

# MỘT PHƯƠNG ÁN TRUYỀN DỮ LIỆU QUA KÊNH THOẠI GSM

Nguyễn Thanh Bình\*, Đặng Văn Trường†, Trần Văn Liên‡

\*Vụ Khoa học – Công nghệ, Ban Cơ yếu Chính phủ

†Viện Nghiên cứu Khoa học – Công nghệ Mật mã, Ban Cơ yếu Chính phủ

‡Tổng công ty Điện tử và Tin học Việt Nam

**Tóm tắt:** Hiện nay điện thoại di động và công nghệ GSM rất phổ biến với người dùng trên toàn thế giới. Nhu cầu bảo mật tín hiệu thoại GSM là rất lớn, trong khi giải pháp bảo mật sẵn có của GSM vẫn chưa đạt độ an toàn cần thiết. Giải pháp bảo mật tốt nhất cho tín hiệu thoại GSM là chuyển tín hiệu sang dạng số qua Vocoder riêng, mã hóa tín hiệu số đó bằng thuật toán mật mã mạnh và truyền tín hiệu đã mã thông qua kênh thoại GSM. Trong bài viết này chúng tôi sẽ trình bày một phương án truyền dữ liệu qua kênh thoại GSM tương thích với các tính chất của kênh thoại GSM, cụ thể là bộ Vocoder GSM và bộ phát hiện tiếng nói tích cực VAD cho GSM.

**Từ khóa:** Công nghệ GSM, GSM Vocoder, kênh thoại GSM, tín hiệu thoại GSM, OFDM, VAD.

## I. YÊU CẦU CỦA VIỆC TRUYỀN DỮ LIỆU QUΑ KÊNH GSM

Kênh thoại GSM có độ ưu tiên cao, độ trễ thấp và chất lượng tốt hơn các dịch vụ thoại cổ điển khác. Mặc dù các công ty viễn thông cam kết đảm bảo tính an toàn thông tin của hệ thống, nhưng thực tế mạng GSM có nhiều lỗ hổng nghiêm trọng. Đối với một số tổ chức và lĩnh vực như tài chính, ngân hàng, an ninh, quốc phòng... thì nội dung các cuộc gọi qua hệ thống GSM càng phải được giữ bí mật. Tuy nhiên, ở đây yêu cầu về mật kích thước, khối lượng phải nằm trong giới hạn nhất định và quá trình mã hóa phải được thực hiện trong thời gian thực.

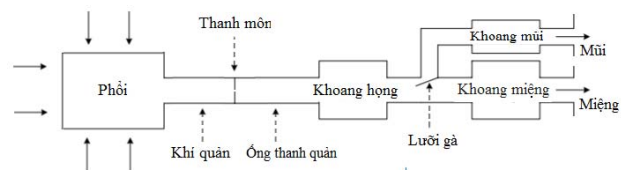
Một số giải pháp được đưa ra là mã hóa tín hiệu thoại trước khi truyền qua hệ thống GSM và giải mã tín hiệu sau đó ở đầu xa. Phương pháp này ngăn chặn được các bên thứ ba, bao gồm cả nhà điều hành mạng tiếp cận nội dung của cuộc thoại. Để làm điều này cần có các modem đặc biệt để truyền dữ liệu qua kênh thoại GSM. Song trong quá trình truyền dữ liệu thoại GSM đã có bộ nén tiếng nói (Vocoder) từ thiết bị đầu cuối. Truyền dữ liệu số qua kênh nén tiếng nói là nhiệm vụ đầy thách thức. Việc nén tiếng nói được thực hiện bằng cách trích xuất các tham số từ lời nói tại bộ mã hóa và sau đó tổng hợp nó tại bộ giải mã. Do đó, tín hiệu đầu vào của kênh có thể hoàn toàn khác với tín hiệu đầu ra. Để giải quyết vấn đề này, một số phương pháp sẽ được thảo luận trong bài viết. Về cơ bản là thực hiện ánh xạ dữ liệu thành tín hiệu giống như lời nói và sau đó truyền dữ liệu qua kênh.

Vấn đề của việc truyền dữ liệu kỹ thuật số thông qua GSM là các kênh này không được thiết kế cho mục đích truyền dữ liệu, mà là để truyền lời nói. Đối với các mô hình truyền dữ liệu qua các tham chiếu của phương pháp phân tích và tổng hợp giọng nói, rất khó để xây dựng mô hình toán học và có thể nói là đến giờ chưa có mô hình nào khả dĩ có thể sử dụng được để tính lỗi bit BER. Lỗi bit BER chỉ được xác định qua thực nghiệm.

## II. MÃ HÓA TIẾNG NÓI TRONG THÔNG TIN DI ĐỘNG GSM

### 2.1 Các tính chất cơ bản của tiếng nói con người

Không khí được ép từ phổi đi qua thanh quản bao gồm các dây thanh âm dao động (theo sự điều khiển của hệ thần kinh) rồi đi dọc theo cơ quan phát âm sẽ tạo ra tiếng nói. Sự dao động của các dây thanh âm tạo ra sự đóng mở tương tự như một cánh cửa (thanh môn). Ngoài sự tác động của thanh quản tạo ra các dao động có tần số cơ bản, các thành phần hài bậc cao của tiếng nói phụ thuộc vào sự thay đổi của cơ quan phát âm gồm: họng, vòm họng, lưỡi, miệng, khoang mũi và mũi tương tự như sự thay đổi tham số của các hốc cộng hưởng. Hình sau là biểu diễn mô hình cơ học của hệ thống phát âm của người [11].



