U-CIMAN: Uncover Spectrum and User Information in LTE Mobile Access Networks

Rui Zou and Wenye Wang
Department of Electrical and Computer Engineering
NC State University
Raleigh, NC 27606, USA
Emails: {rzou, wwang}@ncsu.edu

***Abstract*—With the proliferation of Dynamic Spectrum Access (DSA), Internet of Things (IoT), and Mobile Edge Computing (MEC) technologies, various methods have been proposed to deduce key network and user information in cellular systems, such as available cell bandwidths, as well as user locations and mobility. Not only is such information dominated by cellular networks of vital significance on other systems co-located spectrum-wise and/or geographically, but applications within cellular systems can also benefit remarkably from inferring such information, as exemplified by the endeavours made by video streaming to predict cell bandwidth. Hence, we are motivated to develop a *new* tool to uncover as much information used to be closed to outsiders or user devices as possible with off-the-shelf products. Given the wide-spread deployment of LTE and its continuous evolution to 5G, we design and implement U-CIMAN, a client-side system to accurately *U*n*C*over as much *I*nformation in *M*obile *A*ccess *N*etworks as allowed by LTE encryption. Among the many potential applications of U-CIMAN, we highlight one use case of accurately measuring the spectrum tenancy of a commercial LTE cell. Besides measuring spectrum tenancy in unit of resource blocks, U-CIMAN discovers user mobility and traffic types associated with spectrum usage through decoded control messages and user data bytes. We conduct 4-month detailed accurate spectrum measurement on a commercial LTE cell, and the observations include the predictive power of Modulation and Coding Scheme on spectrum tenancy, and channel off-time bounded under 10 seconds, to name a few.**

I. Introduction

Cellular access networks are becoming increasingly co-located with other wireless systems which often need knowledge of network or user information in cellular systems for various purposes [1]–[4]. For instance, obtaining the spectrum tenancy of cellular networks is important for secondary users in DSA systems to avoid radio channel collisions [1]. Moreover, traces of mobile phone calls are analyzed to locate cellular users to facilitate location based services provided by IoT systems [2]. Additionally, to improve the throughput of cache servers in MEC, a novel cache scheme is proposed to consider the mobility of cellular end users [3].

Not only is cellular network and user information of interests to other co-located systems, but many applications within the cellular system itself also strive to uncover such knowledge which is in low layers and thus not directly accessible [5]–[8]. As the mobile application that generates the largest portion of data traffic, video streaming over wireless links suffer from the highly variable bandwidth, so an algorithm is proposed to predict available bandwidth and encode video contents accordingly [5]. The importance of accurate estimation of available cell bandwidth is also recognized in [6], and it proposes to infer available bandwidth by measuring the energy level of LTE radio resources on the client side to improve video streaming qualities. To save the energy caused by peer discovery in Device-to-Device (D2D) communications, the concept of the proximity area of mobile users is proposed to search peer devices only when success probability is high [7], which utilizes the knowledge of user location to facilitate D2D communications. In another work [8], the knowledge of radio resource usage is achieved by the almost blank subframes to avoid intercell interference induced by D2D users.

Motivated by the huge demands on obtaining network and user information of cellular systems [1]–[8], and inspired by the existing efforts that seek such information from the air interface [6], [9], [10], we develop U-CIMAN to *U*n*C*over network and user *I*nformation in *M*obile *A*ccess *N*etworks. U-CIMAN aims to reveal non-transparent information in LTE access networks since LTE is the most widely deployed mobile system continuously evolving to 5G [11]. Due to data protection mechanisms [12], obtaining all the information aired by eNBs (eNodeBs, i.e., LTE base stations) is not possible. We target the unencrypted information transmitted by eNBs through passively listening to the downlink and decoding raw bytes of both control and data planes. Specifically, U-CIMAN decodes all control messages in the downlink physical layer and sniffs raw bytes of user data based on decoded control information. Compared with prior works [9], [10], U-CIMAN not only reliably unveils all the downlink control, but also user data that is not protected by the LTE encryption. It is worth mentioning that U-CIMAN should cause no security or privacy concerns because it does not jeopardize LTE encryption. The decoded data fields have broad impacts on various applications, such as the aforementioned DSA, D2D, MEC, and video streaming [1]–[8].

Among many potential applications of U-CIMAN, we showcase its capability of measuring spectrum tenancy due to its superiority over the classical energy detection approach

This work is partially supported by NSF CNS-1527696, NSF CNS-1824518 and ARO W911NF-15-2-0102.

[13]. Despite the notorious difficulty of choosing a proper threshold [14], the energy detection approach and its derivatives, such as geo-location databases [15]–[17], cannot accurately distinguish normal radio activities of PUs from those of unauthorized users whose amount will grow significantly [18]. Since U-CIMAN obtains spectrum assignment of an LTE cell from the allocations announced by the eNB, the detection of PU radio activities is not hindered by perceived radio power fluctuations caused by varying noise powers or interference from unauthorized users. Besides accuracy and robustness, U-CIMAN provides important details associated with spectrum tenancy, the location of spectrum occupants and the size of data carried by spectrum slices occupied by PUs. Such details of spectrum sensing have long been recognized as essential information for successful operations of SUs [19], but how to attain such information with energy detection has not been proposed as far as we know.

The design of U-CIMAN and its application to measuring spectrum tenancy of a commercial LTE cell face several challenges. First, existing methods that decipher LTE control messages are prone to errors, and there is no effective validation approach [9], [10]. To decode downlink control messages from eNB, the first step is to obtain Radio Network Temporary Identity (RNTI) by decoding Downlink Control Information (DCI). The current method is based on the structure of DCI messages with payload and trailing bits that are the XOR of RNTI and Cyclic Redundancy Check (CRC) of the payload. A key implicit assumption in such approaches is that DCIs are always received correctly, which however, forfeits the original purpose of error detection mechanism and undermines reliable data reception mechanisms [6]. Second, eNB logs are needed as ground truth to validate decoding accuracy. However, accessing logs of commercial eNBs is not trivial; thus, we set up testbed eNBs with verified supports for LTE specifications as an alternative, and the related details are addressed in section II that presents the necessary LTE background, design rationale, and U-CIMAN implementation and the validations. Lastly, implementing U-CIMAN requires the use of decent SDR devices as radio front ends, since receiving LTE data typically requires analog bandwidth of 20 MHz, a sampling rate of 30.72 MHz or higher, and multiple processing chains for multiple-input and multiple-output (MIMO).

The challenges are addressed in Section II after which we study the most fine-grained LTE spectrum usage, driven by the long-standing desire for accurate PU spectrum tenancy [20]. Compared with existing data, U-CIMAN greatly improves measurement granularity due to its time resolution of 1 ms, the minimum time interval of LTE scheduling, and frequency resolution of 180 kHz, the basic bandwidth unit in LTE spectrum resource assignment. Moreover, U-CIMAN adds essential details of user information associated with spectrum tenancy, including the packet size carried by resource blocks and Time Advance (TA) values of spectrum consumers. Furthermore, we have made several insightful observations many of which we believe are reported for the first time: spectrum tenancy is upper bounded to around 10 seconds, which is consistent with practical systems, but in contrast to analytic results of heavy-tailed distribution [21]–[23]; Modulation and Coding Scheme (MCS) of spectrum slices are highly indicative of occupancy status in the next time slot.

To sum up, our contributions are as follows. First, we propose the design and implementation of U-CIMAN that decodes both control and data plane packets at the physical layer of LTE downlinks. Second, three decoded data fields, DCIs, packet sizes, and TA values, are applied to study the spectrum tenancy of a commercial LTE cell. Compared with prior spectrum measurements, our results have much finer resolutions in both time and frequency domains, and details of spectrum usage including the packet sizes that reflect traffic types and TA values indicating rough user locations. Lastly, enabled by U-CIMAN, we measure the spectrum occupancy of an LTE cell for 4 months, and observe spectrum tenancy patterns that are influential on the design of DSA systems.

