C. Spoofing attacks

It is of interest to consider what happens when an attacker sends a forged message with the intent of getting it accepted by

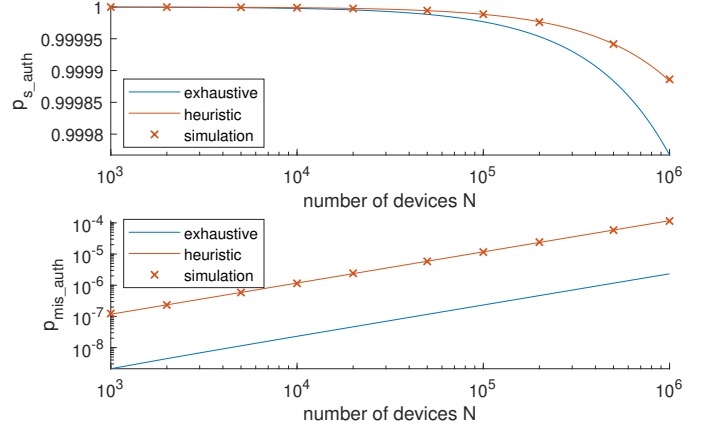


Fig. 3. Probability of successful authentication and probability of mis-authentication as a function of the total number of devices  $N$ . The number of messages is  $K = 100$ ,  $p_{TP} = 0.99$ ,  $p_{FP} = 0.01$ , MAC length  $L = 32$ bits.

the authenticator. Without private keys the attacker is not able to compute the correct MAC for the spoofed *data* and current *nonce* so it has to generate MAC bits at random. As such, from the cryptographic point of view, the message acts as a false positive, and the transmitter cannot target a specific device (i.e. it cannot choose whom it is impersonating). However, from the physical layer point of view this is an actual transmitted codeword, and as such subject to probability of decoding  $p_{TP}$ , so the total probability of successful spoof is

$$p_{s\_spoof} = p_{TP} (1 - (1-p)^N). \quad (10)$$

This is to be compared to the traditional frame structure where the source address is included in the packet. In that case, the authenticator only tries the single MAC associated to that user, and the spoof attack is successful with probability  $p_{TP}p$ .

## V. RESULTS

We start by looking into the physical layer performance. The results were obtained based on the random coding bound given in [4, eq. (3)-(10)]. The codeword length (number of symbols) was chosen to be  $n = 2^{15} = 32768$ . In Fig. 4 we depict the achievable error probability as a function of the energy per codeword  $nP$ . The values are shown for a range of packet sizes  $B$  and for  $K = 50$  and  $K = 150$ . In line with the assumption that two users selecting the same message is considered an error (c.f. Section II), each curve has a floor at  $\binom{K}{2}/M = \binom{K}{2}/2^B$  (visible only for 32 bits). In general, the higher  $B$  and  $K$  are, the steeper the curves become and the transition from almost certain error  $p_{FP} \approx 1$  to very high reliability (such as  $p_{FP} < 10^{-5}$ ) becomes increasingly abrupt. This is even better explained with Fig. 5 which shows the energy per codeword as a function of  $B$  for fixed error rates. Firstly, as the packet size increases, less energy is needed to decrease  $p_{FP}$ . For example, with  $K = 50$ , when  $B = 25$  b improving error rate from  $10^{-1}$  to  $10^{-3}$  requires 1 dB, while with  $B > 100$  b the same shift requires less than 0.5 dB. Secondly, there is a point where the system turns from being noise-limited to interference-limited (curves merging). Such a transition occurs for lower packet sizes the more simultaneous messages  $K$  there are.



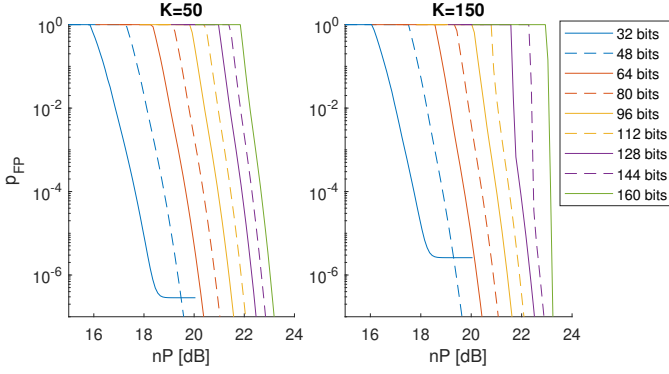


Fig. 4. Achievable physical layer error probability as a function of the energy per codeword and packet size  $B$ . Two different values of  $K$  are considered.