Hình 1. Biểu diễn mô hình cơ học của hệ thống phát âm

Từ mô hình cơ học trên, có thể biểu diễn cơ quan phát âm bằng mô hình gần đúng gồm các ống hình trụ có độ dài bằng nhau nhưng có đường kính khác nhau. Các ống hình trụ này là các hốc cộng hưởng âm thanh với các tần số riêng gọi là tần số formant. Các tần số này tạo ra các âm vị khác nhau tùy theo hình dáng cơ quan phát âm. Mô hình này có thể được biểu diễn một cách khá chính xác bằng hệ phương trình vi phân. Trong quá trình phát âm người ta thấy rằng hình dáng cơ quan phát âm thay đổi rất chậm. Vì vậy người ta có thể biểu diễn hệ thống phát âm bằng một hệ tuyến tính bất biến theo thời gian.

Ngoài ra mô hình hoá quá trình kích thích của luồng không khí từ phổi đi qua thanh quản lên cơ quan phát âm

Tác giả liên hệ: Nguyễn Thanh Bình

Email: [binhnt@bcy.gov.vn](mailto:binhnt@bcy.gov.vn)

Đến tòa soạn: 10/2019, chỉnh sửa 12/2019, chấp nhận đăng 12/2019

cũng rất quan trọng. Tùy theo loại âm thanh mà có cách mô hình hoá thích hợp để tiếng nói sau khi tái tạo đạt được chất lượng theo yêu cầu. Trong kỹ thuật mã hoá tiếng nói, dựa vào dao động của các dây thanh âm có thể chia tiếng nói ra thành hai loại âm chính sau đây:

- *Âm hữu thanh (voiced sound)*: âm hữu thanh được tạo ra khi các dây thanh âm dao động đóng mở làm ngắt quãng luồng không khí và sự ngắt quãng này được xem gần như là tuần hoàn tác động lên cơ quan phát âm. Trong thực tế chu kì tuần hoàn này khoảng từ 2 -20ms. Do đó với âm hữu thanh, tín hiệu kích thích được mô hình hoá là các xung tuần hoàn.

- *Âm vô thanh (unvoiced sound)*: âm vô thanh được tạo ra khi luồng không khí đi qua thanh môn tác động lên cơ quan phát âm không theo một qui luật nào cả (không tuần hoàn). Do đó với âm vô thanh, tín hiệu kích thích được mô hình hoá tương tự như tín hiệu ngẫu nhiên (nhiều).

2.2 Cấu trúc của một bộ mã hoá tiếng nói dùng phương pháp mã hoá lai AbS

Hầu hết các tiêu chuẩn mã hoá tiếng nói trong thông tin di động GSM đều sử dụng phương pháp mã hoá lai AbS.

Trong các bộ mã hoá lai, các thông số của hệ thống sẽ được xác định bằng kỹ thuật dự đoán tuyến tính như trong mã hoá tham số (ở trong phương pháp mã hoá nguồn) và tín hiệu kích thích được xác định bằng một vòng kín (phân tích bằng cách tổng hợp).

Trong hình 3 [5] là sơ đồ khối của một bộ mã hoá lai điển hình. Hệ thống này bao gồm một bộ lọc dự đoán thời gian ngắn (STP)  $A(z)$ , một bộ lọc dự đoán thời gian dài (LTP)  $A_L(z)$ , một bộ lọc nhân  $W(z)$ , một bộ giảm thiểu sai số cung cấp thông tin cần thiết cho bộ tạo tín hiệu kích thích. Trong đó, quan trọng nhất là bộ tạo tín hiệu kích thích vì nó tạo ra hay chọn tín hiệu kích thích sao cho sai số bình phương trung bình sau khi qua  $W(z)$  là nhỏ nhất. Tùy theo mỗi loại mã hoá mà bộ tạo tín hiệu kích thích này khác nhau. Mặc dù sơ đồ trên là chung cho các bộ mã hoá lai nhưng một số loại không sử dụng bộ lọc LTP hoặc

vị trí STP và LTP thay đổi.

2.3 Dự đoán tuyến tính (LP) dựa trên mô hình phát âm

Dự đoán tuyến tính rất quan trọng trong xử lý số tiếng nói. Nó là một công cụ kỹ thuật rất hiệu quả để ước lượng các thông số của tiếng nói như pitch, tần số formant, phổ ... khá chính xác với tốc độ tính toán nhanh. Dựa trên hàm truyền đạt người ta có thể biểu diễn mô hình cơ quan phát âm một cách gần đúng như hình sau [12]:



Hình 2. Sơ đồ ngắn gọn của quá trình tạo tiếng nói

Hàm  $A_L(z)$ ,  $A(z)$  là đa thức thu được trực tiếp từ phép biến đổi  $z$  của phương trình sai phân tuyến tính hoặc từ phép biến đổi Laplace của phương trình vi phân tuyến tính liên tục chuyển qua gián đoạn với khoảng thời gian  $T$  và thay biến  $s$  bằng biến  $z$  (biến đổi song tuyến tính - bilinear transform) [1]:

$$z = \frac{1+sT}{1-sT} \tag{1}$$

Sau khi lấy loga cả hai vế, xấp xỉ bằng chuỗi sẽ thu được cặp biểu thức ở dưới (biến đổi song tuyến tính):

