CHAPTER

ACCESS PROCEDURES

11

The previous chapters have described the LTE uplink and downlink transmission schemes. However, prior to transmission of data, the device needs to connect to the network. This chapter describes the procedures necessary for a device to be able to access an LTE-based network.

11.1 ACQUISITION AND CELL SEARCH

Before an LTE device can communicate with an LTE network it has to do the following:

- Find and acquire synchronization to a cell within the network.
- Receive and decode the information, also referred to as the *cell system information*, needed to communicate with and operate properly within the cell.

The first of these steps, often simply referred to as *cell search*, is discussed in this section. The next section then discusses, in more detail, the means by which the network provides the cell system information.

Once the system information has been correctly decoded, the device can access the cell by means of the random-access procedure as described in Section 11.3.

11.1.1 OVERVIEW OF LTE CELL SEARCH

A device does not only need to carry out cell search at power-up—that is, when initially accessing the system, but, to support mobility, it also needs to continuously search for, synchronize to, and estimate the reception quality of neighboring cells. The reception quality of the neighboring cells, in relation to the reception quality of the current cell, is then evaluated to conclude if a handover (for devices in RRC_CONNECTED) or cell reselection (for devices in RRC_IDLE) should be carried out.

LTE cell search consists of the following basic steps:

- Acquisition of frequency and symbol synchronization to a cell.
- Acquisition of frame timing of the cell—that is, determination of the start of the downlink frame.
- Determination of the physical-layer cell identity of the cell.

4G, LTE-Advanced Pro and The Road to 5G. http://dx.doi.org/10.1016/B978-0-12-804575-6.00011-X

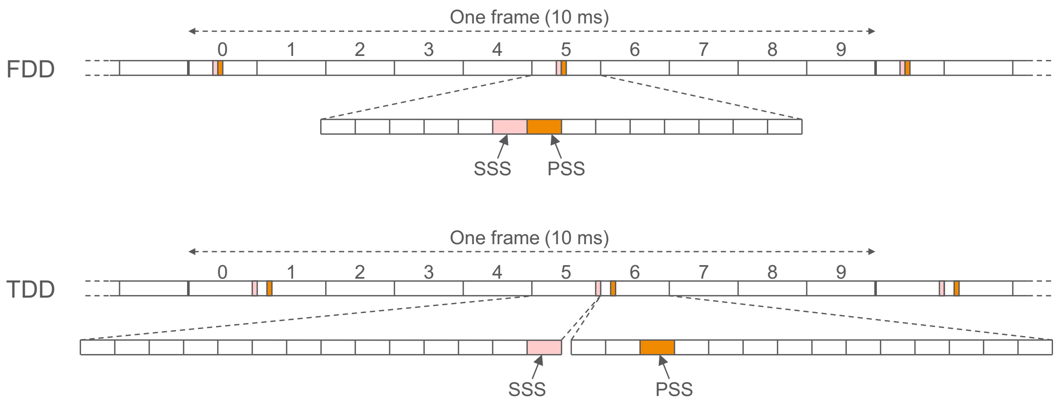

FIGURE 11.1

Time-domain positions of PSS and SSS in case of FDD and TDD.

As already mentioned, for example, in Chapter 6, there are 504 different physical-layer cell identities defined for LTE. The set of physical-layer cell identities is further divided into 168 cell-identity groups, with three cell identities within each group.

To assist the cell search, two special signals are transmitted on each downlink component carrier, the *primary synchronization signal* (PSS) and the *secondary synchronization signal* (SSS). Although having the same detailed structure, the time-domain positions of the synchronization signals within the frame differ somewhat depending on whether the cell is operating in FDD or TDD:

- In the case of FDD (upper part of Figure 11.1), the PSS is transmitted within the last symbol of the first slot of subframes 0 and 5, while the SSS is transmitted within the second last symbol of the same slot—that is, just prior to the PSS.
- In the case of TDD (lower part of Figure 11.1), the PSS is transmitted within the third symbol of subframes 1 and 6—that is, within the DwPTS—while the SSS is transmitted in the last symbol of subframes 0 and 5—that is, three symbols ahead of the PSS.

It should be noted that the difference in PSS/SSS time-domain structure between FDD and TDD allows for the device to detect the duplex mode of the acquired carrier if this is not known in advance.

Within one cell, the two PSSs within a frame are identical. Furthermore, the PSS of a cell can take three different values depending on the physical-layer cell identity of the cell. More specifically, the three cell identities within a cell-identity group always correspond to different PSS. Thus, once the device has detected and identified the PSS of the cell, it has found the following:

- Five-millisecond timing of the cell and thus also the position of the SSS which has a fixed offset relative to the PSS.[1]
- The cell identity within the cell-identity group. However, the device has not yet determined the cell-identity group itself—that is, the number of possible cell identities has been reduced from 504 to 168.

Thus, from the SSS, the position of which is known once the PSS has been detected, the device should find the following:

- Frame timing (two different alternatives given the found position of the PSS)
- The cell-identity group (168 alternatives)

Furthermore, it should be possible for a device to do this by the reception of one single SSS. The reason is that, for example, in the case when the device is searching for cells on other carriers, the search window may not be sufficiently large to cover more than one SSS.

To enable this, each SSS can take 168 different values corresponding to the 168 different cell-identity groups. Furthermore, the set of values valid for the two SSSs within a frame (SSS1 in subframe 0 and SSS2 in subframe 5) are different, implying that, from the detection of a single SSS, the device can determine whether SSS1 or SSS2 has been detected and thus determine frame timing.

Once the device has acquired frame timing and the physical-layer cell identity, it has identified the cell-specific reference signal. The behavior is slightly different depending on whether it is an initial cell search or cell search for the purpose of neighboring cell measurements:

- In the case of initial cell search—that is, the device state is in RRC_IDLE mode—the reference signal will be used for channel estimation and subsequent decoding of the BCH transport channel to obtain the most basic set of system information.
- In the case of mobility measurements—that is, the device is in RRC_CONNECTED mode—the device will measure the received power of the reference signal. If the measurement fulfills a configurable condition, it will trigger sending of a reference signal received power (RSRP) measurement report to the network. Based on the measurement report, the network will conclude whether a handover should take place. The RSRP reports can also be used for component carrier management,[2] for example, whether an

[1]This assumes that the device knows if it has acquired an FDD or a TDD carrier. Otherwise the device needs to try two different hypotheses regarding the SSS position relative to the PSS, thereby also indirectly detecting the duplex mode of the acquired carrier.

[2]As discussed in Chapter 5, the specifications use the terms primary and secondary "cells" instead of primary and secondary "component carriers."

additional component carrier should be configured or if the primary component carrier should be reconfigured.

11.1.2 PSS STRUCTURE

On a more detailed level, the three PSSs are three length-63 Zadoff−Chu (ZC) sequences (see Section 11.2) extended with five zeros at the edges and mapped to the center 73 subcarriers (center six resource blocks) as illustrated in Figure 11.2. It should be noted that the center subcarrier is actually not transmitted as it coincides with the DC subcarrier. Thus, only 62 elements of the length-63 ZC sequences are actually transmitted (element X_{32}^{PSS} is not transmitted).

The PSS thus occupies 72 resource elements (not including the DC carrier) in subframes 0 and 5 (FDD) and subframes 1 and 6 (TDD). These resource elements are then not available for transmission of DL-SCH.

11.1.3 SSS STRUCTURE

Similar to PSS, the SSS occupies the center 72 resource elements (not including the DC carrier) in subframes 0 and 5 (for both FDD and TDD). As described earlier, the SSS should be designed so that:

- The two SSSs (SSS1 in subframe 0 and SSS2 in subframe 5) take their values from sets of 168 possible values corresponding to the 168 different cell-identity groups.
- The set of values applicable for SSS2 is different from the set of values applicable for SSS1 to allow for frame-timing detection from the reception of a single SSS.