II. Design and implement of U-CIMAN

This section presents LTE background closely related to the design of U-CIMAN, and how U-CIMAN decodes control messages and user data bytes. Then, we explain how U-CIMAN is implemented to measure spectrum tenancy accurately, followed by thorough performance validations.

A. LTE primer

By introducing the relevant LTE domain knowledge, this primer explains why some messages aired by eNBs are possible to be understood by outsiders without causing privacy and security concerns, and why such messages can be exploited to achieve accurate LTE spectrum tenancy measurement.

Though LTE is designed with encryption mechanisms to protect data from eavesdropping while being transmitted over the air interface, not all parts of the messages can be protected. There are still two types of unprotected information. *Type I* is left in clear text due to their timing relative to the setup process of data encryption. There are several steps between a UE and the network side to undergo before setting up ciphered control and data channels, so the downlink information transmitted over the air before encryption setup has to be in clear text; *Type II* unencrypted information is found in the messages or headers generated in the protocol layers under the Packet Data Convergence Protocol (PDCP) that is responsible for encryption [24], so the information generated in layers below PDCP cannot be protected. Since both types of unprotected messages either carry information for initial access of an LTE cell or pertain only to operations below PDCP layer, their leakage can hardly be related to specific users whose identifiers in those situations are represented as RNTIs that change rapidly. Though the unencrypted messages cannot be linked to specific users, the information contained therein can still be of vital importance. For example, an SU only needs to know the spectrum tenancy of some LTE users, but do not need to differentiate among different PUs.

As we will showcase the application of U-CIMAN to LTE spectrum tenancy measurement, we explain one of the key

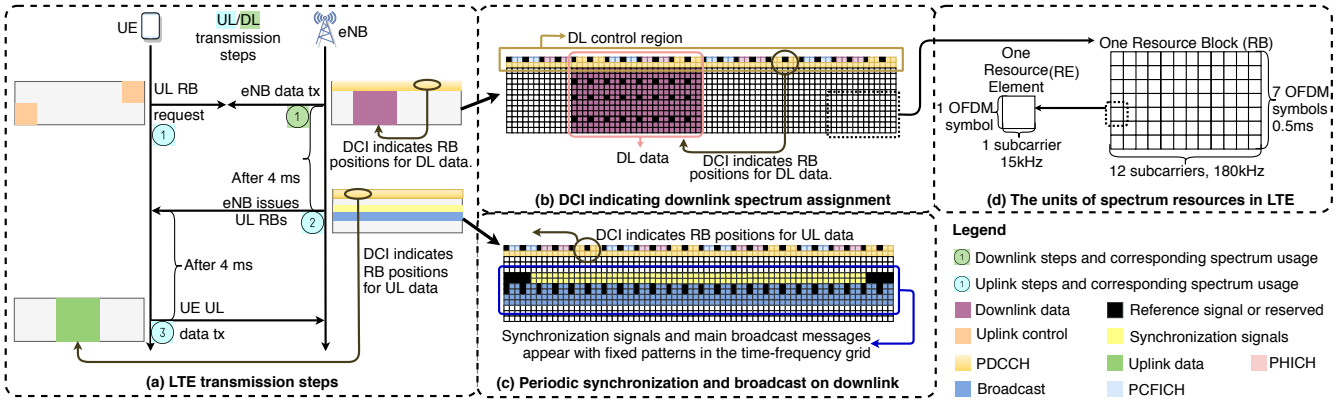

Fig. 1. The steps of LTE data transmissions and the structures of LTE spectrum resources.

unciphered LTE data fields, DCI, and how it reveals the spectrum tenancy. DCIs indicate the assignment of spectrum resources for both uplink and downlink, and the corresponding procedures are illustrated in Fig. 1(a) where the time grows vertically downwards and the frequency increases to the right horizontally. The uplink data transmission steps for an actively connected UE are numbered by blue circles. First, a UE that intends to transmit uplink data sends its request for spectrum resources. The request is carried by uplink control channels that do not require dynamic allocations. After receiving the request, the eNB schedules the uplink spectrum resources and puts the decision in a DCI to inform the UE. DCIs indicate the spectrum resources and MCS for transmissions. Finally, the UE receives the DCI destined for it, and then transmits in accordance with the information in the DCI. For downlink transmissions, only one step is needed as the data originates from the eNB. As shown by the circled green number in Fig. 1(a), the eNB sends the DCI and the corresponding data in the same Transmission Time Interval (TTI). Similar to DCIs for uplink transmission, DCIs for downlink data inform UEs where data is and the MCSs for demodulation. To receive downlink data, a UE blindly searches all possible locations for its DCIs. If DCIs pointing to downlink data are found, UEs locate and decode the spectrum slices according to the resource assignment and MCS values in the DCIs.

Measuring LTE spectrum tenancy by decoding DCIs achieves the finest possible granularity which is the same as the smallest unit of LTE scheduling for spectrum slices. The structure of LTE spectrum resources are shown in Fig. 1(d). In the time domain, a *subframe* is 1 ms, which is the time interval for an eNB to schedule spectrum resources. Thus, a subframe is also known as a TTI. In each subframe, the spectrum resources in two dimensional frequency-time grids are divided into Resource Blocks (RBs) that are the smallest unit of eNB resource assignment [25]. An RB is 180 kHz by 0.5 ms, and it comprises 12 subcarriers each of which typically carries 7 symbols. The smallest spectrum resource is called Resource Element (RE) that carries one symbol on one subcarrier [26]. Though the occupancy of the small amount of REs carrying control messages is not revealed by DCIs, this does not harm the accuracy of deciphering based measurement in general. In the downlink, REs whose occupancy is not assigned by DCIs are ones in the control region, used for broadcast and synchronization, as shown in Fig. 1(b) and (c). The occupancy of broadcast and synchronization REs are designed to be accessible to any LTE compliant devices, so eavesdropper can detect their tenancy in the same way as subscribers. The size of control region is announced in clear texts in the first symbol of every subframe.

In summary, outsiders to an LTE cell can overhear many messages aired by the eNB due to the existence of the two types of unencrypted information. Among the clear text messages, DCI assigns the spectrum resources by specifying the RB location and the corresponding MCS, which can be exploited for accurate measurement of spectrum tenancy.

B. The design of U-CIMAN

Motivated by the fact that DCIs contain RB assignments and the corresponding MCS values in clear texts, U-CIMAN is designed to achieve two goals, decoding the DCIs and raw bytes of the corresponding user data. For the first design goal, we will highlight the differences between our method and those of the others [9], [10]. By achieving the second goal, the unencrypted user data headers that are attached by layers below PDCP can provide essential details of spectrum occupancy, such as protocol fields of TA and the packet size. We would like to emphasize that the payload of user data originates from above the PDCP layer, so eavesdroppers cannot understand those raw bytes even if they achieve them from analogue radio signals properly through demodulation

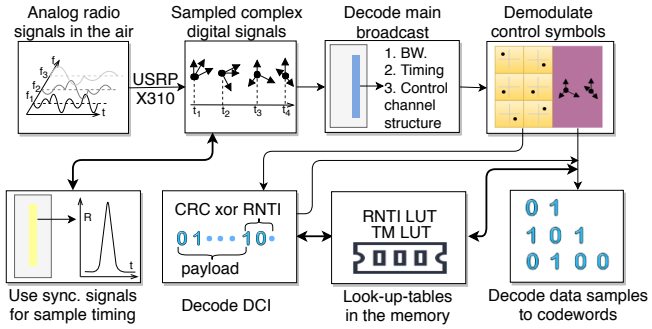

Fig. 2. Data processing steps of U-CIMAN.

and decoding. Thus, the design of U-CIMAN should cause no security breaches or privacy leakage.