$$s \approx \frac{2z-1}{Tz} \text{ và } z \approx \frac{1+sT}{1-sT} \tag{2}$$

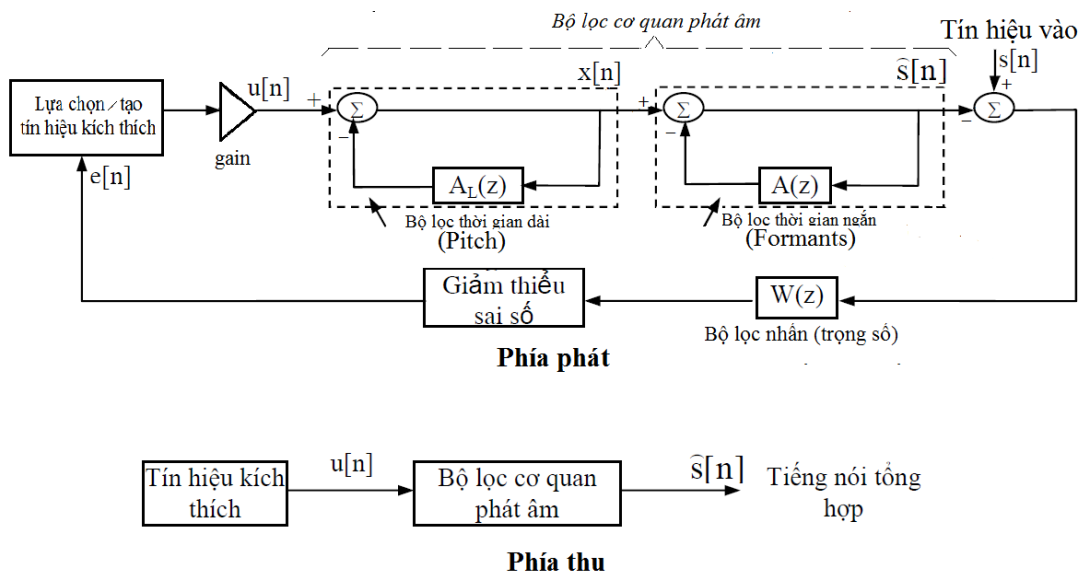
Đa thức  $A_L(z)$ ,  $A(z)$  có dạng sau :

$$\frac{1}{A(z)} = \frac{1}{1 - \sum_{i=1}^p \alpha_i z^{-i}} \tag{3}$$

trong đó  $p$  là bậc của đa thức,  $\alpha$  là hệ số lọc,  $i$  là chỉ số

2.4 Dự đoán thời gian ngắn (STP) và dự đoán thời gian dài (LTP)

Bộ dự đoán thời gian ngắn thực chất là bộ lọc tổng hợp tiếng nói. Bộ lọc này sẽ thực hiện việc tổng hợp tiếng nói khi có tín hiệu kích thích đưa đến đầu vào của nó. Các hệ



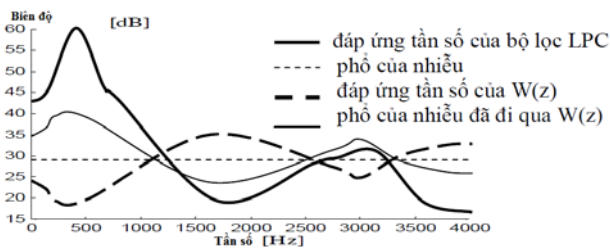
Hình 3. Sơ đồ khối của một bộ mã hoá lai

số của bộ lọc này sẽ được xác định bằng phương pháp dự đoán tuyến tính như đã đề cập ở trên. Các đoạn tiếng nói hữu thanh có dạng sóng tuần hoàn và tính chất tuần hoàn này có thể được khai thác để trợ giúp cho quá trình dự đoán tiếng nói. Từ điều những điều trên người ta đã đưa ra khái niệm về dự đoán thời gian dài hay dự đoán pitch. Cũng giống như các bộ STP, các bộ LTP cũng là các bộ dự đoán tuyến tính nhưng trong khi STP thực hiện việc dự đoán dựa trên các mẫu kế nhau thì LTP dựa trên các mẫu từ một hay nhiều chu kì pitch trước đó. Đây là lý do gọi nó là dự đoán thời gian dài.

Trong thực tế thay vì truyền đi các hệ số của  $A_L(z)$  và  $A(z)$  người ta truyền đi các thành phần là LSF hoặc LSP cùng với biên độ (hay năng lượng) của tiếng nói. Bên thu tái tạo lại đa thức  $A_L(z)$ ,  $A(z)$  cùng với các thông số khác và tổng hợp lại tiếng nói.

### 2.5 Bộ lọc nhân (lọc trọng số $W(z)$ )

Ngoài việc khai thác các tính chất tiếng nói để mã hoá, người ta còn khai thác sự cảm nhận âm thanh của tai người (tai người không cảm nhận được những âm thanh bị che đi bởi các âm thanh khác có năng lượng lớn hơn một mức nhất định – hiệu ứng che lấp) trong mã hoá tiếng nói bằng khái niệm bộ lọc nhân (cảm nhận).



Hình 4. Biểu diễn hiệu của  $W(z)$

Sự tác động của bộ lọc này được biểu diễn trong hình trên. Ta thấy phổ của nhiễu có hai vùng nằm phía trên của đáp ứng tần số của bộ lọc LPC do đó các tần số nằm trong vùng này sẽ bị nhiễu che đi. Bộ lọc nhân  $W(z)$  sẽ nâng biên độ của nhiễu trong vùng tần số formant (vùng đỉnh của đáp ứng tần số bộ lọc LPC) và nén biên độ của nhiễu trong các vùng trũng của đáp ứng tần số. Phổ của nhiễu sau khi đi qua  $W(z)$  sẽ có dạng là đường liền nét mảnh (có hình dạng phổ tương tự như phổ của bộ lọc LPC) và nhiễu sẽ dễ dàng bị các tần số formant che đi (năng lượng các tần số formant che năng lượng nhiễu). Tóm lại,  $W(z)$  sẽ định dạng nhiễu hay các sai số sao cho chúng bị che đi bởi các tần số formant năng lượng cao.