The structure of the two SSSs is illustrated in Figure 11.3. SSS1 is based on the frequency interleaving of two length-31-m sequences X and Y, each of which can take 31 different values (actually 31 different shifts of the same m-sequence). Within a cell, SSS2 is based on exactly the same two sequences as SSS1. However, the two sequences have been swapped in the frequency domain, as outlined in Figure 11.3. The set of valid combinations of X and Y for SSS1 has then been selected so that a swapping of the two sequences in the frequency

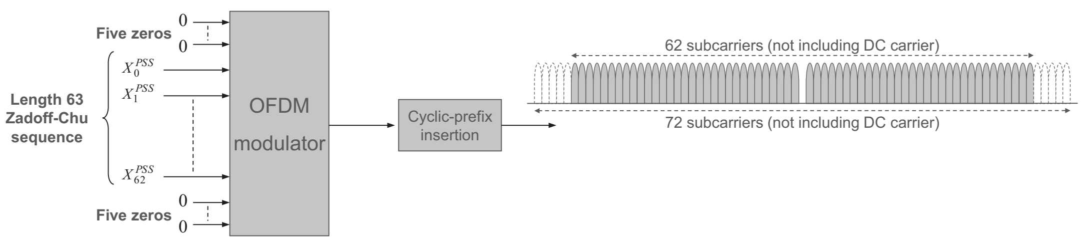

FIGURE 11.2

Definition and structure of PSS.

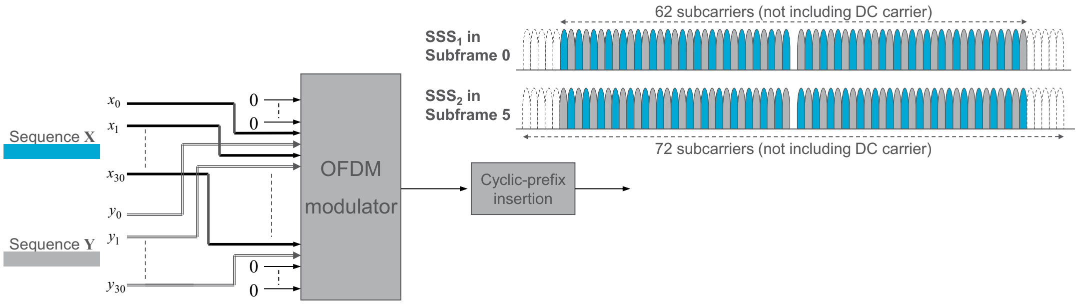

FIGURE 11.3

Definition and structure of SSS.

domain is not a valid combination for SSS1. Thus, the earlier described requirements are fulfilled:

- The set of valid combinations of X and Y for SSS1 (as well as for SSS2) are 168, allowing for detection of the physical-layer cell identity.
- As the sequences X and Y are swapped between SSS1 and SSS2, frame timing can be found.

11.2 SYSTEM INFORMATION

By means of the basic cell-search procedure described in Section 11.1, a device synchronizes to a cell, acquires the physical-layer identity of the cell, and detects the cell frame timing. Once this has been achieved, the device has to acquire the cell system information. This is information that is repeatedly broadcast by the network and which needs to be acquired by devices in order for them to be able to access and, in general, operate properly within the network and within a specific cell. The system information includes, among other things, information about the downlink and uplink cell bandwidths, the uplink/downlink configuration in the case of TDD, detailed parameters related to random-access transmission, and so on.

In LTE, system information is delivered by two different mechanisms relying on two different transport channels:

- A limited amount of system information, corresponding to the so-called *master-information lock* (MIB), is transmitted using the BCH.
- The main part of the system information, corresponding to different so-called *system-information blocks* (SIBs), is transmitted using the downlink shared channel (DL-SCH).

It should be noted that system information in both the MIB and the SIBs corresponds to the BCCH logical channel. Thus, as also illustrated in Figure 8.7, BCCH can be mapped to both BCH and DL-SCH depending on the exact BCCH information.

11.2.1 MIB AND BCH TRANSMISSION

As mentioned in the previous section, the MIB transmitted using BCH consists of a very limited amount of system information, mainly such information that is absolutely needed for a device to be able to read the remaining system information provided using DL-SCH. More specifically, the MIB includes the following information:

- Information about the downlink cell bandwidth. Three bits are available within the MIB to indicate the downlink bandwidth. Thus, up to eight different bandwidths, measured in number of resource blocks, can be defined for each frequency band.
- Information about the PHICH configuration of the cell. As mentioned in Section 6.4.2, the device must know the PHICH configuration to be able to receive the L1/L2 control signaling on PDCCH. The PDCCH information, in turn, is needed to acquire the remaining part of the system information which is carried on the DL-SCH, see later. Thus, information about the PHICH configuration (three bits) is included in the MIB—that is, transmitted using BCH, which can be received and decoded without first receiving any PDCCH.
- The system frame number (SFN) or, more exactly, all bits except the two least significant bits of the SFN are included in the MIB. As described in the following, the device can indirectly acquire the two least significant bits of the SFN from the BCH decoding.

The MIB also includes ten unused or "spare" information bits. The intention with such bits is to be able to include more information in the MIB in later releases while retaining backwards compatibility. As an example, LTE release 13 uses five of the ten spare bits to include additional eMTC-related information in the MIB, see Chapter 20 for more details.

BCH physical-layer processing, such as channel coding and resource mapping, differs quite substantially from the corresponding processing and mapping for DL-SCH outlined in Chapter 6.

As can be seen in Figure 11.4, one BCH transport block, corresponding to the MIB, is transmitted every 40 ms. The BCH transmission time interval (TTI) thus equals 40 ms.

The BCH relies on a 16-bit CRC, in contrast to a 24-bit CRC used for all other downlink transport channels. The reason for the shorter BCH CRC is to reduce the relative CRC overhead, having the very small BCH transport-block size in mind.

BCH channel coding is based on the same rate-1/3 tail-biting convolutional code as is used for the PDCCH control channel. The reason for using convolutional coding for BCH, rather than the Turbo code used for all other transport channels, is the small size of the BCH transport block. With such small blocks, tail-biting convolutional coding actually outperforms Turbo coding. The channel coding is followed by rate matching, in practice, repetition of the coded bits, and bit-level scrambling. QPSK modulation is then applied to the coded and scrambled BCH transport block.

BCH multi-antenna transmission is limited to transmit diversity—that is, SFBC in the case of two antenna ports and combined SFBC/FSTD in the case of four antenna ports. Actually, as