The overall process to achieve these two goals is illustrated in Fig. 2, and the initial steps pose no major difficulties. First, the Software Defined Radio (SDR) front end Universal Software Radio Peripheral (USRP) X310 converts analog radio signals to complex samples, and sends them to the host computer. U-CIMAN in the host computer utilizes synchronization signals in LTE downlink to update sampling time and frequency range of the SDR. After achieving synchronization, U-CIMAN decodes main broadcast messages to discover system time, bandwidth, and downlink control channel structure which is then employed to locate REs carrying downlink control data. Because the Transmission Modes (TMs), i.e., the multiple antenna schemes, of control channels are known from decoding main broadcast messages and the modulation scheme of control channels is fixed as Quadrature Phase Shift Keying (QPSK), U-CIMAN is able to decode the complex samples of DCIs into raw bytes after obtaining the positions of the control channel REs. So far, these steps pose no challenges to outsiders to the cell because synchronization and main broadcast messages are designed to be decodable for any devices executing the corresponding LTE routines.

Challenges emerge in later steps where two user specific configurations, RNTI and TM, are required for verifying DCI messages and decoding raw bytes of downlink user data, respectively. For a normal UE, the eNB assigns an RNTI to the UE during the random access, and the RNTI is used for the validation of DCI decoding and the generation of scrambling sequences for user data protection against burst errors. Unlike downlink control messages transmitted at fixed locations, RNTI assignments are irregular and not necessarily adopted by UEs, making the direct decoding inefficient. Decoding TMs for user data is another challenge, because they are configured by the network side and transmitted to UEs through encrypted messages. With unknown multiple antenna configurations, i.e., the TMs, user data bytes cannot be decoded even if the corresponding DCI is attained. The last three steps in the second row of Fig. 2 illustrate how U-CIMAN obtains RNTIs and TMs to decode control messages and user data bytes, which is explained in later paragraphs in this subsection.

Despite the existence of prior works on RNTI and DCI decoding [9], [10], U-CIMAN differs from them in two important aspects. First, U-CIMAN ensures the validity of the decoded RNTIs. Moreover, U-CIMAN further decodes the raw bytes of user data in addition to RNTI, which is proposed for the first time. In [9], the DCI-based RNTI-derivation method is proposed, where the trailing bits after DCI payload are exploited. Because the last two bytes of DCIs are the XOR of the RNTI and the CRC checksum of DCI payload, RNTI can be obtained by computing the checksum and then XORing it with the last two bytes, assuming that the entire DCI is correctly decoded. This method has been verified to suffer from low reliability, because the internal LTE error detection mechanism for DCI is forfeited [6]. Another decoding method is proposed in [10], where the RNTIs are decoded in the random access stage when they are initially assigned and transmitted to users in Random Access Response (RAR) messages. However, the RNTIs contained in RARs are temporary, and may not necessarily be adopted by UEs [27].

To obtain RNTIs, U-CIMAN first collects a pool of potentially correct DCIs using both methods, decoding RARs and reverse engineering DCIs. Since these results contain invalid RNTIs, we further apply these RNTIs to the decoding of corresponding downlink user data. For an RNTI decoded from RAR messages or DCIs, if there is no subsequent decodable user data that corresponds to the RNTI in the next 10 ms, the RNTI is considered invalid. This is because RNTI is an input to decoding user data which cannot be achieved by erroneous RNTIs. In this way, RNTIs derived from RAR messages or downlink DCIs are validated by decoding user data, eliminating invalid RNTIs. As shown in Fig. 2, the RNTI Look-Up Table (LUT) stores initial results of decoded RNTIs and the ones that have been validated. When U-CIMAN decodes DCIs and user data bytes, it first tries the stored RNTIs before deducing them. If the RNTIs stored in the LUT do not yield successful decoding of DCIs or user data for 10 consecutive TTIs, the RNTIs are removed from the LUT.

For any device capable of LTE physical layer data processing whose functions are specified in the standards, successful decoding of complex samples into raw user data bytes purely depends on the availability of RNTIs, DCIs, and TMs [26], [28]. Now that U-CIMAN has obtained RNTIs and DCIs, it still needs to find TMs in order to achieve the raw user data bytes. Unfortunately, unlike RNTIs or DCIs, TMs are configured in a control message enciphered at PDCP layer, decoding it over the air interface is impossible without breaking LTE encryption. To uncover the TMs in this case, we utilize several LTE mechanisms to deduce TMs more efficiently than trying all possible TMs every time. One is the mapping between TMs and the formats of DCI as summarized in Table 9.2 in [25]. Since many DCI formats map to a very limit set of possible TMs, we use this information first to reduce the size of TM search space. The other LTE mechanism which helps TM inference is that TMs are reconfigured at a much lower rate than that of RNTIs, so we store the TMs corresponding to RNTIs in the LUT as well for later lookup. In this way, U-CIMAN obtains the TMs efficiently. Having achieved the RNTIs, DCIs, and TMs, U-CIMAN decodes user data bytes from the complex samples in the same way as a normal UE. U-CIMAN locates the samples of user data based on the information in DCIs, undoes the precoding and layer mapping according to TM, demodulates the symbols per the MCS incorporated in the DCIs, and unscrambles using the RNTI to obtain codewords. Further processing to convert codewords to raw user data bytes in the transport blocks requires no user specific configurations over the air interface. Though most decoded user data bytes are encrypted, headers added by lower layers are in clear texts and can be understood by U-CIMAN. To discover traffic and user mobility characteristics associated with spectrum tenancy, U-CIMAN exploits two types of user data headers, packet size and TA, and the corresponding

details are explained in the next two subsections.

C. The implementation of U-CIMAN

We implement the overall data flow, and solutions for the two challenges as described in the previous subsection. The implementation of U-CIMAN is facilitated by the open source LTE library srsLTE [29], and the functions for DCI decoding from OWL [10]. Besides implementing the main U-CIMAN design, we decode data fields relevant to LTE spectrum occupancy and record them in files. Three types of data fields are recorded, resource assignments, TA values, and packet sizes in bytes. Resource assignment fields provide the fine granularity of spectrum measurement. TA fields indicate rough locations of users, and the size of the physical layer packets, or codewords, reflects user traffic types.

TABLE I
DESCRIPTIONS FOR RECORDED DATA FIELDS.

SFN	System frame number, in $\{0, \cdots, 1023\}$
Subframe	Index of LTE subframe, in $\{0, \cdots, 9\}$
RNTI	User identifier, in $\{0, \cdots, 65535\}$
Direction	Uplink or downlink, in $\{0, 1\}$
MCS	Modulation, coding scheme, in $\{0, \cdots, 31\}$
Total	Total number of RBs, in $\{1, \cdots, 100\}$
RA type	Resource assignment types, in $\{0, 1, 2\}$
RA1	The first field indicating RB assignment
RA2	The second field indicating RB assignment
RA3	The third field indicating RB assignment
CFI	Size of PDCCH, in $\{1, 2, 3\}$
RAR TA	TA values in RAR, in $\{0, \cdots, 1282\}$
TA	TA updates, in $\{0, \cdots, 63\}$
Length	Packet size in bytes, $\{0, \cdots, 65535\}$

U-CIMAN is implemented to produce one data record per DCI. All possible data fields in a record are listed in Table I. The first 11 fields for spectrum tenancy are decoded from DCIs or other physical layer control channels. Most of them have been explained, except the four whose names include 'RA'. These fields describe the assigned RBs in the same way as DCIs do. Because there are three types of spectrum resource allocation in LTE, and each type adopts different data structures to indicate the assigned RBs, we use four 'RA' fields to record RB assignments. The field for resource allocation type, 'RA type', shows the type of LTE resource allocation, taking values from 0, 1, or 2. For RA type 0, 'RA 1' field is a bitmap indicating the allocated RBs, and 'RA 2' and 'RA 3' fields are left unused. For RA type 1, 'RA 1' field is a different type of bitmap that requires additional information called 'subset' and 'shift' that are stored in fields 'RA 2' and 'RA 3' to describe RB assignments. For RA Type 2, U-CIMAN stores the starting RB position in field 'RA 1' and the number of assigned RBs in 'RA 2'. For details on how to determine the exact index of occupied RBs in each allocation type, interested readers can find them in [28]. The record fields in the TA category include the 'RAR TA' and the 'TA'. 'RAR TA' is the initial TA value obtained from the RAR messages. 'TA' field is the TA update value decoded from the unencrypted headers in downlink user data. The payload size is the number of bytes in decoded downlink user packets at the physical layer. The top 11 fields in Table I are present for every record, while the others may not be. 'RAR TA' is only for RAR messages. TA updates are conducted by eNB at regular time intervals, so they are in the headers of some downwlink packets. The 'Length' field is nonempty when the DCI points to a downlink data packet whose raw bytes are decoded successfully.