### 2.6 Các đặc điểm cơ bản của mạng GSM.

Phần này trình bày các tính chất và đặc điểm cơ bản của tín hiệu tiếng nói truyền qua mạng GSM. Ở đây chỉ xem xét các ảnh hưởng xấu của nó đến quá trình truyền dữ liệu qua kênh thoại GSM trên khía cạnh băng tần, kỹ thuật xử lý mã thoại, chuyển đổi mã thoại mà sẽ không xem xét những tác động khác thuộc về truyền thông trong nội tại mạng GSM hay GSM – PSTN.

- Mất mát tín hiệu do băng tần hẹp:

Kênh thoại mạng GSM được thiết kế để truyền tín hiệu tiếng nói của con người với băng tần hẹp 300-3400Hz. Tín hiệu nghe hiểu được qua bộ lọc thông dải tại đầu vào của máy điện thoại và có thể các giai đoạn trung gian, trong đó những thành phần tần số nằm ngoài băng tần bị loại bỏ. Bộ lọc này làm mất mát một phần dòng tín hiệu nghe hiểu được truyền. Hơn nữa, kênh với băng tần 300-3400Hz là quá hẹp, dẫn đến bị hạn chế tốc độ.

- Mất mát tín hiệu do bộ mã tiếng nói:

Sự mất mát tín hiệu do những bộ mã tiếng nói là thách thức lớn nhất đối với điều chế dữ liệu. Bộ mã tiếng nói số hóa tín hiệu tiếng nói ở đầu phát và tái tạo lại tại đầu thu. Những bộ mã sử dụng trong GSM khai thác triệt để những thuộc tính vốn có trong tín hiệu tiếng nói để thu được hiệu suất nén cao, trong khi vẫn giữ lại chất lượng tiếng nói nghe hiểu theo kinh nghiệm của người nghe. Điều đó dẫn đến tín hiệu không giống tiếng nói bị phá hỏng nặng nề, bởi chúng vi phạm thuộc tính của tiếng nói đã được các bộ mã mặc định. Những bộ mã này xử lý trên cơ sở kỹ thuật nghe hiểu âm thanh cảm quan, mô hình hóa tiếng nói đầu vào và chỉ những đặc tính liên quan đến tiếng nói con người được truyền đi.

Bộ mã mặc định rằng, dạng sóng tiếng nói chỉ có những thay đổi chậm theo thời gian và thực tế, tiếng nói có thể được mô hình hóa giống như những sóng có chu kỳ với một tần số cơ sở có xen kẽ những khoảng lặng và âm bật. Những bộ mã tiếng nói trong mạng GSM xây dựng trên cơ sở mô hình toàn cực (All-Pole Model) và chúng được thiết lập trong mã LPC. LPC ước lượng những thay đổi của tham số tiếng nói và biểu diễn chúng dưới dạng số. Trên cơ sở mô hình tiếng nói LPC xấp xỉ mẫu tiếng nói đầu vào bằng cách tuyến tính một số ít mẫu tiếng nói trước đó. Đặc tính có nhớ này dẫn đến dạng sóng đầu ra có sự khác biệt với dạng sóng đầu vào. Những tín hiệu âm thanh có thay đổi nhanh theo thời gian thì không phù hợp với mô hình tiếng nói trên cơ sở LPC và do vậy rất dễ bị phá hủy.

- Mạch tự động điều chỉnh độ khuếch đại AGC:

Để đảm bảo âm lượng tiếng nói trong cuộc đàm thoại, mạng GSM sử dụng bộ AGC (Automatic Gain Control). Với mức trung bình tín hiệu đầu ra trong vòng phản hồi kín để điều khiển độ lớn biên độ đầu ra. Điều này dẫn đến biên độ của tín hiệu ra có thể khác so với tín hiệu vào.

- Bộ phát hiện tiếng nói – VAD[6]:

Thông thường, xen lẫn tín hiệu tiếng nói là những khoảng lặng. VAD (Voice Activity Detectors) phát hiện tiếng nói có trong dòng tín hiệu nghe hiểu và loại bỏ những khoảng lặng để tiết kiệm băng thông và năng lượng, như vậy, việc truyền dữ liệu có thể bị bỏ qua nếu đó là khoảng lặng. VAD phát hiện tiếng nói qua khai thác đặc tính tiếng nói là những xung (pulse) theo thời gian, trong khi nhiễu nền không có đặc tính này. Do vậy, khi biên độ tín hiệu vào không có tính chất đó, thì có thể bị loại bỏ bởi bộ lọc VAD. Tuy nhiên đây là nhận định có phần đơn giản, thực tế còn phức tạp hơn nhiều, thí dụ

tiếng nói hoàn toàn dùng âm vô thanh như tiếng nói thầm hoặc nói to nhưng lại nói kiểu như nói thầm thì VAD vẫn coi là tiếng nói.