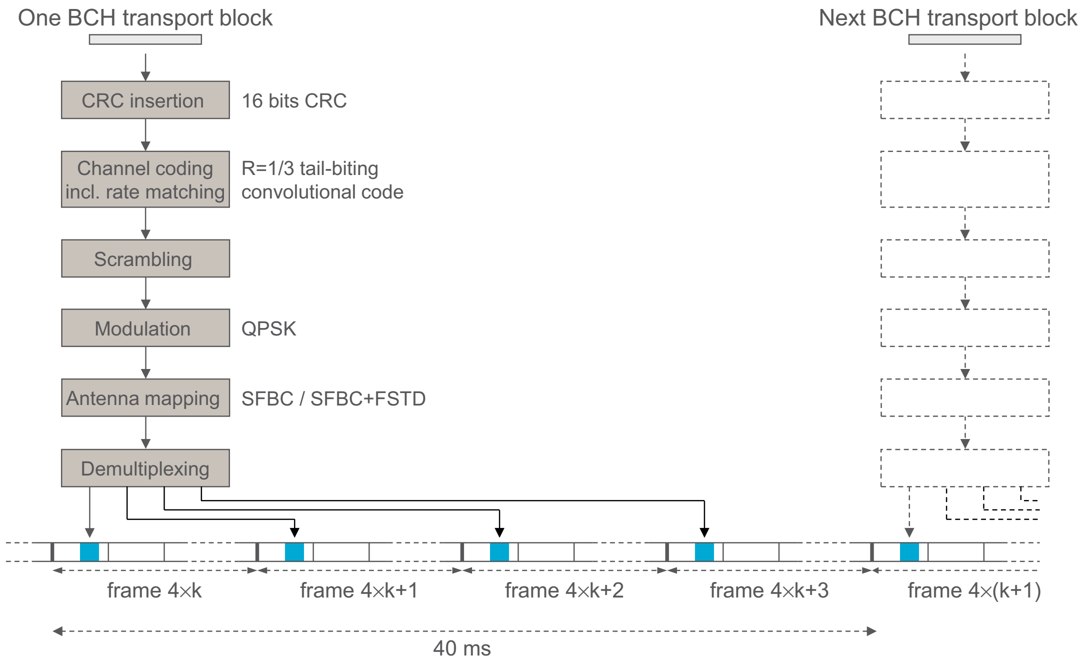

FIGURE 11.4

Channel coding and subframe mapping for the BCH transport channel.

mentioned in Chapter 6, if two antenna ports are available within the cell, SFBC must be used for BCH. Similarly, if four antenna ports are available, combined SFBC/FSTD must be used. Thus, by blindly detecting what transmit-diversity scheme is used for BCH, a device can indirectly determine the number of cell-specific antenna ports within the cell and also the transmit-diversity scheme used for the L1/L2 control signaling.

As can also be seen from Figure 11.4, the coded BCH transport block is mapped to the first subframe of each frame in four consecutive frames. However, as can be seen in Figure 11.5 and in contrast to other downlink transport channels, the BCH is not mapped on a resource-block basis. Instead, the BCH is transmitted within the first four OFDM symbols of the second slot of subframe 0 and only over the 72 center subcarriers.[3] Thus, in the case of FDD, BCH follows immediately after the PSS and SSS in subframe 0. The corresponding resource elements are then not available for DL-SCH transmission.

The reason for limiting the BCH transmission to the 72 center subcarriers, regardless of the cell bandwidth, is that a device may not know the downlink cell bandwidth when receiving BCH. Thus, when first receiving BCH of a cell, the device can assume a cell

[3]Not including the DC carrier.

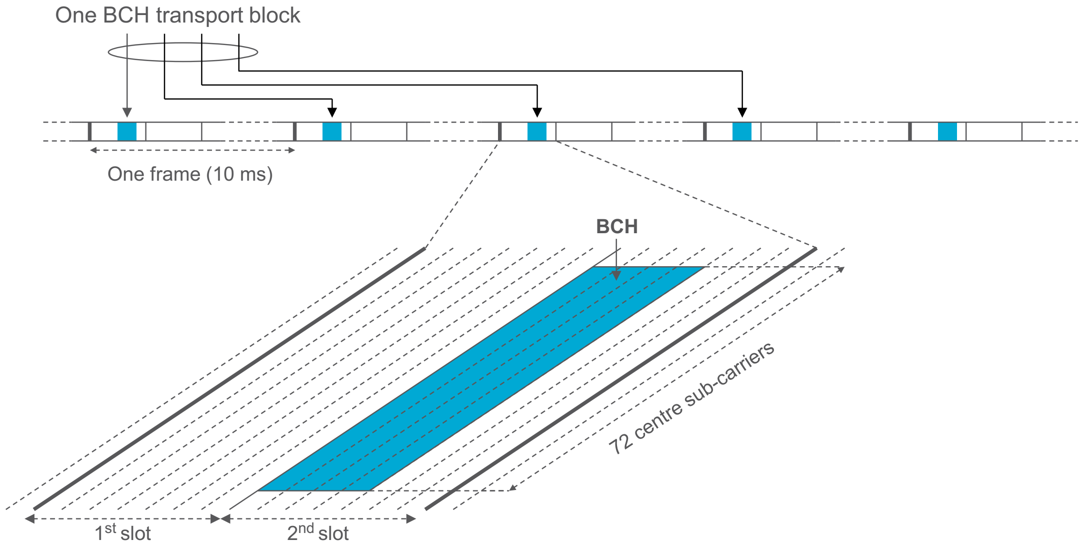

FIGURE 11.5

Detailed resource mapping for the BCH transport channel.

bandwidth equal to the minimum possible downlink bandwidth—that is, six resource blocks corresponding to 72 subcarriers. From the decoded MIB, the device is then informed about the actual downlink cell bandwidth and can adjust the receiver bandwidth accordingly.

The total number of resource elements to which the coded BCH is mapped is very large compared to the size of the BCH transport block, implying extensive repetition coding or, equivalently, massive processing gain for the BCH transmission. Such large processing gain is needed, as it should be possible to receive and correctly decode the BCH also by devices in neighboring cells, implying potentially very low receiver signal-to-interference-and-noise ratio (SINR) when decoding the BCH. At the same time, many devices will receive BCH in much better channel conditions. Such devices then do not need to receive the full set of four subframes over which a BCH transport block is transmitted to acquire sufficient energy for correct decoding of the transport block. Instead, already by receiving only a few or perhaps only a single subframe, the BCH transport block may be decodable.

From the initial cell search, the device has found only the cell frame timing. Thus, when receiving BCH, the device does not know to what set of four subframes a certain BCH transport block is mapped. Instead, a device must try to decode the BCH at four possible timing positions. Depending on which decoding is successful, indicated by a correct CRC check, the device can implicitly determine 40 ms timing or, equivalently, the two least significant bits of the SFN.[4] This is the reason why these bits do not need to be explicitly included in the MIB.

11.2.2 SYSTEM-INFORMATION BLOCKS

As already mentioned, the MIB on the BCH only includes a very limited part of the system information. The main part of the system information is instead included in different SIBs that are transmitted using the DL-SCH. The presence of system information on DL-SCH in a subframe is indicated by the transmission of a corresponding PDCCH marked with a special system-information RNTI (SI-RNTI). Similar to the PDCCH providing the scheduling assignment for "normal" DL-SCH transmission, this PDCCH also indicates the transport format and physical resource (set of resource blocks) used for the system-information transmission.

LTE defines a number of different SIBs characterized by the type of information that is included within them:

- SIB1 includes information mainly related to whether a device is allowed to camp on the cell. This includes information about the operator/operators of the cell, if there are restrictions with regard to which users may access the cell, and so on. SIB1 also includes information about the allocation of subframes to uplink/downlink and configuration of the special subframe in the case of TDD. Finally, SIB1 includes information about the time-domain scheduling of the remaining SIBs (SIB2 and beyond).
- SIB2 includes information that devices need in order to be able to access the cell. This includes information about the uplink cell bandwidth, random-access parameters, and parameters related to uplink power control.
- SIB3 mainly includes information related to cell reselection.
- SIB4—SIB8 include neighboring-cell-related information, including information related to neighboring cells on the same carrier, neighboring cells on different carriers, and neighboring non-LTE cells, such as WCDMA/HSPA, GSM, and CDMA2000 cells.
- SIB9 contains the name of the home-eNodeB.
- SIB10—SIB12 contain public warning messages, for example, earthquake information.
- SIB13 contains information necessary for MBMS reception (see also Chapter 19).
- SIB14 is used to provide enhanced access barring information, controlling the possibilities for devices to access the cell.
- SIB15 contains information necessary for MBMS reception on neighboring carrier frequencies.
- SIB16 contains information related to GPS time and coordinated universal time (UTC).
- SIB17 contains information related to interworking between LTE and WLAN

[4]BCH scrambling is defined with 40 ms periodicity, hence even if the device successfully decodes the BCH after observing only a single transmission instant, it can determine the 40 ms timing.

