Title:	**On integrity protection and the effects of additional segmentation**
Source:	Vodafone
Date:	19th April 2001
Document for:	Discussion

1 Introduction

Over the last TSG GERAN and ad hoc meetings, discussion has taken place regarding the applicability of integrity protection in GERAN. During these discussions, some possible issues have been pointed out: performance of radio procedures, spectral efficiency and impact on speech quality. This paper intends to reflect what Vodafone's position on these potential problems is.

Other issues, like the enhancement of the segmentation mechanisms in RLC/MAC in the control plane, are not addressed here.

2 Discussion

2.1 General

In order to integrity protect a message, a Message Authentication Code (MAC-I) needs to be included in the message so that the receiving end can confirm its origin [1]. In GERAN RRC, this would result in appending the MAC-I (32 or 36 bits) as a new Information Element (IE) to all those layer 3 messages that need to be integrity protected. In GERAN, RRC messages will be transported over layer 2 protocol (LAPDm or RLC), which has a maximum size. Therefore, in some cases the addition of the new IE will mean that the message can no longer be transmitted with the same number of layer 2 blocks and cause (additional) segmentation.

The above also applies to RLC/MAC control messages that are integrity protected.

Three unfavourable effects of the increase in the amount of segmentation have been flagged up as cause for concern when it comes to the applicability of integrity protection in GERAN. They are:

a) Possible decrease in the performance of some procedures over the radio interface

b) Possible decrease in the spectrum efficiency

c) Possible decrease in the speech quality

These three points are addressed in the sections below.

2.2 Performance of radio procedures

2.2.1 General

The fact that some messages used in certain procedures are (additionally) segmented, has been pointed out as being a possible drawback, since the time to send those messages correctly is increased and the procedure takes more time to be performed. Although this is not felt to be a problem in the majority of the procedures, especial consideration needs to be paid to time critical procedures such as the handover.

The following section analyses the performance impact of (additional) segmentation in a loss channel (e.g. like in the radio interface). In particular it attempts to clarify how the rate of the number of retransmissions increases.

2.2.2 Extra delay caused by additional segmentation

2.2.2.1 General

The time T_n that is needed to correctly transmit a message made up by n blocks can be approximated to

$$T_n \approx t_f \cdot M_n \quad \text{(Eq. 1)}$$

where t_f is the duration of a frame and M_n is the average number of frames needed to transmit in order to receive correctly a n-frame message. Decoupling first attempt from retransmissions, (Eq. 1) can be rewritten as

$$T_n \approx t_f \cdot (n + M_n^r) = t_f n + t_f M_n^r \quad \text{(Eq. 2)}$$

Note that this model thus considers a continuous flow of frames, either first attempt or retransmissions, and ignores the idle periods that may occur while waiting for acknowledgements from the receiving entity.

2.2.2.2 Delay for additional block(s)

As seen above, the increase of the number of frames increases the time needed to transmit them for the first time. This increase is deterministic (i.e. product of t_f and n) and is present even in a perfect channel.

2.2.2.3 Delay for retransmissions

The second addend of (Eq. 2) refers to the time needed for retransmissions due to the errors in the radio channel. It can be shown (see Eq. 11 in the annex) that the number of retransmissions is very small when the probability of receiving a frame correctly is high. For instance, for p =0.99 (99% reliability, 1% frame error rate), M_2^r = 2.02×10^{-3} and M_3^r = 3.03×10^{-3}. So, the additional segmentation means that one extra frame has to be re-transmitted approximately once out of a hundred times that the message is sent.

Furthermore, (Eq. 2) shows that the increase of time due to retransmission also grows linearly with the number of segments of the message.

2.2.3 Handover procedure

From this analysis, no procedure should suffer a significant decrease of performance due to the additional segmentation. Vodafone believe that this is valid also for *time critical* procedures like the handover. In addition, lab tests of the performance of the handover procedure subject to execution delays have been performed and the results show the lack of dependency of any performance measure with small delays in the execution of the handover procedure.

2.3 Spectrum efficiency

The degradation to spectrum efficiency (if any) will be a function of the frequency of which an integrity protected message is used. The current exchange of messages during the normal call or transaction is not frequent and as such any additions to the size of the message due to integrity protection, in our opinion, does not impact spectrum efficiency. If there are procedures where the exchange of messages is more frequent (as it may be the case of RLC/MAC control messages), then this procedure is by default spectrum non-efficient and close examination to the particular cases will be needed.

2.4 Speech quality

When the segmented messages are sent on the FACCH, no additional spectrum is used, but speech frames are blanked. It is thus important to consider how the (additional) segmentation can affect the speech quality. Since the additional number of blocks is not significant, nor the amount of retransmissions and the messages sent in the FACCH are not *very* frequent, Vodafone tend to believe that the impact on speech quality in this case is not significant.