- Hoạt động tự thích nghi của những bộ nén :

Những bộ mã này thực hiện theo những thuật toán đã biết, theo một mô hình rõ ràng và bù phần mất mát xuất hiện do bộ mã gây ra trong quá trình điều chế dữ liệu. Một vấn đề được đặt ra là: từ một tín hiệu nghe cho trước, chúng ta có thể mô hình hóa và dự báo những mất mát do bộ nén tiếng nói gây ra hay không. Rất không may, giải pháp này là không thực hiện được vì nhiều bộ nén thay đổi tham số và cơ chế trong khi đang hoạt động. Xem xét AMR GSM có nhiều chế độ hoạt động và nó sẽ lựa chọn chế độ hoạt động tốt nhất theo chất lượng đường truyền vô tuyến và năng lực có thể. Khi đường truyền liên kết vô tuyến hay hữu tuyến là xấu, nó giảm mã nguồn và gia tăng mã kênh để gửi đi tín hiệu tiếng nói chất lượng thấp qua đường truyền. Những nhà cung cấp có thể gia tăng số cuộc gọi thông qua dùng trạm cơ sở để buộc máy điện thoại sử dụng chế độ nén với tốc độ bit thấp. Vì bộ mã có thể thay đổi chế độ khi đang hoạt động và mỗi chế độ có những đặc tính riêng biệt, nên để dự báo chính xác những sai lệch mà bộ mã gây ra là không thể.

- Liên kết mạng không đồng nhất:

Tín hiệu thoại trải qua nhiều giai đoạn từ đầu cuối đến đầu cuối. Liên kết trung gian có thể chuyển đổi mã tín hiệu tiếng nói tùy theo năng lực, trạng thái đường truyền và băng thông cho phép, nên có thể gây ra những mất mát thông tin ở các khâu mà nó đi qua. Ví dụ, tín hiệu có thể chuyển đổi sang chế độ tốc độ bit thấp khi chuyển qua vệ tinh và cũng có thể từ GSM qua PSTN rồi ngược lại. Những chuyển đổi này có thể xảy ra ở bất kỳ giai đoạn nào và gây ra những mất mát thông tin đối với tín hiệu. Hơn nữa, ở mỗi thiết bị đầu cuối có thể sử dụng những thuật toán nén tiếng nói khác nhau, có các bộ lọc cân bằng equalizer bổ chính cho loa và tai nghe khác nhau vì thế tạo ra những mất mát trong kênh thoại không thể dự báo được.

### III. ĐIỀU CHẾ VÀ GIẢI ĐIỀU CHẾ ĐỂ TRUYỀN DỮ LIỆU QUA KÊNH GSM

#### 3.1 Phương pháp điều chế tín hiệu tựa tiếng nói

Phương pháp điều chế tín hiệu tựa tiếng nói speech-like waveform đã được thử nghiệm và mô tả trong một số bài báo của các tác giả khác nhau. Đây là phương pháp truyền dữ liệu dưới dạng tổng hợp thành tiếng nói và cơ bản sử dụng 3 đặc tính chính:

- Đường bao của phổ tiếng nói được biểu diễn bởi các tần số phổ vạch (LSF).
- Tần số cơ bản hoặc cao độ của giọng nói (pitch)
- Hình dạng và năng lượng kích thích ACELP

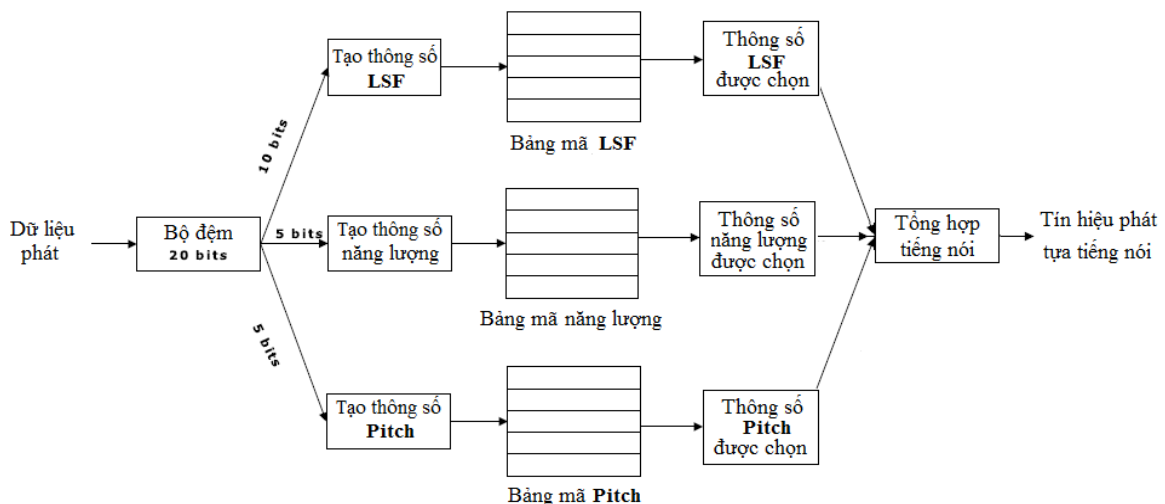
Các thông số nêu trên được bảo tồn khi truyền qua kênh thoại GSM và PSTN.

Dữ liệu đầu vào được ánh xạ tới các thông số trên bằng 3 bảng mã - codebook và sau đó được nhập vào bộ tổng hợp (xem hình 5). Tiếng nói tổng hợp này không phải là ngôn ngữ của bất kỳ cư dân nào trên thế giới mà nó chỉ có cùng tính chất của tiếng nói trên phương diện nén và giải nén mà thôi.

Tiếng nói với các thông số chủ yếu được tổng hợp và phát đi. Bên thu có bộ phân tích tiếng nói sẽ tách ra các thông số, kiểm tra tính tương thích rồi tra trong bảng mã để lấy ra dữ liệu.

Dữ liệu được gán như sau : 10 bit cho LSF, 5 bit cho cao độ và 5 bit cho năng lượng. Tổng cộng là 20 bit được truyền trong 20 ms. Điều này sẽ cho tốc độ bit là 1 kbps. Tốc độ bit cao hơn sẽ đạt được bởi các bảng mã lớn hơn.