- SIB18 and SIB19 contain information related to sidelink (direct device-to-device) connectivity, see Chapter 21.
- SIB20 contains information related to single-cell point to multipoint, see Chapter 19.

Not all the SIBs need to be present. For example, SIB9 is not relevant for an operator-deployed node and SIB13 is not necessary if MBMS is not provided in the cell.

Similar to the MIB, the SIBs are broadcasted repeatedly. How often a certain SIB needs to be transmitted depends on how quickly devices need to acquire the corresponding system information when entering the cell. In general, a lower-order SIB is more time critical and is thus transmitted more often compared to a higher-order SIB. SIB1 is transmitted every 80 ms, whereas the transmission period for the higher-order SIBs is flexible and can be different for different networks.

The SIBs represent the basic system information to be transmitted. The different SIBs are then mapped to different system-information messages (SIs), which correspond to the actual transport blocks to be transmitted on DL-SCH. SIB1 is always mapped, by itself, on to the first system-information message SI-1,[5] whereas the remaining SIBs may be group-wise multiplexed on to the same SI subject to the following constraints:

- The SIBs mapped to the same SI must have the same transmission period. Thus, as an example, two SIBs with a transmission period of 320 ms can be mapped to the same SI, whereas an SIB with a transmission period of 160 ms must be mapped to a different SI.
- The total number of information bits that are mapped to a single SI must not exceed what is possible to transmit within a transport block.

It should be noted that the transmission period for a given SIB might be different in different networks. For example, different operators may have different requirements concerning the period when different types of neighboring-cell information needs to be transmitted. Furthermore, the amount of information that can fit into a single transport block very much depends on the exact deployment situation, such as cell bandwidth, cell size, and so on.

Thus, in general, the SIB-to-SI mapping for SIBs beyond SIB1 is flexible and may be different for different networks or even within a network. An example of SIB-to-SI mapping is illustrated in Figure 11.6. In this case, SIB2 is mapped to SI-2 with a transmission period of 160 ms. SIB3 and SIB4 are multiplexed into SI-3 with a transmission period of 320 ms, whereas SIB5, which also requires a transmission period of 320 ms, is mapped to a separate SI (SI-4). Finally, SIB6, SIB7, and SIB8 are multiplexed into SI-5 with a transmission period of 640 ms. Information about the detailed SIB-to-SI mapping, as well as the transmission period of the different SIs, is provided in SIB1.

Regarding the more detailed transmission of the different SIs there is a difference between the transmission of SI-1, corresponding to SIB1, and the transmission of the remaining SIs.

[5]Strictly speaking, as SIB1 is not multiplexed with any other SIBs, it is not even said to be mapped to an SI. Rather, SIB1 in itself directly corresponds to the transport block.

FIGURE 11.6

Example of mapping of SIBs to SIs.

The transmission of SI-1 has only a limited flexibility. More specifically, SI-1 is always transmitted within subframe 5. However, the bandwidth or, in general, the set of resource blocks over which SI-1 is transmitted, as well as other aspects of the transport format, may vary and is signaled on the associated PDCCH.

For the remaining SIs, the scheduling on DL-SCH is more flexible in the sense that each SI can, in principle, be transmitted in any subframe within time windows with well-defined starting points and durations. The starting point and duration of the time window of each SI are provided in SIB-1. It should be noted that an SI does not need to be transmitted on consecutive subframes within the time window, as is illustrated in Figure 11.7. Within the time window, the presence of system information in a subframe is indicated by the SI-RNTI on PDCCH, which also provides the frequency-domain scheduling as well as other parameters related to the system-information transmission.

Different SIs have different nonoverlapping time windows. Thus, a device knows what SI is being received without the need for any specific identifier for each SI.

In case of a relatively small SI and a relatively large system bandwidth, a single subframe may be sufficient for the transmission of the SI. In other cases, multiple subframes may be needed for the transmission of a single SI. In the latter case, instead of segmenting each SI into sufficiently small blocks that are separately channel coded and transmitted in separate subframes, the complete SI is channel coded and mapped to multiple, not necessarily consecutive, subframes.

Similar to the case of the BCH, devices that are experiencing good channel conditions may then be able to decode the complete SI after receiving only a subset of the subframes to which

FIGURE 11.7

Transmission window for the transmission of an SI.

the coded SI is mapped, while devices in bad positions need to receive more subframes for proper decoding of the SI. This approach has two benefits:

- Similar to BCH decoding, devices in good positions need to receive fewer subframes, implying the possibility for reduced device power consumption.
- The use of larger code blocks in combination with Turbo coding leads to improved channel-coding gain.

Strictly speaking the single transport block containing the SI is not transmitted over multiple subframes. Rather, the subsequent SI transmissions are seen as autonomous hybrid-ARQ retransmissions of the first SI transmission—that is, retransmissions taking place without any explicit feedback signaling provided on the uplink.

For devices capable of carrier aggregation, the system information for the primary component carrier is obtained as discussed. For secondary component carriers, the device does not need to read the SIBs but assumes that the information obtained for the primary component carrier also holds for the secondary component carriers. System information specific for the secondary component carrier is provided through dedicated RRC signaling as part of the procedure to configure an additional secondary component carrier. Using dedicated signaling instead of reading the system information on the secondary component carrier enables faster activation of secondary component carriers as the device otherwise would have to wait until the relevant system information had been transmitted.

11.3 RANDOM ACCESS

A fundamental requirement for any cellular system is the possibility for the device to request a connection setup, commonly referred to as random access. In LTE, random access is used for several purposes, including:

- for initial access when establishing a radio link (moving from RRC_IDLE to RRC_CONNECTED; see Chapter 4 for a discussion on different device states);
- to reestablish a radio link after radio-link failure;
- for handover when uplink synchronization needs to be established to the new cell;
- to establish uplink synchronization if uplink or downlink data arrives when the device is in RRC_CONNECTED and the uplink is not synchronized;
- for the purpose of positioning using positioning methods based on uplink measurements;
- as a scheduling request if no dedicated scheduling-request resources have been configured on PUCCH (see Chapter 9 for a discussion on uplink scheduling procedures).

Acquisition of uplink timing is a main objective for all these cases; when establishing an initial radio link (i.e., when moving from RRC_IDLE to RRC_CONNECTED), the random-access procedure also serves the purpose of assigning a unique identity, the C-RNTI, to the device.

Either a contention-based or a contention-free scheme can be used, depending on the purpose. Contention-based random access can be used for all previously discussed purposes, while contention-free random access can only be used for reestablishing uplink synchronization upon downlink data arrival, uplink synchronization of secondary component carriers, handover, and positioning. The basis for the random access is the four-step procedure illustrated in Figure 11.8, with the following steps:

1. The device transmission of a random-access preamble, allowing the network to estimate the transmission timing of the device. Uplink synchronization is necessary as the device otherwise cannot transmit any uplink data.
2. The network transmission of a timing advance command to adjust the device transmit timing, based on the timing estimate obtained in the first step. In addition to establishing uplink synchronization, the second step also assigns uplink resources to the device to be used in the third step in the random-access procedure.