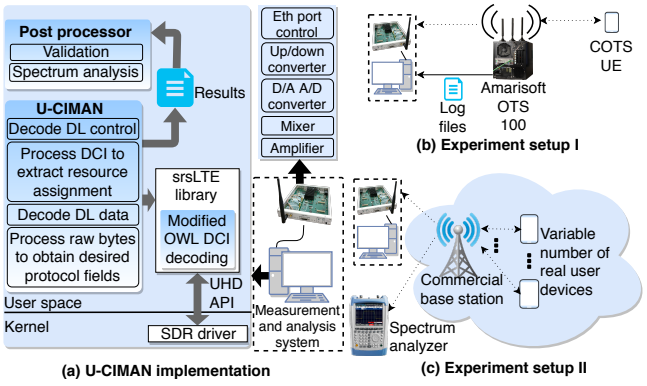

Fig. 3. Implementation and experiment setups.

As shown in Fig. 3(a), U-CIMAN is implemented in the user space of a Linux computer. The four main function blocks of U-CIMAN are decoding the raw bytes of DCI, parsing DCI messages, decoding raw bytes of user data, and parsing some unencrypted headers of user data. U-CIMAN calls the srsLTE library for existing LTE functions, including DCI decoding routines provided by OWL. The measurement results of U-CIMAN are written into files that are fed to post-processing scripts for performance validation and result analysis. Through the SDR driver API, U-CIMAN calls USRP Hardware Driver (UHD) [30] version 3.9.7 to communicate with the SDR front end. The SDR system includes a USRP X310 mother board [31] and two SBX-120 wide-band daughter-boards [32]. The SDR boards contain function blocks to convert analog signals to complex samples. The host computer has a quad-core CPU and 16 GB memory, running Ubuntu 16.04. A Gigabit Ethernet cable connects the host computer and the USRP.

D. Validation

Thorough experiments are conducted to demonstrate that U-CIMAN decodes downlink control messages and user data bytes with high accuracy. Based on decoded data, the application of U-CIMAN to spectrum tenancy measurement achieves accurate spectrum measurement at the frequency-time granularity of 180 kHz by 1 ms. Besides spectrum tenancy, packet sizes and TA values associated with occupied RBs are also revealed through decoding headers of user data. We also show that the distributions of packet sizes are highly indicative of traffic types and TA values reflect user mobility.

Experiments validating U-CIMAN performance are conducted in two setups shown in Fig. 3(b) and (c). Experiment setup I includes a Commercial Off-The-Shelf (COTS) UE, and a commercial grade LTE network realized by Amarisoft OTS 100 [33] that works as the core network and the eNB. This setup allows access to log files of the LTE network, so decoding accuracy can be validated. Though Amarisoft system

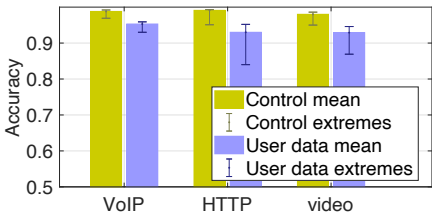

Fig. 4. Accuracy of U-CIMAN decoding.

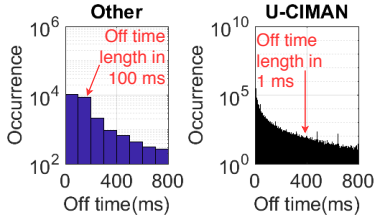

Fig. 5. Time granularity comparison.

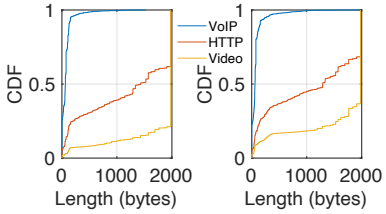

Fig. 6. Distributions of packet lengths.

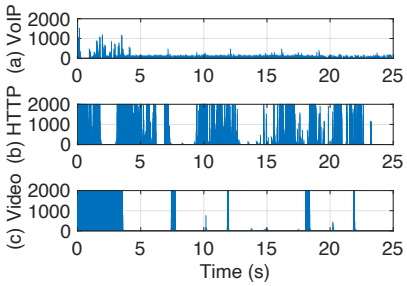

Fig. 7. Packet length under low traffic.

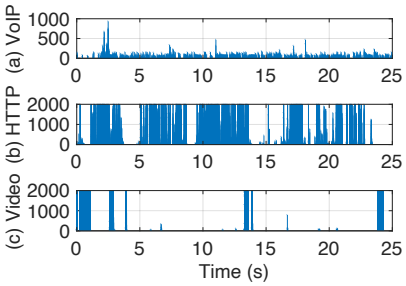

Fig. 8. Packet length under high traffic.

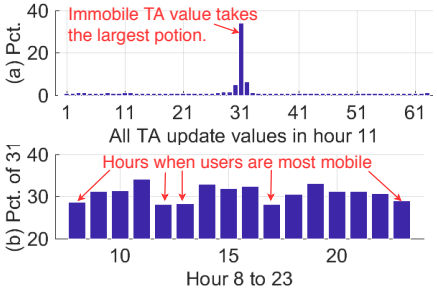

Fig. 9. Distributions of TA upates.

is the same as commercial LTE networks in terms of the compliance to the standards, it comes with one COTS UE and has limited radio coverage, which cannot reflect statistical characteristics of spectrum tenancy and mobility of real users in an LTE cell. To overcome these limitations, U-CIMAN is also validated by decoding the downlink of a commercial eNB owned by operators, which is experiment setup II.

We run three different applications on the UE multiple times in setup I, and compare the downlink control messages and user data bytes decoded by U-CIMAN with the ground truth in log files of Amarisoft OTS. The three applications each generate VoIP, HTTP, and video streaming traffic. Each application runs around 25 seconds in every round, and we ensure that no applications other than the one under test generate wireless traffic. Ten rounds of experiments are conducted in total. The accuracy for control messages is the ratio between the number of correctly decoded DCIs and the total number of DCIs. The accuracy for user data bytes is the ratio between the number of correctly decoded codewords and the total number of codewords. The results are shown in Fig. 4 where the bars show the mean values and the caps are the extremes. The yellow plots are the accuracy for control messages while the blue ones show the accuracy for user data. U-CIMAN achieves over 95% accuracy for control messages, and over 90% accuracy for user data. The results show that U-CIMAN is capable of decoding downlink control messages and user data with high accuracy for different traffic.

To demonstrate the benefits of fine measurement granularity, we compare the distributions of time length when an RB is not occupied, the off time, since it is the only LTE channel usage statistics at RB frequency granularity reported in other studies to our best knowledge. In [1], LTE spectrum occupancy is measured with good accuracy since the frequency resolution is the same with that of an LTE RB and time resolution is 100 ms. In comparison, U-CIMAN measures spectrum tenancy at time resolution of an LTE subframe, or 1 ms. Shown in Fig. 5 are off time distributions of RB 5 according to the measurement in [1], and that obtained by U-CIMAN in setup II. The system bandwidth in the two measurements are both 10 MHz. Off times over 800 ms are omitted due to their small percentage. Though 100 ms time resolution is good compared to most other measurements summarized in [20], the majority of the off time falls in only eight bins of 100ms. In comparison, the off time distribution achieved by our measurements is much more fine grained. Our off time distribution shows that off times are mostly under 30 ms, which cannot be observed with 100 ms time resolution. Thus, we claim that U-CIMAN provides an accessible way to achieve the highly accurate spectrum tenancy measurement which has long been desired [34].