Có hai nhiệm vụ chính cần thực hiện trong phương pháp này. Một là chọn loại mã hóa tiếng nói nào sẽ được sử dụng và hai là từ đó thiết kế các bảng mã. Do hệ thống GSM dùng mã nén tiếng nói theo thuật toán CELP – ACELP nên loại mã hóa cùng loại sẽ được chọn. Tại sao nên chọn mã nén cùng loại với hệ thống kênh truyền, là vì các thông số mà dữ liệu ánh xạ vào khi truyền qua hệ thống ít bị sai lệch hơn. Có thể chọn một trong số các mã GSM-HR (VSELP), GSM-EFR, GSM-ARM (ACELP),



Hình 5. Sơ đồ khối của phương pháp điều chế tín hiệu tựa tiếng nói

CELP... hoặc Speex. GSM-EFR là loại được ưu tiên lựa chọn vì việc triển khai đơn giản hơn. Từ loại mã nén được chọn sẽ quyết định việc thực hiện nhiệm vụ thứ hai là thiết kế bảng mã như thế nào.

Thiết kế bảng mã là công việc phức tạp và tốn nhiều thời gian nhất. Bảng mã thực hiện ánh xạ dữ liệu vào các thông số và sau đó nhập chúng vào bộ tổng hợp tiếng nói.

### 3.2 Phương pháp điều chế tín hiệu kiểu truyền thống có cấu trúc phổ gần giống phổ của tiếng nói.

#### a) Điều chế tín hiệu kiểu viễn thông truyền thống

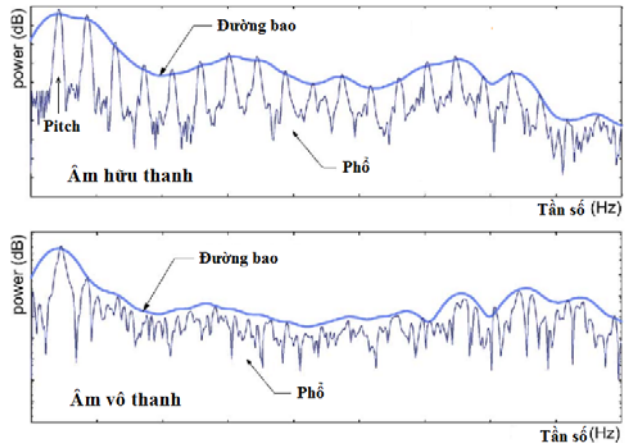
Điều chế theo phương thức viễn thông truyền thống được nhiều tác giả nghiên cứu. Thực nghiệm cho thấy điều chế (số) khóa pha (dịch pha) PSK tốt hơn so với điều chế (số) khóa biên độ (dịch biên) ASK và điều chế (số) khóa tần số (dịch tần) FSK. ASK thay đổi biên độ, trong trường hợp này, bộ mã hóa tiếng nói của GSM có AGC và nó phát hiện các thay đổi về biên độ này sẽ thực hiện việc bù, điều này sẽ gây ra lỗi trong máy thu. FSK cũng không phải là một lựa chọn tốt ở đây vì băng thông rất hạn chế (4 kHz). Điều chế (số) dịch pha vì sai DPSK thường được chọn vì tính đơn giản khi thực hiện và không cần bộ thu kết hợp.

Kênh bị giới hạn băng tần 4 kHz vì tần số lấy mẫu là 8 kHz. Trong hệ thống điện thoại thường có các bộ lọc thông thấp và thông cao, vì vậy tốt nhất trong thực tế chọn tần số sóng mang là tần số trung tâm của băng thông và có thể được xác định bằng các thiết bị đo lường. Với dải tần 300-3400Hz tần số sóng mang được chọn là 1.8 kHz.

Hạn chế của phương pháp điều chế tín hiệu kiểu viễn thông truyền thống là tốc độ truyền thấp và hiện tượng mất tín hiệu do VAD. Với GSM thời gian đáp ứng của bộ lọc dự đoán thời gian ngắn STP là 5 ms, bộ lọc dự đoán thời gian dài LTP là 20 ms. Như thế thời gian truyền một ký hiệu – symbol không dưới 5 ms. Tần số truyền ký hiệu cực đại sẽ là 200 Hz (1/5ms). Nếu dùng điều chế DPSK thì tốc độ truyền chỉ là 200 bps. Để tăng tốc độ truyền phải tăng mức điều chế và khi đó sai số BER sẽ tăng. Tác động của VAD cũng cần phải được xem xét. Đối với truyền dữ liệu không yêu cầu thời gian thực thì khoảng lặng xuất hiện không thành vấn đề, nhưng với yêu cầu truyền dữ liệu thời gian thực như mật mã thoại chẳng hạn thì không thể chấp nhận được. Để khắc phục người ta thường chèn những đoạn tín hiệu có tính xung để “đánh lừa” bộ VAD. Khi đó ta phải trả giá bằng tốc độ truyền giảm và không phải lúc nào giải pháp cũng thực hiện tốt.

#### b) Điều chế tín hiệu kiểu viễn thông truyền thống có cấu trúc phổ gần giống phổ của tiếng nói

Hình dưới cho thấy phổ của đoạn tiếng nói với âm hữu thanh là phổ có hình răng lược với tần số là bội nguyên lần của tần số của giọng nói pitch. Với đoạn âm vô thanh phổ là phổ của nhiễu. Đường bao trong cả hai trường hợp, là đặc tuyến tần số của các hốc cộng hưởng của cơ quan phát âm (bộ lọc cơ quan phát âm).