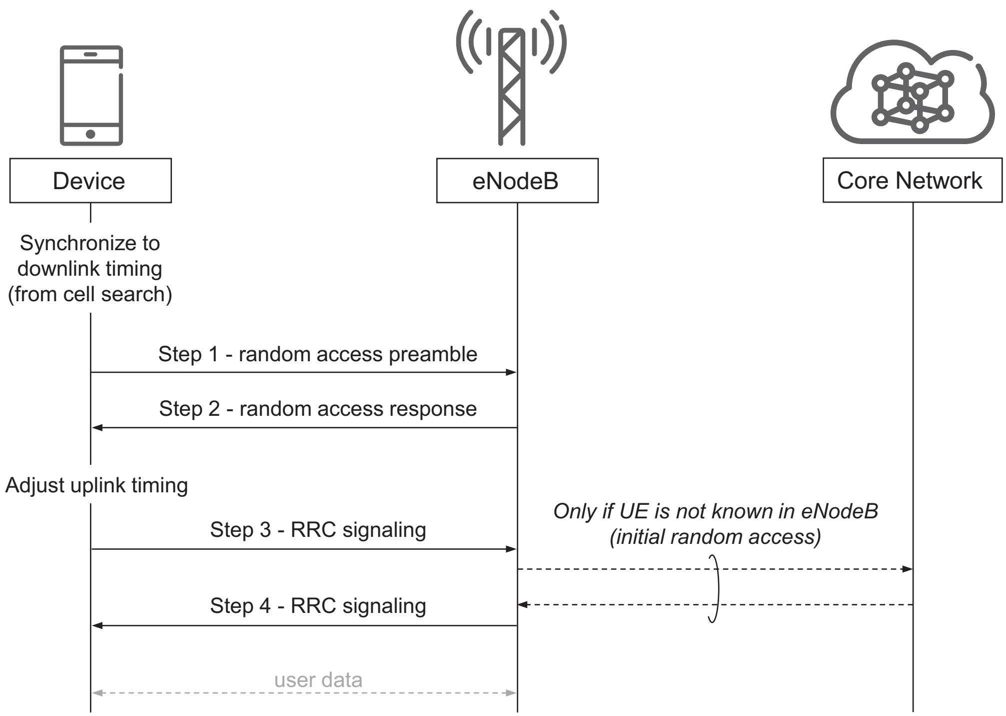

FIGURE 11.8

Overview of the random-access procedure.

3. The device transmission of the mobile-device identity to the network using the UL-SCH similar to normal scheduled data. The exact content of this signaling depends on the state of the device, in particular whether it is previously known to the network or not.
4. The network transmissions of a contention-resolution message from the network to the device on the DL-SCH. This step also resolves any contention due to multiple devices trying to access the system using the same random-access resource.

Only the first step uses physical-layer processing specifically designed for random access. The subsequent three steps utilize the same physical-layer processing as used for normal uplink and downlink data transmission. In the following, each of these steps is described in more detail. Only the first two steps of the preceding procedure are used for contention-free random access as there is no need for contention resolution in a contention-free scheme.

Both the device and the network can initiate a random-access attempt. In the latter case, RRC signaling or a so-called PDCCH order is used. A PDCCH order is a specific message transmitted on the PDCCH, containing information about when to initiate the random-access procedure and, in case of a contention-free random access, the preamble to use. PDCCH orders are primarily intended as a tool for the network to reestablish uplink synchronization but can be used for other purposes as well. A device may perform random access on its primary component carrier only[6] except for establishing uplink timing alignment for a secondary timing advance group.

11.3.1 STEP 1: RANDOM-ACCESS PREAMBLE TRANSMISSION

The first step in the random-access procedure is the transmission of a random-access preamble. The main purpose of the preamble transmission is to indicate to the base station the presence of a random-access attempt and to allow the base station to estimate the delay between the eNodeB and the device. The delay estimate will be used in the second step to adjust the uplink timing.

The time–frequency resource on which the random-access preamble is transmitted is known as the physical random-access channel (PRACH). The network broadcasts information to all devices in which time–frequency resource random-access preamble transmission is allowed (i.e., the PRACH resources), in SIB-2. As part of the first step of the random-access procedure, the device selects one preamble to transmit on the PRACH.

In each cell, there are 64 preamble sequences available. Two subsets of the 64 sequences are defined as illustrated in Figure 11.9, where the set of sequences in each subset is signaled as part of the system information. When performing a (contention-based) random-access attempt, the device selects at random one sequence in one of the subsets. As long as no other device is performing a random-access attempt using the same sequence at the same time

[6] The primary component carrier is device-specific as already discussed; hence, from an eNodeB perspective, random access may occur on multiple component carriers.

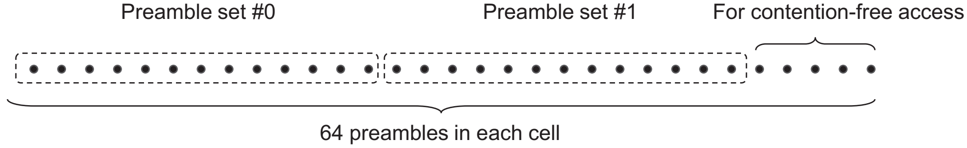

FIGURE 11.9

Preamble subsets.

instant, no collisions will occur and the attempt will, with a high likelihood, be detected by the eNodeB.

The subset to select the preamble sequence from is given by the amount of data the device would like to (and from a power perspective can) transmit on the UL-SCH in the third random-access step. Hence, from the preamble the device used, the eNodeB, will get some guidance on the amount of uplink resources to be granted to the device.

If the device has been requested to perform a contention-free random access, for example, for handover to a new cell, the preamble to use is explicitly indicated from the eNodeB. To avoid collisions, the eNodeB should preferably select the contention-free preamble from sequences outside the two subsets used for contention-based random access.

11.3.1.1 PRACH Time–Frequency Resources

In the frequency domain, the PRACH resource, illustrated in Figure 11.10, has a bandwidth corresponding to six resource blocks (1.08 MHz). This nicely matches the smallest uplink cell bandwidth of six resource blocks in which LTE can operate. Hence, the same random-access preamble structure can be used, regardless of the transmission bandwidth in the cell.

In the time domain, the length of the preamble region depends on configured preamble, as will be discussed later. The basic random-access resource is 1 ms in duration, but it is also possible to configure longer preambles. Also, note that the eNodeB uplink scheduler in principle can reserve an arbitrary long-random-access region by simply avoiding scheduling devices in multiple subsequent subframes.

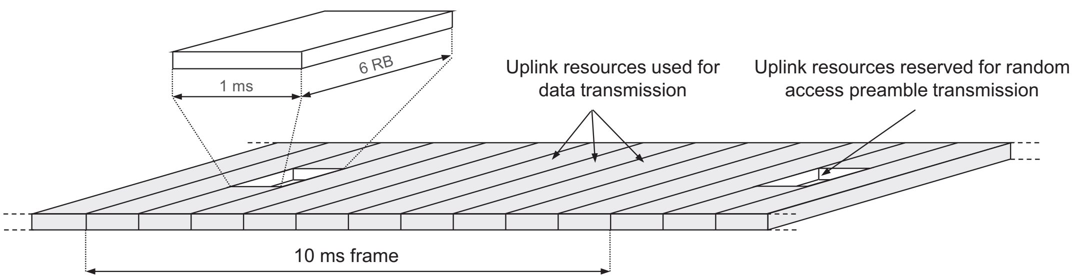

FIGURE 11.10

Principal illustration of random-access-preamble transmission.

Typically, the eNodeB avoids scheduling any uplink transmissions in the time–frequency resources used for random access, resulting in the random-access preamble being orthogonal to user data. This avoids interference between UL-SCH transmissions and random-access attempts from different devices. However, from a specification perspective, nothing prevents the uplink scheduler from scheduling transmissions in the random-access region. Hybrid-ARQ retransmissions are examples of this; synchronous nonadaptive hybrid-ARQ retransmissions may overlap with the random-access region and it is up to the implementation to handle this, either by moving the retransmissions in the frequency domain as discussed in Chapter 8 or by handling the interference at the eNodeB receiver.