Besides high accuracy, the spectrum tenancy measurements achieved by U-CIMAN also provide insight on mobile applications. Since codeword sizes are obtained by decoding user data, traffic types can be inferred from the distribution of the codeword sizes. We use setup I, and run three applications generating VoIP, HTTP, and video streaming traffic one at a time for around 25 seconds on a COTS UE. In the first scenario, there is no background traffic, so the only mobile traffic in the cell is generated by the COTS UE. We plot length of codewords in bytes versus time for a single run of the three applications in Fig. 7. For the VoIP traffic, there are a few large packets in the beginning, and the packets are short afterwards. The codeword sizes of HTTP traffic have a wide range. For video streaming, the packet sizes are mostly very large. We redo the same experiment in another scenario with heavy background traffic realized by adding one set of USRP and PC that emulates large amount of user traffic by running Amarisoft UE 100 [35]. According to Fig. 8, the codeword size versus time plots show similar trends to those in Fig. 7, so cell traffic load has little effects on codeword size characteristics of different applications. In addition to distinctive trends of codeword sizes along the time horizon, distributions of packet sizes of different traffic demonstrate clear separations as shown in Fig. 6. The plot on the left shows

the distribution of codeword sizes under heavy traffic, and the one on the right shows the distribution in light traffic. Though the increased traffic load slightly shifts packet sizes of HTTP and video traffic to the low end, packet size distributions of various applications show clear differences in both scenarios with diverse traffic loads. Hence, the decoded packet sizes at physical layer are highly indicative of user application types, regardless of varying cell traffic volume. Detecting PU spectrum occupancy and the traffic type at the same time can serve as the enabling function for many DSA proposals that are traffic pattern or application dependent [36]–[38].

To validate the correctness of TA data fields decoded by U-CIMAN, we conduct experiments in setup II where mobility is generated by real users. Due to the lack of actual UE coordinates, our experiments focus on the validation of TA updates which correspond to user mobility. In mobile networks, downlink synchronizations are achieved by mobile devices individually; on the uplink, however, the base station centrally adjusts the timing of user transmissions so that they arrive at the same time, making it possible for the base station to synchronize with all uplink transmissions at the same time. In LTE, uplink transmission time is adjusted by TA values that command UEs far from eNB to transmit with larger time advance than the nearby ones, so TA values reveal the distance of a UE to the eNB. According to LTE standards, the maximum cell radius is 100 km, mapping to the largest TA value of 1282, so the UE-eNB distance of around 78 meters maps to one in TA value. When a UE attempts random access, the eNB sets the initial TA value in RAR. Afterwards, the eNB adjusts TA with TA update commands placed in user data headers as needed due to user mobility. A TA update is a six-bit offset value, and the updated TA value is the sum of the original value, TA update, and -31. Thus, a TA update of 31 means that the UE remains its previous distance from the eNB. The farther the TA update values are from 31, the quicker the UE moves.

Fig. 9 shows two distributions of TA update values. The upper figure shows how TA update values in the 11th hour of a day distribute across all the 64 possible values, 0 to 63. The observation is that most TA updates are static or near static. Since TA update value 31 takes the largest portion, we plot how the percentage of TA update value 31 varies across different hours. The result in the lower plot agrees with life experience in that the portion of static users are small during commute hours 8, 12, 13, 17, and 23. Thus, U-CIMAN single handedly succeeds in obtaining the spectrum tenancy of all users in a cell together with their rough location and mobility as shown by TA values and TA updates, which requires less equipment than triangulation [9] while serving the needs of location based DSA algorithms [39], [40].

III. Spectrum tenancy measurement

The measurement setup is illustrated in Fig. 3(c). We search LTE bands assigned to operators near our lab, and find the cell with the strongest signal using a spectrum analyzer. The downlink bandwidth is 10 MHz, accommodating 50 LTE RBs. We collect LTE spectrum tenancy data of the cell in band 17 with U-CIMAN for four months, and conduct the post-processing to present the measurement results. Due to space limitations, only downlink results are presented. As far as we know, this is the first long time LTE spectrum occupancy measurement at RB granularity with packet sizes and TA values. The results are presented from three aspects, the total number of occupied RBs in each subframe, the frequency-wise spectrum tenancy, and the channel occupancy time. The observations and the driven forces behind them are discussed along the results. Since spectrum prediction is of fundamental significance to DSA systems and an important application of spectrum measurement [41], we highlight how our results are applicable to predictions of spectrum occupancy.

Total tenancy per subframe. Three metrics, the size of PDCCH, the number of DCIs, and the number of occupied RBs in each TTI are studied to reveal the controversy that the trends of the three metrics are alike in hourly average but their correlations are low in finer time scale. We define stochastic processes X_h^{PDH}, X_h^{DCI}, and X_h^{nRB}, where $h \in \{0, \cdots, 23\}$ is the time index of hours. For an hour h_0, the random variables, $X_{h_0}^{PDH}$, $X_{h_0}^{DCI}$, and $X_{h_0}^{nRB}$, are defined on the sample space $\Omega_{h_0}^{TT}$ where each element ω is the spectrum tenancy measurement for an LTE subframe in hour h_0.

PDCCH size is the number of symbols in a subframe assigned for DCIs, taking values from one to three. Thus, the probability mass function (pmf) for $X_{h_0}^{PDH}$ is $\mathbb{P}(X_{h_0}^{PDH} = x) = |\{\omega \in \Omega_{h_0}^{TT} | X_{h_0}^{PDH}(\omega) = x\}|/|\Omega_{h_0}^{TT}|$, where $x \in \{1, 2, 3\}$ and $|\cdot|$ is the size of a set. Fig. 10 shows distributions of $X_{h_0}^{PDH}$ before normalized by the size of the sample space. PDCCH sizes map to three colors, blue, green, and yellow. There are 3.6×10^6 subframes each hour, so $|\Omega_{h_0}^{TT}| = 3.6 \times 10^6$, which is the flat top in Fig. 10. Name time periods 3 to 6, 11 to 16, and 18 to 22 as stable periods s1, s2, and s3, and the hours in between them as transition periods t1, t2, and t3, as illustrated in Fig. 11 and 12. According to Fig. 10, the size of PDCCH is stable in s1 through s3, and changes smoothly between different levels during transition periods t1 through t3. The changes of the PDCCH sizes reflect varying spectrum activity levels that correspond to human activity intensities. In s2, the PDCCH sizes are the largest due to increased level of human activities during class time. There are lower levels of activities in s3 in the evening, because fewer people are on campus. During s1, the spectrum activities are at the lowest level.

Carried in PDCCH, DCIs are transmitted to convey spectrum resource assignments, so the number of DCIs in each subframe reflects the level of transmission activities. Define random process X_h^{DCI} as the number of DCIs in a TTI in an hour. Therefore, the pmf of $X_{h_0}^{DCI}$ for hour h_0 is $\mathbb{P}(X_{h_0}^{DCI} = x) = |\{\omega \in \Omega_{h_0}^{TT} | X_{h_0}^{DCI}(\omega) = x\}|/|\Omega_{h_0}^{TT}|$, where $x \in \mathbb{N}$. The distribution function is $F_{X_{h_0}^{DCI}}(x) = \sum_{x_i \le x} \mathbb{P}(X_{h_0}^{DCI} = x_i), x \ge 0$. Fig. 11 shows the minimum $F^{-1}_{X_{h_0}^{DCI}}(0) + 1$, the 15th percentile $F^{-1}_{X_{h_0}^{DCI}}(0.15)$, the mean $\mathbb{E}(X_{h_0}^{DCI})$, the

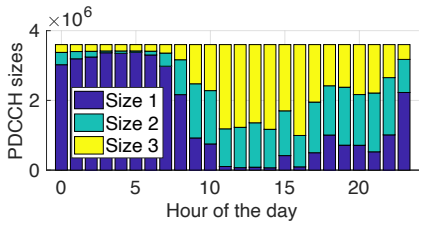

Fig. 10. PDCCH size distribution in each hour.

