iWOW Connections Pte Ltd

SMS Message in PDU Encoding

A. Introduction

Our product supports Short Message Service (SMS) in both Text and PDU mode. In Text mode, alphabets in a message will be supported based in the GSM Default Alphabets tables. AT command - AT+CMGF=1 enables application to support SMS for sending and receiving in Text mode.

PDU encoding is used in order to supports SMS that may contain alphabets within or beyond GSM Default Alphabets, Concatenated SMS or other special SMS services.

SMS may be coded in 8 bits (7 bits packed characters format), 8 bits or 16 bits. AT command - AT+CMGF=0 switches application to PDU mode.

The 7 bits packing format is used for characters in a message contains only GSM Default Alphabets. This allows maximum characters in a message to be 160 characters.

For 8 bits coding, the maximum characters in a message is 140.

16 bits coding is used when the message contains UCS2 alphabets. Since two bytes is needed to represent a character, this results maximum of 70 characters is allowed in a message.

Further information regarding Short Message Services is documented in GSM Specifications 03.38, 03.40, and 03.41. The focus of this document will be on the formatting of the most common cases for PDU encoding of SMS-DELIVER and SMS-SUBMIT.

Table1: GSM Default Alphabet

b4	b3	b2	b1		b7→0 b6→0 b5→0 (0)	0 0 1 (1)	0 1 0 (2)	0 1 1 (3)	1 0 0 (4)	1 0 1 (5)	1 1 0 (6)	1 1 1 (7)
0	0	0	0	0	@	Δ	SP	0	¡	P	¿	p
0	0	0	1	1	£	_	!	1	A	Q	a	q
0	0	1	0	2	$	Φ	"	2	B	R	b	r
0	0	1	1	3	¥	Γ	#	3	C	S	c	s
0	1	0	0	4	è	Λ	¤	4	D	T	d	t
0	1	0	1	5	é	Ω	%	5	E	U	e	u
0	1	1	0	6	ù	Π	&	6	F	V	f	v
0	1	1	1	7	ì	Ψ	'	7	G	W	g	w
1	0	0	0	8	ò	Σ	(	8	H	X	h	x
1	0	0	1	9	Ç	Θ	)	9	I	Y	i	y
1	0	1	0	10	LF	Ξ	*	:	J	Z	j	z
1	0	1	1	11	Ø	1)	+	;	K	Ä	k	ä
1	1	0	0	12	ø	Æ	,	<	L	Ö	l	ö
1	1	0	1	13	CR	æ	-	=	M	Ñ	m	ñ
1	1	1	0	14	Å	ß	.	>	N	Ü	n	ü
1	1	1	1	15	å	É	/	?	O	§	o	à

1) This code is an escape to an extension of the 7 bit default alphabet table. A receiving entity which does not understand the meaning of this escape mechanism shall display it as a space character.

Overview: <u>SMS-SUBMIT (Mobile Originated) PDU Encoding</u>

1-12 octets	1 octet	1 octet	2-12 octets	1 octet	1 octet	0,1 or 7 octet(s)	1 octet	0-140 octet(s)
SCA	PDU Type	MR	DA	PID	DCS	VP (Optional)	UDL	UD (Optional)

bit 7	bit8	bit9	bit4	bit3	bit2	bit1	bit0
RP	UDHI	SRR	VPF		RD	MTI	

	(Optional)	(Optional)				

Overview: <u>**SMS-DELIVER (Mobile Terminated) PDU Encoding**</u>

1-12 octets	1 octet	2-12 octets	1 octet	1 octet	0,1 or 7 octet(s)	1 octet	0-140 octet(s)
SCA	PDU Type	DO	PID	DCS	SCTS	UDL	UD (Optional)

bit 7	bit8	bit9	bit4	bit3	bit2	bit1	bit0
RP	UDHI (Optional)	SRI (Optional)			MMS		MTI

Abbreviation	Reference	Description
SCA	Service Center Address	Phone number for the Short Message Service Center
PDU Type	Protocol Data Unit Type	
RP	Reply Path	Parameter indicating the request for Reply Path
UDHI	User Data Header Indicator	Parameter indicating that the UD field contains a Header
SRR	Status Report Request	Parameter indicating if the MS is requesting a status report.
SRI	Status Report Indicator	Parameter indicating if the Short Message Entity (SME) has requested a status report.
VPF	Validity Period Format	Parameter indicating whether or not the VP field is present
RD	Reject Duplicates	Parameter indicating whether or not the SC shall accept an SMS-SUBMIT for an SM still held in the SC which has the same MR and the same DA as a previously submitted SM from the same OA
MMS	More Message to Send	Parameter indicating whether or not there are more messages to send
MTI	Message Type Indicator	Parameter describing the message type
MR	Message Reference	Parameter identifying the SMS-SUBMIT
DA	Destination Address	Address of the destination SME
PID	Protocol Identifier	Parameter identifying the above layer protocol, if any
DCS	Data Coding Scheme	Parameter identifying the coding scheme within the User Data
SCTS	Service Center Time Stamp	Parameter identifying the time when SMSC received the message
VP	Validity Period	Parameter identifying the time from where the message is no longer valid
UDL	User Data Length	Parameter indicating the length of the User Data field to follow
UD	User Data	Data for the SM