For FDD, there is at most one random-access region per subframe—that is, multiple random-access attempts are not multiplexed in the frequency domain. From a delay perspective, it is better to spread out the random-access opportunities in the time domain to minimize the average waiting time before a random-access attempt can be initialized.

For TDD, multiple random-access regions can be configured in a single subframe. The reason is the smaller number of uplink subframes per radio frame in TDD. To maintain the same random-access capacity as in FDD, frequency-domain multiplexing is sometimes necessary. The number of random-access regions is configurable and can vary from one per 20 ms to one per 1 ms for FDD; for TDD up to six attempts per 10 ms radio frame can be configured.

11.3.1.2 Preamble Structure and Sequence Selection

The preamble consists of two parts:

- A preamble sequence
- A cyclic prefix

Furthermore, the preamble transmission uses a guard period to handle the timing uncertainty. Prior to starting the random-access procedure, the device has obtained downlink synchronization from the cell-search procedure. However, as uplink synchronization has not yet been established prior to random access, there is an uncertainty in the uplink timing[7] as the location of the device in the cell is not known. The uplink timing uncertainty is proportional to the cell size and amounts to 6.7 μs/km. To account for the timing uncertainty and to avoid interference with subsequent subframes not used for random access, a guard time is used as part of the preamble transmission—that is, the length of the actual preamble is shorter than 1 ms.

Including a cyclic prefix as part of the preamble is beneficial as it allows for frequency-domain processing at the base station (discussed later in this chapter), which can be advantageous from a complexity perspective. Preferably, the length of the cyclic prefix is approximately equal to the length of the guard period. With a preamble sequence length of approximately 0.8 ms, there is 0.1 ms cyclic prefix and 0.1 ms guard time. This allows for cell

[7]The start of an uplink frame at the device is defined relative to the start of a downlink frame received at the device.

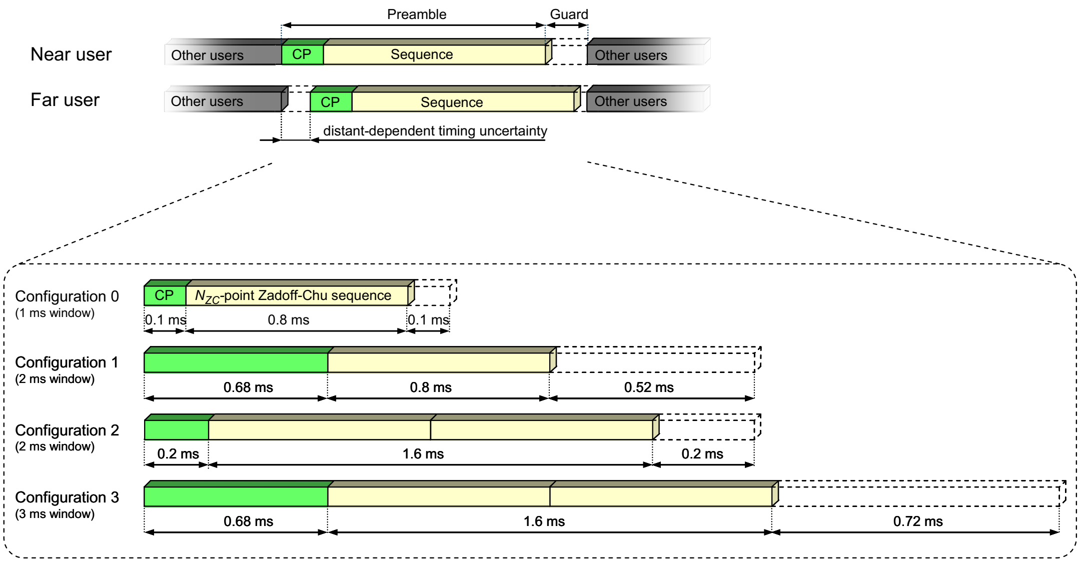

FIGURE 11.11

Different preamble formats.

sizes up to 15 km and is the typical random-access configuration, configuration 0 in Figure 11.11. To handle larger cells, where the timing uncertainty is larger, preamble configurations 1–3 can be used. Some of these configurations also support a longer preamble sequence to increase the preamble energy at the detector, which can be beneficial in larger cells. The preamble configuration used in a cell is signaled as part of the system information. Finally, note that guard times larger than those in Figure 11.11 can easily be created by not scheduling any uplink transmissions in the subframe following the random-access resource.

The preamble formats in Figure 11.11 are applicable to both FDD and TDD. However, for TDD, there is an additional fourth preamble configuration for random access. In this configuration, the random-access preamble is transmitted in the UpPTS field of the special subframe instead of in a normal subframe. Since this field is at most two OFDM symbols long, the preamble and the possible guard time are substantially shorter than the preamble formats described earlier. Hence, format 4 is applicable to very small cells only. The location of the UpPTS, next to the downlink-to-uplink switch for TDD, also implies that the interference from distant base stations may interfere with this short random-access format, which limits its usage to small cells and certain deployment scenarios.

11.3.1.3 PRACH Power Setting

The basis for setting the transmission power of the random-access preamble is a downlink path-loss estimate obtained from measuring the cell-specific reference signals on the primary

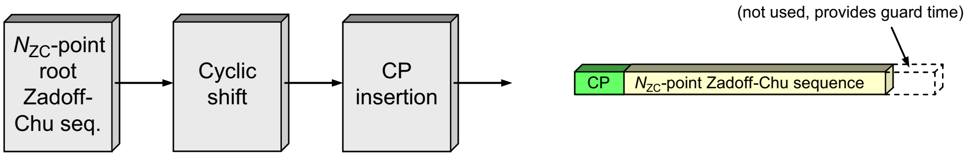

FIGURE 11.12

Random-access preamble generation.

downlink component carrier. From this path-loss estimate, the initial PRACH transmission power is obtained by adding a configurable offset.

The LTE random-access mechanism allows power ramping where the actual PRACH transmission power is increased for each unsuccessful random-access attempt. For the first attempt, the PRACH transmission power is set to the initial PRACH power. In most cases, this is sufficient for the random-access attempts to be successful. However, if the random-access attempt fails (random-access failures are detected at the second of four random-access steps, as described in the following sections), the PRACH transmission power for the next attempt is increased by a configurable step size to increase the likelihood of the next attempt being successful.

Since the random-access preamble is typically orthogonal to the user data, the need for power ramping to control intra-cell interference is smaller than in other systems with non-orthogonal random access and in many cases the transmission power is set such that the first random-access attempt with a high likelihood is successful. This is beneficial from a delay perspective.

11.3.1.4 Preamble Sequence Generation

The preamble sequences are generated from cyclic shifts of root Zadoff–Chu sequences [34]. Zadoff–Chu sequences are also used for creating the uplink reference signals as described in Chapter 7, where the structure of those sequences is described. From each root Zadoff–Chu sequence $X_{ZC}^{(u)}(k)$, $\lfloor N_{ZC}/N_{CS} \rfloor$ cyclically shifted[8] sequences are obtained by cyclic shifts of N_{CS} each, where N_{ZC} is the length of the root Zadoff–Chu sequence. The generation of the random-access preamble is illustrated in Figure 11.12. Although the figure illustrates generation in the time domain, frequency-domain generation can equally well be used in an implementation.