Hình 6. Phổ của âm hữu thanh và âm vô thanh

Kỹ thuật điều chế ghép kênh phân chia theo tần số trực giao – OFDM do R.W Chang phát minh năm 1966. OFDM là một trường hợp đặc biệt của phương pháp điều chế đa sóng mang, trong đó các sóng mang phụ trực giao với nhau, nhờ vậy phổ tính hiệu ở các sóng mang phụ cho phép chồng lấn lên nhau mà phía thu vẫn có thể khôi phục lại tín hiệu ban đầu. Sự chồng lấn phổ tín hiệu làm cho hệ thống OFDM có hiệu suất sử dụng phổ lớn hơn nhiều so với kỹ thuật điều chế thông thường.

Từ phân tích ở trên, chúng tôi lựa chọn phương pháp điều chế ghép kênh phân chia theo tần số trực giao - OFDM để truyền dữ liệu qua kênh GSM. OFDM đã được nhiều công trình trong và ngoài nước nghiên cứu khá kỹ càng cả trên phương diện lý thuyết và ứng dụng. Dưới đây chúng tôi xin trình bày nội dung liên quan đến việc lựa chọn các thông số và thực hiện điều chế OFDM sao cho có thể truyền dữ liệu qua kênh thoại GSM. Những vấn đề mới được đề cập là :

#### - Lựa chọn các thông số

Như đã nêu ở phần trên để tín hiệu truyền qua kênh thoại GSM được bảo toàn thì tín hiệu phải có cấu trúc phổ giống phổ của tiếng nói và trên phương diện khác cần phải xem xét các đặc điểm xử lý tiếng nói của hệ thống GSM.

°Thứ nhất dải phổ của OFDM phải nằm trong dải thoại 300 – 3400 Hz (khi truyền qua GSM)

°Thứ hai số vạch phổ không nên nhiều quá vì GSM sử dụng kỹ thuật nén dựa trên cơ sở LPC với bậc lọc cố định. Về mặt toán học người ta có thể biểu diễn một hàm đi qua n điểm cho trước bằng một đa thức có (n – 1) bậc. LPC trong GSM có bậc lọc là 10 nên số vạch phổ tốt nhất sẽ là 11. Tuy nhiên ở đây mỗi sóng mang con được điều chế số về pha hoặc biên độ với số mức là 2 (hoặc 4) nên số điểm có thể tăng lên mà ít ảnh hưởng đến chất lượng tín hiệu sau khi khôi phục. Ở đây số sóng mang con chúng tôi lựa chọn là 16 vạch.

°Thứ ba khoảng thời gian truyền một ký hiệu (symbol) không được ngắn hơn thời gian giữa 2 supeframe (bao gồm 4 frame) trong GSM là 20ms, tương ứng với tốc độ truyền symbol là 50Hz. Như vậy khoảng cách ngắn nhất

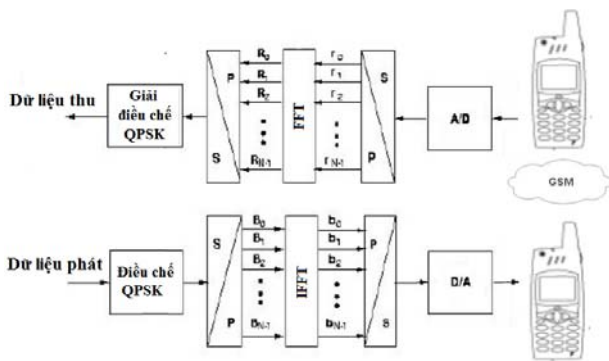
giữa các vạch phổ sóng mang phụ của OFDM là 50Hz. Trong trường hợp này ta chọn khoảng cách là 75Hz, để dư 25Hz làm khoảng bảo vệ.

Dải tần của OFDM sẽ là :  $75\text{Hz} \times 16 = 1200\text{Hz}$  ta chọn tần số trung tâm là 1500 Hz như vậy dải phổ của OFDM từ 900Hz đến 2100Hz thỏa mãn điều kiện thứ nhất là dải phổ nằm trong dải phổ của thoại GSM từ 300 đến 3400 Hz.

Thứ tư chọn phương thức điều chế. Như phân tích ở trên điều biên ảnh hưởng của bộ AGC, còn điều tần phổ quá rộng vì thế chỉ còn điều pha. Điều pha QPSK được chọn cho điều chế OFDM truyền qua kênh thoại GSM.

- Thực hiện điều chế OFDM với QPSK :

Việc thiết kế, chế tạo modem OFDM với QPSK có thể được thực hiện ngay trên PC theo sơ đồ dưới đây. Phần còn lại là bài toán kinh điển của OFDM, nó đã được nhiều tài liệu khác nhau thể hiện nên không được trình bày ở đây.



Hình 7. Sơ đồ nguyên lý modem QPSK – OFDM

**IV. KẾT LUẬN**

Về mặt lý thuyết thì phương pháp điều chế tín hiệu tựa tiếng nói để truyền dữ liệu qua kênh thoại GSM sẽ cho kết quả tốt nhất. Tuy nhiên trong thực tế rất khó thực hiện và nếu có thể thực hiện được thì chất lượng cũng không cao như đã phân tích ở trên. Qua các thực nghiệm của nhóm tác giả cho thấy khi lựa chọn điều chế tín hiệu bằng OFDM ta được tín hiệu điều chế có cấu trúc phổ gần giống phổ của tiếng nói, với ưu điểm không bị chặn bởi VAD, dễ thực hiện, kết quả truyền tin khá tốt. Trường hợp kênh truyền có băng thông tối đa  $BER < 0.05\%$ , trường hợp kênh truyền xấu  $BER$  không quá 2%. Dựa trên kết quả này, có thể khẳng định rằng với các chỉnh sửa codebook và các lựa chọn tham số của nhóm tác giả, chúng ta có thể sử dụng OFDM để truyền dữ liệu qua kênh thoại GSM.