During the discussions on integrity protection and segmentation, the issue of speech quality has been brought up due to the increase of the window size of the layer 2 protocol. However, this issue is not directly correlated to the segmentation and it is addressed in [3].

3 Conclusions

The addition of information elements (e.g. a MAC-I for integrity protection) may cause the segmentation of some messages. Possible effects of this segmentation have been studied in this paper: performance of radio procedures,

spectral efficiency and impact on speech quality. It is expected that these three be affected unfavourably by the increase of the segmentation, although not in a significant way in the case of RRC messages.

Vodafone tend to believe that the benefits of integrity protection outweigh the side effects of possible (additional) segmentation and it is requested that the use of integrity protection in GERAN be adopted as the working assumption, unless other significant impacts are found. Further study is felt to be needed before this is ratified for the case of RLC/MAC control messages.

References

[1] 3GPP TS 33.102; "3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; 3G Security; Security Architecture"

[2] 3GPP TS 45.005; "3rd Generation Partnership Project; Technical Specification Group GERAN; Digital cellular telecommunications system (Phase 2+); Radio transmission and reception"

[3] TD GP#4(01)0769, "Modified 'leaky bucket' to limit FACCH usage", Vodafone, 3GPP TSG GERAN meeting #4, Biarritz (France), 2nd – 6th April 2001.

[4] "Probability and Statistical Inference", R V Hogg, E A Tanis; Ed. Prentice Hall; 5th edition, 1997; pp. 135-137.

Annex: calculation of the average number of retransmissions in a non-perfect channel

The analysis here performed assumes a channel where the probabilities of receiving correctly a layer 2 frame are not correlated, i.e. whether one frame is received correctly or not is considered to be independent of the correct reception of the previous and following frames. This assumption ignores the effect of slow fading, where the correlation of the erroneous reception of two consecutive frames is not negligible.

With the simplification above, the correct reception of one layer 2 frame can be modelled as a Bernoulli experiment characterised by *p*, the probability of receiving a frame correctly, where $p \in [0,1]$.

From this, it can be derived [4] that the number of layer 2 frames *n* correctly received out of *m* trials follows a binomial distribution:

$$f(n) = \binom{m}{n} p^n (1-p)^{m-n}, \quad n = 0, 1, 2, ..., m. \tag{Eq. 3}$$

where

$$\binom{m}{n} = \frac{m!}{n!(m-n)!} \tag{Eq. 4}$$

is the number of combinations of *n* frames received successfully out of *m* attempts. The probability of needing *m* to send frames to correctly receive a *n*-frame message can be expressed as the probability of receiving correctly $(n-1)$ frames out of $(m-1)$ attempts plus the correct reception of the last one:

$$P(n,m) = \left[\binom{m-1}{n-1} p^{n-1} (1-p)^{(m-1)-(n-1)}\right] \cdot p = \binom{m-1}{n-1} p^n (1-p)^{m-n} \tag{Eq. 5}$$

Then, the average number of frames can be easily calculated as

$$M_n = \sum_{m=n}^{\infty} m \cdot P(n,m) = \sum_{m=n}^{\infty} m \cdot \binom{m-1}{n-1} p^n (1-p)^{m-n} \tag{Eq. 6}$$

Due to the assumed independence of all transmissions, it can be derived that

$$M_n = n \cdot M_1 \tag{Eq. 7}$$

Particularising (Eq. 6) for *n* = 1

$$\begin{aligned} M_1 &= \sum_{m=1}^{\infty} m \cdot \binom{m-1}{0} p(1-p)^{m-1} = \sum_{m=1}^{\infty} m \cdot p(1-p)^{m-1} = -p \sum_{m=1}^{\infty} m \cdot (1-p)^{m-1} = -p \sum_{m=1}^{\infty} \frac{d}{dp}(1-p)^m = \\ &= -p \frac{d}{dp} \sum_{m=1}^{\infty} (1-p)^m = -p \frac{d}{dp}\left(\frac{1-p}{1-(1-p)}\right) = -p \frac{d}{dp}\left(\frac{1}{p} - 1\right) = -p \frac{d}{dp}\left(\frac{1}{p}\right) = \frac{1}{p} \end{aligned} \tag{Eq. 8}$$

Thus, substituting in (Eq. 7)

$$M_n = n \cdot M_1 = n \frac{1}{p} \tag{Eq. 9}$$

or, separating first frames and their repetitions as in (Eq. 2),

$$M_n = n \frac{1}{p} = n + n \cdot \left(\frac{1}{p} - 1\right) = n + M_n^r \tag{Eq. 10}$$

Therefore, in addition to the extra frame that needs to be transmitted always when a message is segmented further, the number of additional frame repetitions also increases. However, note that the dependence of M_n^r with n is linear and that its value is very small for realistic values of p:

$$\lim_{p \to 1} M_n^r = \lim_{p \to 1} n \cdot \left(\frac{1}{p} - 1 \right) = 0 \qquad \text{(Eq. 11)}$$