B. Parameter Description

1. Service Center Address (SCA)

It provides the SCA information. SMSC Address consists of 3 parts:

- Address Length
  It specifies the total number of bytes in the phone number plus Type of Address in hex value.
- Type of Address
  It provides the instruction on the interpretation of the phone number. 0x91 is used for international numbering and 0x81 is used for national numbering.
- Address Value
  This is a phone number entry in BCD format (each pair of digits is swapped) which excluding the "+" sign if international phone number prefix is used. "F" is filled at the end of the phone number if the Address Length is an odd number before BDC formatting is performed.

Note: For SCA that has been pre-configured onto the SIM card, the SCA field can be replaced with 0x00. It informs the application to use the default SCA stored in SIM card. "AT+CSCA" command is used to check or configure the SCA. Contact your network operator for the correct SCA.

Examples: SCA

Service Center Address	Type of Address	Address Value (BDC format)	Address Length	SCA
+65987654321F	0x91	5689674523F1	0x07	07**915689674523F1**
+6017645321	0x91	0671463512	0x06	06**910671463512**
pre-configured	-	-	-	0x00

2. PDU Type

Bit No.	PDU Type	Possible Entry	Explanation
7	RP	0	Reply Path parameter is not set in this SMS-SUBMIT
		1	Reply Path parameter is set in this SMS-SUBMIT
6	UDHI	0	User Data field contains only the short message
		1	The beginning of the UD field contains a Header in addition to the short message
5	SRR	0	A status report is not requested
		1	A status report is requested
5	SRI (Set by SMSC)	0	A status report will not be returned to the SME

		1	A status report will be returned to the SME
4 3	VPF	0 0	Validity Period not present
		1 0	Validity Period present- relative format
		0 1	Validity Period present - enhanced format (reserved)
		1 1	Validity Period present - absolute format
2	RD	0	Instruct the SC to accept SMS-SUBMIT for a SM still held in the SC which has the same MR and the same DA as a previously submitted SM from the same OA
		1	Instruct the SC to reject an SMS-SUBMIT for an SM still held in the SC which has the same MR and the same DA as the previously submitted SM from the same OA. In this case an appropriate FCS value will be returned in the SMS-SUBMIT-REPORT
2	MMS (Set by SMSC)	0	More Message are waiting for the MS in the SMSC
		1	No more Message are waiting for the MS in the SMSC
1 0	MTI	0 0	SMS-DELIVER (in the direction SC to MS) SMS-DELIVER REPORT (in the direction MS to SC)
		1 0	SMS-STATUS-REPORT (in the direction SC to MS) SMS-COMMAND (in the direction MS to SC
		0 1	SMS-SUBMIT (in the direction MS to SC) SMS-SUBMIT-REPORT (in the direction SC to MS)
		1 1	Reserved

Example: PDU Type

bit 7	bit8	bit9	bit4	bit3	bit2	bit1	bit0	PDU Type
RP	UDHI	SRR	VPF		RD	MTI		
0	0	0	0	0	0	0	1	0x01
0	0	0	1	0	0	0	1	0x11

3. **Message Reference (MR)**

The MR field gives an integer representation of a reference number of the SMS-SUBMIT. The MR is automatically incremented from 0 to 255 for each SMS-SUBMIT by the application. After each SMS-SUBMIT has been submitted to the network, the Last-Used-MR value in the SIM is updated with the MR that was used in the SMS-SUBMIT operation. The value in the MR assigned by the MS is the same value which is received at the SC.

Example MR: 0x00

4. **Destination Address (DO) or Origination Address (DA)**

Address field indicates the phone number of the recipient for SMS-SUBMIT or sender for SMS-DELIVER. Address field consists of 3 parts:
- Address Length expressed in number of digits
  It specifies the total number of digits in the phone number itself excludes Type of Address in hex value.
- Type of Address
  It provides the instruction on the interpretation of the phone number. 0x91 is used for international numbering and 0x81 is used for national numbering.
- Address Value
  This is a phone number entry in BCD format (each pair of digits is swapped) which excluding the "+" sign if international phone number prefix is used. "F" is filled at the end of the phone number if the Address Length is an odd number before BDC formatting is performed.