Cyclically shifted Zadoff–Chu sequences possess several attractive properties. The amplitude of the sequences is constant, which ensures efficient power amplifier utilization and maintains the low PAR properties of the single-carrier uplink. The sequences also have ideal cyclic auto-correlation, which is important for obtaining an accurate timing estimation at the eNodeB. Finally, the cross-correlation between different preambles based on cyclic

[8]The cyclic shift is in the time domain. Similar to the uplink reference signals and control signaling, this can equivalently be described as a phase rotation in the frequency domain.

shifts of the same Zadoff–Chu root sequence is zero at the receiver as long as the cyclic shift N_{CS} used when generating the preambles is larger than the maximum round-trip propagation time in the cell plus the maximum delay spread of the channel. Therefore, due to the ideal cross-correlation property, there is no intra-cell interference from multiple random-access attempts using preambles derived from the same Zadoff–Chu root sequence.

To handle different cell sizes, the cyclic shift N_{CS} is signaled as part of the system information. Thus, in smaller cells, a small cyclic shift can be configured, resulting in a larger number of cyclically shifted sequences being generated from each root sequence. For cell sizes below 1.5 km, all 64 preambles can be generated from a single root sequence. In larger cells, a larger cyclic shift needs to be configured and to generate the 64 preamble sequences, multiple root Zadoff–Chu sequences must be used in the cell. Although the larger number of root sequences is not a problem in itself, the zero cross-correlation property only holds between shifts of the same root sequence and from an interference perspective it is therefore beneficial to use as few root sequences as possible.

Reception of the random-access preamble is discussed later in this chapter. In principle, it is based on correlation of the received signal with the root Zadoff–Chu sequences. One disadvantage of Zadoff–Chu sequences is the difficulties in separating a frequency offset from the distance-dependent delay. A frequency offset results in an additional correlation peak in the time domain—a correlation peak that corresponds to a spurious device-to-base-station distance. In addition, the true correlation peak is attenuated. At low-frequency offsets, this effect is small and the detection performance is hardly affected. However, at high Doppler frequencies, the spurious correlation peak can be larger than the true peak. This results in erroneous detection; the correct preamble may not be detected or the delay estimate may be incorrect.

To avoid the ambiguities from spurious correlation peaks, the set of preamble sequences generated from each root sequence can be restricted. Such restrictions imply that only some of the sequences that can be generated from a root sequence are used to define random-access preambles. Whether such restrictions should be applied or not to the preamble generation is signaled as part of the system information. The location of the spurious correlation peak relative to the "true" peak depends on the root sequence and hence different restrictions have to be applied to different root sequences.

11.3.1.5 Preamble Detection

The base-station processing is implementation specific, but due to the cyclic prefix included in the preamble, low-complexity frequency-domain processing is possible. An example hereof is shown in Figure 11.13. Samples taken in a time-domain window are collected and converted into the frequency-domain representation using an FFT. The window length is 0.8 ms, which is equal to the length of the Zadoff–Chu sequence without a cyclic prefix. This allows handling timing uncertainties up to 0.1 ms and matches the guard time defined for the basic preamble configuration.

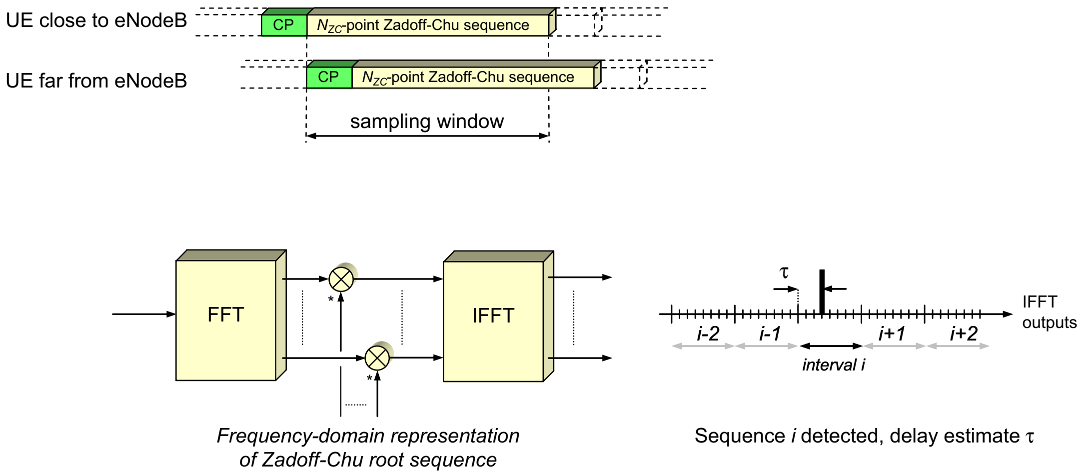

FIGURE 11.13

Random-access preamble detection in the frequency domain.

The output of the FFT, representing the received signal in the frequency domain, is multiplied by the complex-conjugate frequency-domain representation of the root Zadoff–Chu sequence and the result is fed through an IFFT. By observing the IFFT outputs, it is possible to detect which of the shifts of the root Zadoff–Chu sequence has been transmitted and its delay. Basically, a peak of the IFFT output in interval *i* corresponds to the *i*th cyclically shifted sequence and the delay is given by the position of the peak within the interval. This frequency-domain implementation is computationally efficient and allows simultaneous detection of multiple random-access attempts using different cyclically shifted sequences generated from the same root Zadoff–Chu sequence; in the case of multiple attempts there will simply be a peak in each of the corresponding intervals.

11.3.2 STEP 2: RANDOM-ACCESS RESPONSE

In response to the detected random-access attempt, the network will, as the second step of the random-access procedure, transmit a message on the DL-SCH, containing:

- The index of the random-access preamble sequences the network detected and for which the response is valid.
- The timing correction calculated by the random-access-preamble receiver.
- A scheduling grant, indicating what resources the device should use for the transmission of the message in the third step.
- A temporary identity, the TC-RNTI, used for further communication between the device and the network.

If the network detects multiple random-access attempts (from different devices), the individual response messages of multiple devices can be combined in a single transmission. Therefore, the response message is scheduled on the DL-SCH and indicated on a PDCCH using an identity reserved for random-access response, the RA-RNTI.[9] The usage of the RA-RNTI is also necessary as a device may not have a unique identity in the form of a C-RNTI allocated at this stage. All devices that have transmitted a preamble monitor the L1/L2 control channels for random-access response within a configurable time window. The timing of the response message is not fixed in the specification in order to be able to respond to many simultaneous accesses. It also provides some flexibility in the base-station implementation. If the device does not detect a random-access response within the time window, the attempt will be declared as failed and the procedure will repeat from the first step again, possibly with an increased preamble transmission power.

As long as the devices that performed random access in the same resource used different preambles, no collision will occur and from the downlink signaling it is clear to which device(s) the information is related. However, there is a certain probability of contention—that is, multiple devices using the same random-access preamble at the same time. In this case, multiple devices will react upon the same downlink response message and a collision occurs. Resolving these collisions is part of the subsequent steps, as discussed in the following. Contention is also one of the reasons why hybrid ARQ is not used for transmission of the random-access response. A device receiving a random-access response intended for another device will have incorrect uplink timing. If hybrid ARQ were used, the timing of the hybrid-ARQ acknowledgment for such a device would be incorrect and may disturb uplink control signaling from other users.

Upon reception of the random-access response in the second step, the device will adjust its uplink transmission timing and continue to the third step. If contention-free random access using a dedicated preamble is used, this is the last step of the random-access procedure as there is no need to handle contention in this case. Furthermore, the device already has a unique identity allocated in the form of a C-RNTI.

11.3.3 STEP 3: DEVICE IDENTIFICATION

After the second step, the uplink of the device is time synchronized. However, before user data can be transmitted to/from the device, a unique identity within the cell, the C-RNTI, must be assigned to the device (unless the device already has a C-RNTI assigned). Depending on the device state, there may also be a need for additional message exchange for setting up the connection.