In Fig. 6 we combine all the earlier insights and look into the total probability of mis-authentication that takes into account both the physical and cryptographic layer performance. These results are obtained assuming a population of  $N = 10^5$  users and  $K = 100$  messages. In the plots, the blue line represents a packet consisting solely of the information bits, i.e.  $B = D$ . Since there is no additional means of authentication, every decoded packet is accepted and hence  $p_{\text{miss\_auth}} = p_{\text{FP}}$ . The red line denotes our proposed scheme in which the packet consists of information bits and a MAC, that is, in total  $B = D + L$ , where  $L = 32$  b is fixed. The values reported here correspond to the exhaustive search variant, hence, the total probability of mis-authentication is  $p_{\text{miss\_auth}} = p_{\text{FP}} p_{\text{fp\_auth}}$  with  $p_{\text{fp\_auth}}$  given by (5). Lastly, the yellow curve represents the classic packet structure, where the address is also included (here  $A = 32$  b as well) which yields  $B = D + L + A$ . In such case, the receiver checks only one key corresponding to the given address, hence we have that  $p_{\text{miss\_auth}} = p_{\text{FP}}$ . It is important to keep in mind that the most basic mode of operation (blue) does not provide any way of identifying the users, and as such it is not directly comparable with the other two. Furthermore, it might not provide sufficient level of reliability when the packets are very short, which is due to the floor on  $p_{\text{FP}}$ . What might be surprising, is that the classic packet structure actually performs slightly worse than our proposed scheme (at least until  $10^{-20}$  level which should be more than enough). This is because even though the probability of MAC collision is significantly lower, the packet needs to be larger to accommodate the address, which requires higher energy.

## VI. CONCLUSIONS

In this paper we proposed a method to introduce identification and authentication capabilities to algorithms that follow the framework of unsourced random access. Our scheme adds very limited amount of metadata to the communication, which is especially important for short IoT packets. Furthermore, as a consequence of not including explicit user identification, the packets are fully anonymous to everyone except the BS, which opens the door to new use cases and applications. This is in contrast to traditional protocols, where only the message content is assumed to be secret while the identities are public. The extra functionalities come at the cost of increased

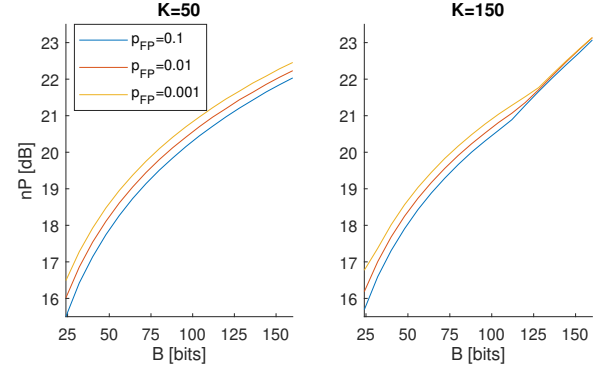


Fig. 5. Minimum energy required to achieve fixed physical layer error probability as a function of the number of information bits  $B$ .

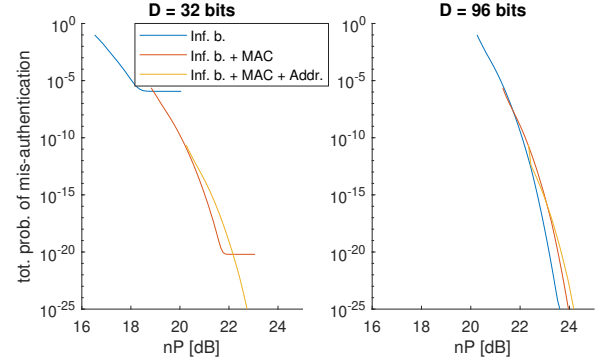


Fig. 6. Total probability of mis-authentication as a function of energy, taking into account both physical and cryptographic layer. The number of messages is  $K = 100$  and  $N = 10^5$ .

processing at the receiver. However, our results show that by avoiding the address we can simultaneously improve the spectral efficiency, and, for a given given energy per codeword, decrease the overall mis-authentication probability compared to the case where the address is included in the packet.

## REFERENCES

- [1] C. Bockelmann *et al.*, “Towards massive connectivity support for scalable mMTC communications in 5G networks,” *IEEE Access*, vol. 6, pp. 28 969–28 992, 2018.
- [2] N. Abramson, “THE ALOHA SYSTEM: another alternative for computer communications,” in *Proceedings of the November 17-19, 1970, fall joint computer conference*. ACM, pp. 281–285, 1970.
- [3] X. Chen, T.-Y. Chen, and D. Guo, “Capacity of gaussian many-access channels,” *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3516–3539, 2017.
- [4] Y. Polyanskiy, “A perspective on massive random-access,” in *IEEE Int. Symp. Inf. Theory (ISIT)*, June 2017, pp. 2523–2527.
- [5] G. Durisi, T. Koch, and P. Popovski, “Toward massive, ultrareliable, and low-latency wireless communication with short packets,” *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, Sept 2016.
- [6] S. S. Kowshik and Y. Polyanskiy, “Fundamental limits of many-user MAC with finite payloads and fading,” *IEEE Transactions on Information Theory*, vol. 67, no. 9, pp. 5853–5884, 2021.
- [7] A. Fengler, G. Caire, P. Jung, and S. Haghighatshoar, “Massive MIMO unsourced random access,” *preprint arXiv:1901.00828*, 2019.
- [8] K. Stern, A. E. Kalør, B. Soret, and P. Popovski, “Massive random access with common alarm messages,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, July 2019, pp. 1–5.
- [9] ETSI, “Security architecture and procedures for 5G system,” TS 133 501 v16.6.0, 2021.
- [10] A. Decurninge, I. Land, and M. Guillaud, “Tensor-based modulation for unsourced massive random access,” *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 552–556, 2021.