Example: Destination Address or Origination Address

Phone number	Address Length	Type of Address	Address Value (BDC format)	DO or DA
94751802	0x08	0x81	49578120	088149578120
+6597481802	0x0A	0x91	5649578120	0A915649578120
01222312345F	0x0B	0x81	1011322143F5	0B811011322143F5
+60173456789F	0x0B	0x91	0671436587F9	0B910671436587F9

5. **Protocol Identifier (PID)**

The PID field serves the purpose which the Transport Layer either refers to the higher layer protocol being used, or indicates interworking with a certain type of telematic device. Although there are a lot of PID definitions (refer to GSM 03.40 section), 0x00 is used. This informed SC that the PDU message must be treated as short message.

Example: PID: 0x00

6. **Data Coding Scheme (DCS)**

Coding Group bits 7 6 5 4	Use of bits 3 2 1 0		
0 0 X X	General Data Coding indication Bits 5..0 indicate the following :		
		Bit 5	Compress/ uncompressed text
		0	Indicates the text is uncompressed

	1 — Indicates the text is compressed using the GSM standard compression algorithm
	Bit 4 — Presence of message class **0** — Indicates that bits 1 to 0 are reserved and have no message class meaning **1** — Indicates that bits 1 to 0 have a message class meaning

Bit 1	Bit 0	Message Class
0	0	Class 0 (immediate display)
0	1	Class 1 (ME specific)
1	0	Class 2 (SIM specific)
1	1	Class 3 (TE specific)

Bits 3 and 2 indicate the alphabet being used, as follows :

Bit 3	Bit 2	Alphabet
0	**0**	**Default alphabet (7bit)**
0	**1**	**8 bit data**
1	**0**	**UCS2 (16bit)**
1	1	Reserved

NOTE: The special case of bits 7..0 being 0000 0000 indicates the Default Alphabet as in Phase 2

0 1 0 0..1 0 1 1	Reserved coding groups
1 1 1 1	Data coding/message class Bit 3 is reserved, set to 0

Bit 2	Message coding
0	Default alphabet
1	8-bit data

Bit 1	Bit 0	Message Class
0	0	Class 0 (immediate display)
0	1	Class 1 (ME specific)
1	0	Class 2 (SIM specific)
1	1	Class 3 (TE specific)

Default alphabet indicates the UD is coded in 7-bit alphabet format with refer GSM Default Alphabets table. When this coding is used, eight characters of the message are packed in 7 octets result a maximum of characters of 160 in a message.

8 bit data indicates the UD is coded in 8-bit format and result a maximum of characters of 140 in a message.

UCS2 data indicates the UD is coded in 16-bit format and a maximum of characters of 70 in a message.

For Class 0 (immediate display), in case ME incapable of displaying the mobile terminated message immediately; the ME will treat the message as though there was no message class.

As for the Class 1, the MS shall normally store the message in the ME by default, if that is possible, but otherwise the message may be stored elsewhere, e.g. in the SIM.

Class 2 and 3 mobile terminated message will be stored in SIM and TE respectively.

Examples: DCS

bit 7	bit 8	bit 9	bit 4	bit 3	bit 2	bit 1	bit 0	DCS	Note
0	0	0	0	0	0	0	0	0x00	7-bit alphabet
0	0	0	0	0	1	0	0	0x04	8-bit alphabet
0	0	0	0	1	0	0	0	0x08	16-bit alphabet

7. Service Center Time Stamp (SCTS)

The SCTS is the Information element offering the recipient of a short message the information of when the message arrived at the SM-TL entity of the SC. The time of arrival comprises the year, month, day, hour, minute, second and time zone. The SCTS is included on every SMS-DELIVERY being delivered to the SMSC. **The Time Zone indicates the difference, expressed in quarters of an hour, between the local time and GMT.**

Example: SCTS

	Year		Month		Day		Hour		Minute		Second		Time Zone	
Digits: (Semi-octets)	2		2		2		2		2		2		2	
	2	1	2	1	2	1	2	1	2	1	2	1	2	1
SCTS	5	0	0	3	1	3	6	1	7	0	9	3	2	3
Decoded	05		03		31		16		07		39		32	

Decode SCTS: 2005 March 31 16:07:39 GMT+08:00

8. Validity Period (VP)

The VP comprises 1 octet in integer representation in relative format gives the length of the VP, counted from when the SMS-SUBMIT is received by the SC.