Hướng tiếp theo là tích hợp toàn bộ modem này vào chip để có thể lắp vào điện thoại di động. Đây là công việc gian nan và cần có thời gian. Lập trình trên chip với không gian chật hẹp, tài nguyên hạn chế nên yêu cầu phải tối ưu hóa về tốc độ, về kích thước mã chương trình, về không gian vùng nhớ dữ liệu và vùng nhớ phục vụ thao tác tính toán. Hướng khác là tích hợp chức năng modem

vào phần mềm của điện thoại di động thông minh. Công việc này cũng khó khăn không kém công việc tích hợp vào chip .

**TÀI LIỆU THAM KHẢO**

- [1] John G.Proakis, Dimitris G.Manolakis. “Digital Signal Processing : Principles, Algorithms, and Applications”, 4th Edition. Prentice-Hall, Inc., 2007
- [2] A.M. Kondoz, “Digital Speech”, 2nd Edition, Wiley, 2004.
- [3] La Hữu Phúc, Lê Mỹ Tú. Truyền dữ liệu qua kênh thoại GSM với CD-FSK. Chuyên san Nghiên cứu Khoa học và Công nghệ trong lĩnh vực An toàn thông tin. Số 1.CS (01) 2015.
- [4] G. Cattaneo, G. De Maio, and U. F. Petrillo. Security issues and attacks on the GSM standard: a review. Journal of Universal Computer Science, vol. 19, no. 16, pp. 2437–2452, 2013.
- [5] ETSI, “Digital cellular telecommunications system (phase 2+); enhanced full rate (EFR) speech transcoding” (GSM 06.60 version 8.0.1 release 1999), Technical report, ETSI, 1999.
- [6] ETSI. Digital cellular telecommunications system (phase 2+); voice activity detector (VAD) for enhance full rate (EFR) speech traffic channels; (GSM 06.82 version 8.0.1 release 1999). Technical report, ETSI, 2000.
- [7] ETSI. Digital cellular telecommunications system (phase 2+); channel coding (3gpp ts 05.03 version 8.9.0 release 1999), Technical report, ETSI, 2005.
- [8] Mahdi Boloursaz Mashhadi, Fereidoon Behnia. Efficient codebook design for digital communication through compressed voice channels. IET Communications Volume 10, Issue 18, pp 2614-2620, 2016
- [9] Sigurdur Sverrisson, Xiaoyun Liang. *Digital Communication over Speech Compressed Channel*. CHALMERS UNIVERSITY OF TECHNOLOGY, Goteborg, Sweden. EXE028/2008.
- [10] N. Katugampala, S. Villette, and A.M. Kondoz. Secure voice over GSM and other low bit rate systems. Technical report, Center for Communication Systems Research, 2003.
- [11] Stephane Pierre Villette, Simon N. Fabri, and Ahmet M. Kondoz. Data transmission over a compressed speech channel. Patent Application: WO 03/071521 A1, Feb 2003.
- [12] N.N. Katugampala, K.T. Al-Naimi, S. Villette, and A.M. Kondoz. Real time data transmission over GSM voice channel for secure voice and data applications. Technical report, University of Surrey, United Kingdom, August 2005.
- [13] John G. Proakis and Masoud Salehi. Communication System Engineering, Second Edition Volume. Prentice Hall, 2002
- [14] Heinrich Meyr, Marc Moeneclaey, and Stefna A. Fecthel. Digital Communication Receivers. John Wiley & Sons, INC, 1998.

**A METHOD OF TRANSFER DATA OVER GSM VOICE CHANNEL**

**Abstracts:** Mobile phones and GSM technology are now very popular around the world. The demand for GSM voice signal security is urgent, while GSM's existing security solution has not yet reached the required safety level. The best solution for GSM voice signal security is to convert the signal into a digital form via a separate Vocoder, encrypt that digital signal using a strong encryption algorithm and transmit the encrypted signal through the GSM voice channel. In this contribution, we will present a solution for transmitting data via GSM voice channel which compatible with the properties of GSM voice channel, in

this case, they are GSM Vocoder and voice activity detection (VAD) for GSM.

**Keywords:** GSM technology, GSM Vocoder, GSM voice channel, GSM voice signal, OFDM, VAD.



**Nguyễn Thanh Bình**, Nghiên cứu sinh tại Học viện Công nghệ Bưu chính Viễn thông. Hiện công tác tại Vụ Khoa học Công nghệ - Ban Cơ yếu Chính phủ. Lĩnh vực nghiên cứu: Thiết kế thiết bị mật mã, công nghệ xử lý tín hiệu số và công nghệ truyền thông.



**Đặng Vân Trường**, Nghiên cứu sinh tại Học viện Công nghệ Bưu chính Viễn thông. Hiện công tác tại Viện Khoa học Công nghệ Mật mã - Ban Cơ yếu Chính phủ. Lĩnh vực nghiên cứu: Phát triển phần mềm bảo mật dữ liệu và bảo mật mạng, Công nghệ giao tiếp thiết bị ngoại vi.



**Trần Văn Liên**, kỹ sư vô tuyến điện. Hiện công tác tại Tổng công ty Điện tử và Tin học Việt Nam. Lĩnh vực nghiên cứu: Công nghệ xử lý tín hiệu số và công nghệ truyền thông.