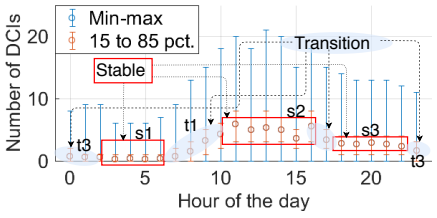

Fig. 11. Per hour statistics of DCIs.

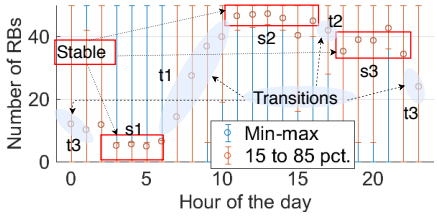

Fig. 12. Per hour statistics of RBs.

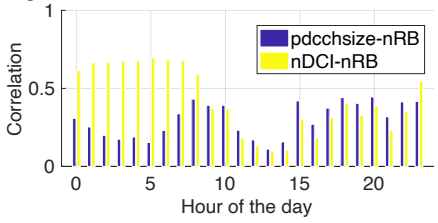

Fig. 13. Correlations between RB and 2 params.

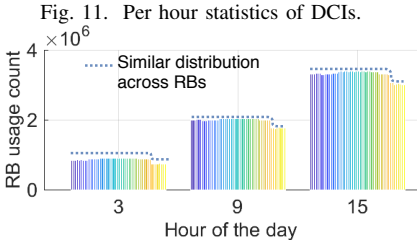

Fig. 14. Per hour total usage times of each RB.

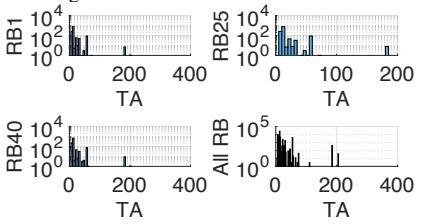

Fig. 15. TA distributions conditioned on RB.

85th percentile $F^{-1}_{X^{DCI}_{h_0}}(0.85)$, and the maximum $F^{-1}_{X^{DCI}_{h_0}}(1)$ for $h_0 \in \{0, \cdots, 23\}$, where $F^{-1}(\cdot)$ is the inverse function. When inverting the distribution function of a discrete random variable, $F^{-1}(x)$ may map to multiple values. In this case, the smallest value is chosen. The per hour average number of DCIs, $\mathbb{E}(X^{DCI}_{h_0})$, is shown by the red circle. The top red cap shows the 85th percentile, $F^{-1}_{X^{DCI}_{h_0}}(0.85)$, while the lower one is the 15th percentile, $F^{-1}_{X^{DCI}_{h_0}}(0.15)$. Blue caps are the extreme values on two ends.

To study the number of occupied RBs in a subframe, we consider the two RBs co-located frequency-wise as one and refer to them using the same RB index, because they have the same occupancy status. Thus, the total number of occupied RBs in a subframe ranges from 0 to 50, though there are technically 100 RBs per TTI in a 10 MHz LTE system. The pmf for $X^{nRB}_{h_0}$ in hour h_0 is $\mathbb{P}(X^{nRB}_{h_0} = x) = |\{\omega \in \Omega^{TT}_{h_0} | X^{nRB}_{h_0}(\omega) = x\}|/|\Omega^{TT}_{h_0}|$, where $x \in \{0, \cdots, 50\}$. The distribution function is $F_{X^{nRB}_{h_0}}(x) = \sum_{x_i \le x} \mathbb{P}(X^{nRB}_{h_0} = x_i)$, , where $x \ge -1$. Fig. 12 shows five hourly statistics of the number of occupied RBs per TTI, and the correspondence between markings in Fig. 12 and the statistics are the same with those in Fig. 11. The three stable periods and the three transition periods are also observed, and the phenomenon is again due to varying human activities.

Remark 1: The number of occupied RBs, the size of PDCCH, and the number of DCIs per subframe are determined by user activities, so the three metrics have very similar trends when averaged hourly, as shown in Fig. 10 to 12. In contrast, their correlations in millisecond time scale in most hours are low, as shown in Fig. 13. The cross correlation of two random processes at the same time index is $\rho_{X,Y}(h_0) = \mathbb{E}[(X_{h_0} - \mathbb{E}(X_{h_0}))(Y_{h_0} - \mathbb{E}(Y_{h_0}))]/(\sigma_X(h_0)\sigma_Y(h_0))$, where $\sigma_\cdot(h_0)$ is the standard deviation at time index h_0. The yellow bars illustrate $\rho_{X^{nRB},X^{DCI}}(h_0)$, and the blue bars show $\rho_{X^{nRB},X^{PDH}}(h_0)$. Since the correlations are below 0.5 except few hours, the size of PDCCH and the number of DCIs are weak predictors in millisecond level, but good at forecasting long time averages of occupied spectrum slices.

Tenancy characteristics along RBs. Having examined the spectrum occupancy along the time axis, we now study spectrum tenancy in different frequency. Define the random process X^{Chn}_h as the channel index of an occupied RB in an hour. The sample space for $X^{Chn}_{h_0}$ is the set of the indexes for all the occupied RBs $\Omega^{Chn}_{h_0}$ in hour h_0, and the probability $\mathbb{P}(X^{Chn}_{h_0} = x) = |\{\omega \in \Omega^{Chn}_{h_0} | X^{Chn}_{h_0}(\omega) = x\}|/|\Omega^{Chn}_{h_0}|$, where $x \in \{1, \cdots, 50\}$. Fig. 14 shows the distributions of $X^{Chn}_{h_0}$ before normalized by the size of sample space $\Omega^{Chn}_{h_0}$, where $h_0 \in \{3, 9, 15\}$. The bars stand for $|\{\omega \in \Omega^{Chn}_{h_0} | X^{Chn}_{h_0}(\omega) = x\}| = |\Omega^{Chn}_{h_0}|\mathbb{P}(X^{Chn}_{h_0} = x)$. Colored from blue to yellow, the bars represent the hourly usage times of RB 1 to 50. All RBs are used for similar number of times per hour. The evenly distributed total usage times suggest that the 50 RBs have similar channel conditions when evaluated hourly, and the eNB schedules them for similar times.

LTE scheduling does not favor any frequency slices, nor does it take user locations or traffic types into consideration. We consider the sample space Ω^{RP} of RBs allocated for 2000 randomly chosen packets each of which may be carried by multiple RBs. Define the random variable X^{TA} to be the TA value of a UE receiving packets carried by the RBs in the sample space. The pmf of X^{TA} is $\mathbb{P}(X^{TA} = x) = |\{\omega \in \Omega^{RP} | X^{TA}(\omega) = x\}|/|\Omega^{RP}|$, where $x \in \{0, \cdots, 1282\}$. The bars in the bottom right of Fig. 15 show the distribution of X^{TA} before normalized by the size of sample space Ω^{TA}. The heights of the bars are $|\{\omega \in \Omega^{RP} | X^{TA}(\omega) = x\}| = |\Omega^{RP}|\mathbb{P}(X^{TA} = x)$. The other three plots depict the distributions of X^{TA} conditioned on another random variable, the index of the RBs, X^{iRB}. The probability of TA values conditioned on RB indexes $\mathbb{P}(X^{TA} = x | X^{iRB} = y)$ is $\mathbb{P}(x|y) = |\{\omega \in \Omega^{RP} | X^{TA}(\omega) = x, X^{iRB}(\omega) = y\}|/|\{\omega \in \Omega^{RP} | X^{iRB}(\omega) = y\}|$, where $y \in \{1, \cdots, 50\}$. The conditional probability mass, $\mathbb{P}(X^{TA} | X^{iRB} = y_0)$ is similar to $\mathbb{P}(X^{TA})$, and three examples are given in Fig. 15 where $y_0 \in \{1, 25, 40\}$, showing that TA values are independent of RB indexes, i.e., $X^{TA} \perp\!\!\!\perp X^{iRB}$.