In another case, the Validity Period comprises 7 octets in semi octet representation giving the absolute time of the VP termination. The representation of time is identical to the representation of the Service Center Time Stamp (SCTS).

The VP is present if the MR in the PDU Type is set to be present else this is omitted.

VP value	Validity period value
0 to 143	(VP + 1) x 5 minutes (i.e. 5 minutes intervals up to 12 hours)
144 to 167	12 hours + ((VP -143) x 30 minutes)
168 to 196	(VP - 166) x 1 day
197 to 255	(VP - 192) x 1 week

Example: 0xAA: 170 in decimal: (170-166) x 1 day= 4 days

9. **User Data Length (UDL)**

   If the User Data is coded using the 7-bit GSM default alphabet, the User Data Length field gives an integer representation of the number of septets within the User Data field to follow. If a User Data Header field is present, then the User Data Length value is the sum of the number of septets in the User Data Header field (including any padding) and the number of septets in the User Data field which follows.

   If the User Data is coded using 8-bit data, the User Data Length field gives an integer representation of the number of octets within the User Data field to follow. If a User Data Header field is present, then the User Data Length value is the sum of the number of octets in the User Data Header field and the number of octets in the User Data field which follows.

   If the User Data is coded using UCS2 data, the User Data Length field gives an integer representation of the number of octets within the User Data field to follow. If a User Data Header field is present, then the User Data Length value is the sum of the number of octets in the User Data Header field and the number of octets in the User Data field which follows.

Data Coding Scheme	UD	Note	UDL
00 (7-bit)	Howdy : C8F79D9C07	5 septets	0x05
04 (8-bit)	ABCDEFGHIJ :4142434445464748494A	10 bytes	0x0A
08 (16-bit)	你好！: 4F60597DFF01	6 bytes	0x06

10. **User Data (UD)**

    User Data is the message data coded depends on the Data Coding Scheme (DCS) being set.

    a. UD coded in 7-bit alphabet (GSM Default Alphabet)
       Special formatting needed in order to pack the 7-bit alphabet so that 140 octets are possible in a UD field. This formatting method is called "7-bit packed as 8-bit format".

       To code the UD in 7-bit, first determine the 7-bit binary alphabet representation with refer to GSM Default Alphabet. Secondly, follow the "7-bit packed as 8-bit format" packing procedure.

Characters	Hex Format	7-bit Binary format
H	0x48	1001000 (1a,1b,1c,1d,1e,1f,1g)
o	0x6F	1101111 (2a,2b,2c,2d,2e,2f,2g)
w	0x77	1110111(3a,3b,3c,3d,3e,3f,3g)
d	0x64	1100100 (4a,4b,4c,4d,4e,4f,4g)
y	0x79	1111001(5a,5b,5c,5d,5e,5f,5g)

   "7-bit packed as 8-bit format" packing procedure:

Octet No.	7	6	5	4	3	2	1	0	7-bit packed as 8-bit format
1	2g 1	**1a** 1	1b 0	1c 0	1d 1	1e 0	1f 0	1g 0	C8

2	3f 1	3g 1	**2a 1**	2b 1	2c 0	2d 1	2e 1	2f 1	F7
3	4e 1	4f 0	4g 0	**3a 1**	3b 1	3c 1	3d 0	3e 1	9D
4	5d 1	5e 0	5f 0	5g 1	**4a 1**	4b 1	4c 0	4d 0	9C
5	6c **0**	6d **0**	6e **0**	6f **0**	6g **0**	**5a 1**	5b 1	5c 1	07
6	7b	7c	7d	7e	7f	7g	**6a**	6b	-
7	**8a**	8b	8c	8d	8e	8f	8g	**7a**	**-**
8	10g	**9a**	9b	9c	9d	9e	9f	9g	-
. .	. .	. .	. .	. .	. .	. .	. .	. .	-

Result: UD for "Howdy" is C8F79D9C07

b. UD coded in 8-bit alphabet
This formatting is related to Intel ASCII-HEX format where one byte is used to represent two characters. Since 140 octets are allocated for UD field, 140 characters are the maximum characters when coded in 8-bit format.

ASCII characters	Hex Format
H	0x48
o	0x6F
w	0x77
d	0x64
y	0x79

Result: UD for "Howdy" is 486F776479

c. UD coded in UCS2 (16-bit)
For maximum of 140 octets in UD field results only 70 characters allowable.