In the third step, the device transmits the necessary messages to the eNodeB using the UL-SCH resources assigned in the random-access response in the second step. Transmitting the

[9]There are actually several RA-RNTIs defined. Which RA-RNTI a device is listening to is given by the time and frequency resource upon which the random-access preamble was transmitted.

uplink message in the same manner as scheduled uplink data instead of attaching it to the preamble in the first step is beneficial for several reasons. First, the amount of information transmitted in the absence of uplink synchronization should be minimized, as the need for a large guard time makes such transmissions relatively costly. Secondly, the use of the "normal" uplink transmission scheme for message transmission allows the grant size and modulation scheme to be adjusted to, for example, different radio conditions. Finally, it allows for hybrid ARQ with soft combining for the uplink message. The latter is an important aspect, especially in coverage-limited scenarios, as it allows for the use of one or several retransmissions to collect sufficient energy for the uplink signaling to ensure a sufficiently high probability of successful transmission. Note that RLC retransmissions are not used for the uplink RRC signaling in step 3.

An important part of the uplink message is the inclusion of a device identity, as this identity is used as part of the contention-resolution mechanism in the fourth step. If the device is in the RRC_CONNECTED state—that is, connected to a known cell and therefore has a C-RNTI assigned—this C-RNTI is used as the device identity in the uplink message.[10] Otherwise, a core-network device identifier is used and the eNodeB needs to involve the core network prior to responding to the uplink message in step 3.

Device-specific scrambling is used for transmission on UL-SCH, as described in Chapter 7. However, as the device may not yet have been allocated its final identity, the scrambling cannot be based on the C-RNTI. Instead, a temporary identity is used (TC-RNTI).

11.3.4 STEP 4: CONTENTION RESOLUTION

The last step in the random-access procedure consists of a downlink message for contention resolution. Note that, from the second step, multiple devices performing simultaneous random-access attempts using the same preamble sequence in the first step listen to the same response message in the second step and therefore have the same temporary identifier. Hence, the fourth step in the random-access procedure is a contention-resolution step to ensure that a device does not incorrectly use the identity of another device. The contention resolution mechanism differs somewhat depending on whether the device already has a valid identity in the form of a C-RNTI or not. Note that the network knows from the uplink message received in step 3 whether or not the device has a valid C-RNTI.

If the device already had a C-RNTI assigned, contention resolution is handled by addressing the device on the PDCCH using the C-RNTI. Upon detection of its C-RNTI on the PDCCH the device will declare the random-access attempt successful and there is no need for contention-resolution-related information on the DL-SCH. Since the C-RNTI is unique to one device, unintended devices will ignore this PDCCH transmission.

If the device does not have a valid C-RNTI, the contention resolution message is addressed using the TC-RNTI and the associated DL-SCH contains the contention-resolution message.

[10] The device identity is included as a MAC control element on the UL-SCH.

The device will compare the identity in the message with the identity transmitted in the third step. Only a device which observes a match between the identity received in the fourth step and the identity transmitted as part of the third step will declare the random-access procedure successful and promote the TC-RNTI from the second step to the C-RNTI. Since uplink synchronization has already been established, hybrid ARQ is applied to the downlink signaling in this step and devices with a match between the identity they transmitted in the third step and the message received in the fourth step will transmit a hybrid-ARQ acknowledgment in the uplink.

Devices that do not detect PDCCH transmission with their C-RNTI or do not find a match between the identity received in the fourth step and the respective identity transmitted as part of the third step are considered to have failed the random-access procedure and need to restart the procedure from the first step. No hybrid-ARQ feedback is transmitted from these devices. Furthermore, a device that has not received the downlink message in step 4 within a certain time from the transmission of the uplink message in step 3 will declare the random-access procedure as failed and need to restart from the first step.

11.4 PAGING

Paging is used for network-initiated connection setup when the device is in RRC_IDLE. In LTE, the same mechanism as for "normal" downlink data transmission on the DL-SCH is used and the mobile device monitors the L1/L2 control signaling for downlink scheduling assignments related to paging. Since the location of the device typically is not known on a cell level, the paging message is typically transmitted across multiple cells in the so-called tracking area (the tracking area is controlled by the MME; see [5] for a discussion on tracking areas).

An efficient paging procedure should allow the device to sleep with no receiver processing most of the time and to briefly wake up at predefined time intervals to monitor paging information from the network. Therefore, a paging cycle is defined, allowing the device to sleep most of the time and only briefly wake up to monitor the L1/L2 control signaling. If the device detects a group identity used for paging (the P-RNTI) when it wakes up, it will process the corresponding downlink paging message transmitted on the PCH. The paging message includes the identity of the device(s) being paged, and a device not finding its identity will discard the received information and sleep according to the DRX cycle. As the uplink timing is unknown during the DRX cycles, no hybrid-ARQ acknowledgments can be transmitted and consequently hybrid ARQ with soft combining is not used for paging messages.

The network configures in which subframes a device should wake up and listen for paging. Typically, the configuration is cell specific, although there is a possibility to complement the setting by device-specific configuration. In which frame a given device should wake up and search for the P-RNTI on a PDCCH is determined by an equation taking as input the identity of the device as well as a cell-specific and (optionally) a device-specific paging cycle.

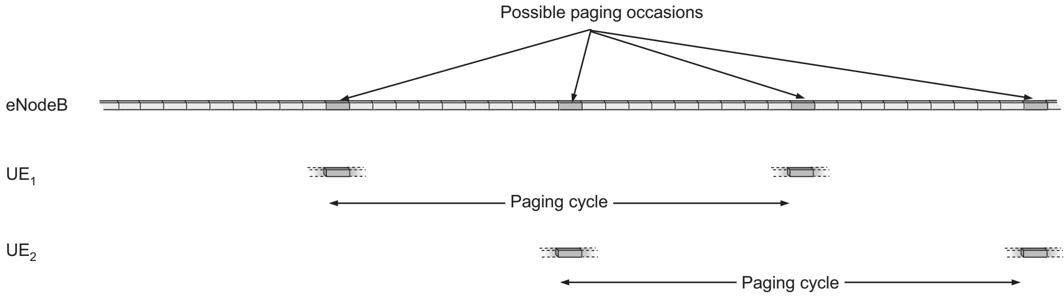

FIGURE 11.14

Illustration of paging cycles.

Table 11.1 Paging Cycles and Paging Subframes

		Number of Paging Subframes per Paging Cycle							
		1/32	**1/16**	**1/8**	**1/4**	**1/2**	**1**	**2**	**4**
Paging subframes in a paging frame	FDD	9	9	9	9	9	9	4, 9	0, 4, 5, 9
	TDD	0	0	0	0	0	0	0, 5	0, 1, 5, 6

The identity used is the so-called IMSI, an identity coupled to the subscription, as an idle mode device does not have a C-RNTI allocated, and the paging cycle for a device can range from once per 256 up to once per 32 frames. The subframe within a frame to monitor for paging is also derived from the IMSI. Since different devices have different IMSI, they will compute different paging instances. Hence, from a network perspective, paging may be transmitted more often than once per 32 frames, although not all devices can be paged at all paging occasions as they are distributed across the possible paging instances as shown in Figure 11.14.

Paging messages can only be transmitted in some subframes, ranging from one subframe per 32 frames up to a very high paging capacity with paging in four subframes in every frame. The configurations are shown in Table 11.1. Note that, from a network perspective, the cost of a short paging cycle is minimal as resources not used for paging can be used for normal data transmission and are not wasted. However, from a device perspective, a short paging cycle increases the power consumption as the device needs to wake up frequently to monitor the paging instants.

In addition to initiating connection to devices being in RRC_IDLE, paging can also be used to inform devices in RRC_IDLE as well as RRC_CONNECTED about changes of system information. A device being paged for this reason knows that the system information will change and therefore needs to acquire the update system information as described in Section 11.2.