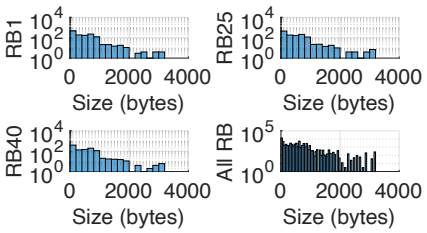

Fig. 16. Pkt. size distributions conditioned on RB.

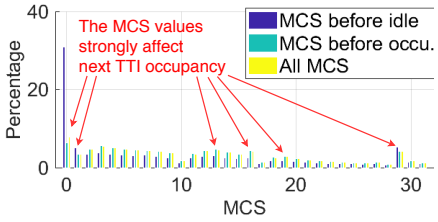

Fig. 17. Conditional MCS distributions.

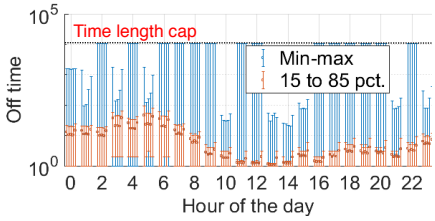

Fig. 18. 24-hour off time statistics of 5 RBs

Having studied the relationship between RB usage and UE-eNB distance, we investigate the impact of packet size on the frequency of assigned RBs. Define the random variable X^{PS} for packet sizes of RBs in the sample space Ω^{RP}. The packet size of an RB is the size of the packet carried by the RB. If multiple RBs carry the same packet, they are considered as having the same size. The pmf of X^{PS} is $\mathbb{P}(X^{PS} = x) = |\{\omega \in \Omega^{RP} | X^{PS}(\omega) = x\}|/|\Omega^{RP}|$, where $x \in \mathbb{N}^+$. Bars in the bottom right of Fig. 16 show the distribution of X^{PS} before normalized by the size of sample space Ω^{RP}, which is $|\{\omega \in \Omega^{RP} | X^{PS}(\omega) = x\}| = |\Omega^{RP}|\mathbb{P}(X^{PS} = x)$. The other three plots show distributions of X^{PS} conditioned on RB indexes of packets, X^{iRB}. The probability of packet sizes conditioned on RB indexes $\mathbb{P}(X^{PS} = x|X^{iRB} = y)$ is $\mathbb{P}(x|y) = |\{\omega \in \Omega^{RP} | X^{PS}(\omega) = x, X^{iRB}(\omega) = y\}|/|\{\omega \in \Omega^{RP} | X^{iRB}(\omega) = y\}|$, where $y \in \{1, \cdots, 50\}$. The conditional probability mass, $\mathbb{P}(X^{PS}|X^{iRB} = y_0)$ is similar to that of X^{PS}, and three examples are given in Fig. 16 where $y_0 \in \{1, 25, 40\}$, indicating that packet sizes do not depend RB indexes, $X^{PS} \perp\!\!\!\perp X^{iRB}$.

Thus, the RB occupancy does not depend on its frequency, UE-eNB distance, and packet size. However, MCS values strongly affect RB occupancy. MCS is specified along with spectrum resource assignments in DCIs. It is five-bit long, taking values from 0 to 31. Define X^M as the random variable for MCS values of RBs. The sample space is the set of all occupied RBs in a day, Ω^M. Define another random variable X^{NO} for RB occupancy in the next TTI on the same sample space. The pmf for X^M, $\mathbb{P}(X^M = x) = |\{\omega \in \Omega^M | X^M(\omega) = x\}|/|\Omega^M|, x \in \{0, \cdots, 31\}$, shown by yellow bars in Fig. 17 in percentage. The blue and the green bars show the probability of X^M conditioned on X^{NO} being 0 for unoccupied, or 1 for occupied in percentage. The conditional pmf, $\mathbb{P}(X^M = x|X^{NO} = y)$, is $\mathbb{P}(x|y) = |\{\omega \in \Omega^M | X^M(\omega) = x, X^{NO}(\omega) = y\}|/|\{\omega \in \Omega^M | X^{NO}(\omega) = y\}|$. Fig. 17 illustrates that MCS values of an occupied RB have different distributions conditioning on whether the RBs are occupied in the next TTI, and the differences are quite obvious at some MCS values. For example, the MCS values of 1 and 29 are much more likely to appear when RBs turn idle next TTI than in RBs continuing to be occupied. Hence, MCS values of currently occupied RBs strongly correlate with their tenancy in the next TTI, i.e. $X^M \not\perp\!\!\!\perp X^{NO}$.

Remark 2: We find that eNB treats spectrum resources on different frequency similarly, irrespective of UE-eNB distance, and packet sizes. Hence, they provide no information on the frequency of occupied spectrum. However, MCS values are highly indicative of channel occupancy in the next subframe.

Time characteristics of RB occupancy. We study the off time of an RB, defined as the time period from the TTI when the RB becomes idle to the subframe right before it turns occupied. Define a group of stochastic processes $X^{Off}_{c,h}$ for off time of all RBs. For every channel $c_0 \in \{1, \cdots, 50\}$, $X^{Off}_{c_0,h}$ is a stochastic process. For the hour $h_0 \in \{0, \cdots, 23\}$, $X^{Off}_{c_0,h_0}$ is a random variable on the sample space $\Omega^T_{c_0,h_0}$ where each element ω is an off time of RB c_0 in hour h_0. The pmf of $X^{Off}_{c_0,h_0}$ is $\mathbb{P}(X^{Off}_{c_0,h_0} = x) = |\{\omega \in \Omega^T_{c_0,h_0} | X^{Off}_{c_0,h_0}(\omega) = x\}|/|\Omega^T_{c_0,h_0}|, x \in \mathbb{N}^+$. The distribution function of $X^{Off}_{c_0,h_0}$ is $F_{X^{Off}_{c_0,h_0}}(x) = \sum_{x_i \le x} \mathbb{P}(X^{Off}_{c_0,h_0} = x_i), x \ge 0$. Fig. 18 presents the hourly statistics of off time for 5 RBs, $c_0 \in \{1, 13, 25, 37, 50\}$ in a day. The markings in Fig. 18 are the same with those in Fig. 11.

Remark 3: Observing Fig. 18, we find that the daily off time patterns also show the same stable and transition periods as those in Fig. 10 to 12. Thus, the off time is affected by the user activity levels as well. Another observation is that the off time distributions of different RBs in the same hour are similar. This is especially true for RB 1 and 50 in hour 7, because the four points, $x_0 \in \{0, 0.15, 0.85, 1\}$, in their distributions shown in Fig. 18 are almost the same, $F^{-1}_{X^{Off}_{1,7}}(x_0) \approx F^{-1}_{X^{Off}_{50,7}}(x_0)$. The similar off time among different RBs in the same hour stems from the scheduling that treats the spectrum resources in different frequency equally, as suggested by the even distributions of $X^{nRB}_{h_0}$ in Fig. 14. Still another observation is that off time exhibits a common upper bound for all hours and RBs. Fig. 18 shows the off time capped around 10^4 ms. The upper bound of off time is observed in all measurement data, not limited to off time in the day shown in Fig. 18.

IV. Conclusion

We design and implement a new LTE sniffing tool U-CIMAN that is capable of decoding both downlink control messages and raw bytes of user data. U-CIMAN is applied to the four-month measurement of spectrum tenancy of a commercial LTE cell. Compared with traditional measurements, our results are more accurate in terms of time-frequency granularity, and provide important details of spectrum users, such as the inferred traffic types and their rough locations. We observe spectrum occupancy characteristics from both time and frequency perspectives, and obtain valuable insights for spectrum tenancy predictions.

REFERENCES

[1] T. A. Hall, A. Sahoo, C. Hagwood, and S. Streett, "Exploiting lte white space using dynamic spectrum access algorithms based on survival analysis," in *Communications (ICC), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1–7.

[2] D. Zhang, S. Zhao, L. T. Yang, M. Chen, Y. Wang, and H. Liu, "Nextme: Localization using cellular traces in internet of things," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 2, pp. 302–312, 2015.