C characters	UCS2 Format
你	4F60
好	597D
！	FF01

Result: UD for "你好！" is 4F60597DFF01

C. Examples for complete PDU Short Message

1. SMS-SUBMIT (Mobile Originated)

SCA	PDU Type	MR	DA	PID	DCS	VP	UDL	UD
00 - Default SCA is used	11 - SMS-SUBMIT MS to SMSC - VP field is present	00 - Auto – increment is preformed for the Last-Used-MR	0A915649578120 - (DA)10 digits in the phone number - (91) International phone number -5649578120 is BCD of 6594751802	00 - PDU is treated as short message	00 - GSM Default Alphabet (7-bit)	AA - 4 days of validity period	05 - 5 septets in the UD field	C8F79D9C07 - Message content is "Howdy"
00	01 - SMS-SUBMIT to SMSC - VP field is not present	00	0A915649578120	00	00	-	0C - 12 septets in the UD	D4F29C0E6A96E7F3F0B90C - Message content is "Test Message"
06915669489599 - SCA +6596845999	11	00	0A915649578120	00	00	AA	05	C8F79D9C07 - Message content is "Howdy"
00	11	00	0A915649578120	00	04 8-bit	AA	0A - 10 octets in the UD	4142434445464748494A - Message content is "ABCDEFGHIJ"
00	11	00	0A915649578120	00	08 USCS 16-bit	AA	06 - 6 octet in the UD	4F60597DFF01 - Message content is "你好！"
00	11	00	0A915649578120	00	08	AA	10 - 16 octet in the UD	4F60597DFF010048006F007700640079 - Message content is "你好！Howdy"

2. AT Command Interface for Sending Message
Command: Send Message (AT+CMGS)

Description: This command is use to send message. To send a message, type Ctrl-Z (ASCII Char (26)). To abort the message, type ESC (Escape/ ASCII Char (27)).
The total bytes of PDU message is excluded the SCA address field expressed in decimal value.

Format: AT+CMGS=<total bytes of PDU Message><CR /ASCII Char (13)>
<PDU Message><Ctrl-Z>

Command	Possible Response(s)
AT+CMGF=0	OK

Note: Enable PDU mode	
AT+CMGS=18 >0011000A9156495781200000AA05C8F79D9C07 Note: Send a PDU message	+CMGS: 10 OK Note: Successfully sent
AT+CMGS=18 **06915669489599**11000A915649578120000 0AA05C8F79D9C07 Note: SCA address is supplied	+CMGS: 11 OK Note: Successfully sent
	+CMS ERROR: 308 Note: invalid PDU parameter; error could be due to incorrect PDU parameter set

3. SMS-DELIVER (Mobile Originated)

SCA	PDU Type	DO	PID	DCS	SCTS	UDL	UD
06915669489569	04 - SMS-DELIVER SMSC to MS - No more message in the SMSC	0A915649578120 - (OA)10 digits in the phone number - (91) International phone number - 5649578120 is BCD of 6594751802	00 - short message must be treated as short message	00 - GSM Default Alphabet (7-bit)	50301391612023 - Time Stamp:2005 March 31 19:16:02 GMT+08:00	05 - 5 septets in UD field	C8F79D9C07 - Message content "Howdy"
06915669489569	04	0A915649578120	00	08 -UCS2 Alphabet (16-bit)	50301391612023	06 - 6 octets in UD field	4F60597DFF01 - Message content is "你好！"

4. AT Command Interface for Received Message
Command: Read Message (AT+CMGR)

Description: Short Message is sent to the ME if there is enough memory storage to hold the message. Use AT+CNMI command to turn on the new message indication.

Format: AT+CMGR=<index>

Command	Possible Response(s)
AT+CMGF=1	OK
AT+CNMI=2,1,0,0,0 Note: enable new message indication	OK
	+CMTI: "SM",3 Note: new message received, stored at SIM card with Index 3
AT+CPMS="SM" Note: Select SIM card storage for the preferred message storage for reading, writing etc.	OK
AT+CMGR=3	+CMGR: 0,,23

Note: Read message	06915669489569040A91564957812000 **00**5040104111852**305C8F79D9C07** OK Note: Message "Howdy" is coded in 7-bit is read

References

1. Digital cellular telecommunications system (Phase 2+); Alphabets and language-specific information (GSM 03.38)
2. Digital cellular telecommunications system (Phase 2+); Technical realization of Short Message Service Cell Broadcast (SMSCB) (GSM 03.41)
3. Digital cellular telecommunications system (Phase 2+); Technical realization of the Short Message Service (SMS); Point-to-Point (PP) (GSM 03.40)
4. Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Technical realization of Short Message Service (SMS) (3GPP TS 23.040 version 3.10.0
5. http://www.unicode.org/