[3] X. Liu, J. Zhang, X. Zhang, and W. Wang, "Mobility-aware coded probabilistic caching scheme for mec-enabled small cell networks," *IEEE Access*, vol. 5, pp. 17 824–17 833, 2017.

[4] R. Zou and W. Wang, "Change detection based segmentation and modeling of LTE spectrum tenancy," in *2019 IEEE Global Communications Conference: Cognitive Radio and AI-Enabled Network Symposium (Globecom2019 CRAEN)*, Waikoloa, USA, Dec. 2019.

[5] A. Elgabli, V. Aggarwal, S. Hao, F. Qian, and S. Sen, "Lbp: Robust rate adaptation algorithm for svc video streaming," *IEEE/ACM Transactions on Networking*, vol. 26, no. 4, pp. 1633–1645, 2018.

[6] X. Xie, X. Zhang, S. Kumar, and L. E. Li, "pistream: Physical layer informed adaptive video streaming over lte," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015, pp. 413–425.

[7] A. Prasad, A. Kunz, G. Velev, K. Samdanis, and J. Song, "Energy-efficient d2d discovery for proximity services in 3gpp lte-advanced networks: Prose discovery mechanisms," *IEEE vehicular technology magazine*, vol. 9, no. 4, pp. 40–50, 2014.

[8] L. Babun, M. Simsek, and I. Güvenc, "Intercell interference coordination for d2d discovery in lte-a hetnets," in *2014 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2014, pp. 2202–2207.

[9] S. Kumar, E. Hamed, D. Katabi, and L. Erran Li, "Lte radio analytics made easy and accessible," in *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 4. ACM, 2014, pp. 211–222.

[10] N. Bui and J. Widmer, "OWL: a Reliable Online Watcher for LTE Control Channel Measurements," in *ACM All Things Cellular (MobiCom Workshop)*, Nov. 2016.

[11] J. Lee, Y. Kim, Y. Kwak, J. Zhang, A. Papasakellariou, T. Novlan, C. Sun, and Y. Li, "Lte-advanced in 3gpp rel-13/14: an evolution toward 5g," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 36–42, 2016.

[12] L. He, Z. Yan, and M. Atiquzzaman, "Lte/lte-a network security data collection and analysis for security measurement: a survey," *IEEE Access*, vol. 6, pp. 4220–4242, 2018.

[13] B. Li, M. Sun, X. Li, A. Nallanathan, and C. Zhao, "Energy detection based spectrum sensing for cognitive radios over time-frequency doubly selective fading channels," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 402–417, 2015.

[14] S. Saleem and K. Shahzad, "Performance evaluation of energy detection based spectrum sensing technique for wireless channel," *International journal of multidisciplinary sciences and engineering*, vol. 3, no. 5, pp. 31–34, 2012.

[15] H. B. Yilmaz and T. Tugcu, "Location estimation-based radio environment map construction in fading channels," *Wireless communications and mobile computing*, vol. 15, no. 3, pp. 561–570, 2015.

[16] Y. Ma, X. Zhang, and Y. Gao, "Joint sub-nyquist spectrum sensing scheme with geolocation database over tv white space," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 3998–4007, 2017.

[17] A. Adebayo, D. B. Rawat, J. Li, and M. Garuba, "Group-query-as-a-service for secure dynamic spectrum access in geolocation-enabled database-driven opportunistic wireless communications in roar framework," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2018, pp. 842–847.

[18] J. Wang, W. Wang, and C. Wang, "Sas: Modeling and analysis of spectrum activity surveillance in wireless overlay networks," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2019, pp. 2143–2151.

[19] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer networks*, vol. 50, no. 13, pp. 2127–2159, 2006.

[20] M. Höyhtyä, A. Mämmelä, M. Eskola, M. Matinmikko, J. Kalliovaara, J. Ojaniemi, J. Suutala, R. Ekman, R. Bacchus, and D. Roberson, "Spectrum occupancy measurements: A survey and use of interference maps," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2386–2414, 2016.

[21] S. Sengottuvelan, J. Ansari, P. Mähönen, T. Venkatesh, and M. Petrova, "Channel selection algorithm for cognitive radio networks with heavy-tailed idle times," *IEEE Transactions on Mobile Computing*, vol. 16, no. 5, pp. 1258–1271, 2016.

[22] T. A. Hall, A. Sahoo, C. Hagwood, and S. Streett, "Dynamic spectrum access algorithms based on survival analysis," *IEEE transactions on cognitive communications and networking*, vol. 3, no. 4, pp. 740–751, 2017.

[23] H. Eltom, S. Kandeepan, R. J. Evans, Y. C. Liang, and B. Ristic, "Statistical spectrum occupancy prediction for dynamic spectrum access: a classification," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 29, 2018.

[24] *Evolved Universal Terrestrial Radio Access (E-UTRA); Packet Data Convergence Protocol (PDCP) specification (3GPP TS 36.323 version 14.3.0 Release 14)*, ETSI 3GPP, Jul. 2017, version 14.3.0.

[25] S. Sesia, M. Baker, and I. Toufik, *LTE-the UMTS long term evolution: from theory to practice*. John Wiley & Sons, 2011.

[26] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation (3GPP TS 36.211 version 14.2.0 Release 14)*, ETSI 3GPP, Apr. 2017, version 14.2.0.

[27] *Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification (3GPP TS 36.321 version 13.5.0 Release 13)*, ETSI 3GPP, Apr. 2017, version 13.5.0.

[28] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (3GPP TS 36.213 version 13.0.0 Release 13)*, ETSI 3GPP, May 2016, version 13.0.0.

[29] I. Gomez-Miguelez, A. Garcia-Saavedra, P. D. Sutton, P. Serrano, C. Cano, and D. J. Leith, "srslte: an open-source platform for lte evolution and experimentation," in *Proceedings of the Tenth ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation, and Characterization*. ACM, 2016, pp. 25–32.

[30] E. Research, "Uhd," November 2017. [Online]. Available: https://kb.ettus.com/UHD

[31] ——, "x310-kit," February 2019. [Online]. Available: https://www.ettus.com/all-products/x310-kit/

[32] ——, "Sbx120," February 2019. [Online]. Available: https://www.ettus.com/all-products/SBX120/

[33] Amarisoft, "Amarisoft ots 100," December 2017. [Online]. Available: https://www.amarisoft.com/2016/12/28/amarisoft-presents-amari-ots-100/

[34] R. Chen, J.-M. Park, and K. Bian, "Robust distributed spectrum sensing in cognitive radio networks," in *IEEE INFOCOM 2008-The 27th Conference on Computer Communications*. IEEE, 2008, pp. 1876–1884.

[35] Amarisoft, "Amarisoft ue 100," December 2017. [Online]. Available: https://www.amarisoft.com/products/test-measurements/amari-ue-simbox/

[36] M. E. Ahmed, J. B. Song, and Z. Han, "Traffic pattern-based reward maximization for secondary user in dynamic spectrum access," in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2013, pp. 291–296.

[37] J. Huang, H. Wang, Y. Qian, and C. Wang, "Priority-based traffic scheduling and utility optimization for cognitive radio communication infrastructure-based smart grid," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 78–86, 2013.

[38] M. Hoyhtya, S. Pollin, and A. Mammela, "Classification-based predictive channel selection for cognitive radios," in *2010 IEEE International Conference on Communications*. IEEE, 2010, pp. 1–6.

[39] H. Hu, H. Zhang, and N. Li, "Location-information-assisted joint spectrum sensing and power allocation for cognitive radio networks with primary-user outage constraint," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 658–672, 2015.

[40] X. Xu, L. Li, Y. Cai, X. Chen, and M. Zhao, "Transmission rate optimization in cooperative location-aware cognitive radio networks," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–6.

[41] G. Ding, Y. Jiao, J. Wang, Y. Zou, Q. Wu, Y.-D. Yao, and L. Hanzo, "Spectrum inference in cognitive radio networks: Algorithms and applications," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 150–182, 2017.