GIÁO TRÌNH

LÝ THUYẾT THÔNG TIN

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

BÀI GIẢNG

LÝ THUYẾT THÔNG TIN

Biên soạn : PGS.Ts. NGUYỄN BÌNH

Lưu hành nội bộ

HÀ NỘI - 2006

LỜI NÓI ĐẦU

Giáo trình Lý thuyết thông tin là một giáo trình cơ sở dùng cho sinh viên chuyên ngành Điện tử – Viễn thông và Công nghệ thông tin của Học viện Công nghệ Bưu chính Viễn thông. Đây cũng là một tài liệu tham khảo hữu ích cho các sinh viên chuyên ngành Điện - Điện tử.

Giáo trình này nhằm chuẩn bị tốt kiến thức cơ sở cho sinh viên để học tập và nắm vững các môn kỹ thuật chuyên ngành, đảm bảo cho sinh viên có thể đánh giá các chỉ tiêu chất lượng cơ bản của một hệ thống truyền tin một cách có căn cứ khoa học.

Giáo trình gồm 6 chương, ngoài chương I có tính chất giới thiệu chung, các chương còn lại được chia thành 4 phần chính:

Phần I: Lý thuyết tín hiệu ngẫu nhiên và nhiễu (Chương 2)

Phần II: Lý thuyết thông tin và mã hóa (Chương 3 và Chương 4)

Phần III: Lý thuyết thu tối ưu (Chương 5)

Phần IV: Mật mã (Chương 6)

Phần I: (C***hương II***). Nhằm cung cấp các công cụ toán học cần thiết cho các chương sau.

Phần II: Gồm hai chương với các nội dungchủ yếu sau:

- Chương III: Cung cấp những khái niệm cơ bản của lý thuyết thông tin Shannon trong hệ truyền tin rời rạc và mở rộng cho các hệ truyền tin liên tục.

- Chương IV: Trình bày hai hướng kiến thiết cho hai định lý mã hóa của Shannon. Vì khuôn khổ có hạn của giáo trình, các hướng này (mã nguồn và mã kênh) chỉ được trình bày ở mức độ các hiểu biết cơ bản. Để có thể tìm hiểu sâu hơn những kết quả mới và các ứng dụng cụ thể sinh viên cần phải xem thêm trong các tài liệu tham khảo.

Phần III: (C***hương V***) Trình bày vấn đề xây dựng các hệ thống thu tối ưu đảm bảo tốc độ truyền tin và độ chính xác đạt được các giá trị giới hạn. Theo truyền thống bao trùm lên toàn bộ giáo trình là việc trình bày hai bài toán phân tích và tổng hợp. Các ví dụ trong giáo trình được chọn lọc kỹ nhằm giúp cho sinh viên hiểu được các khái niệm một cách sâu sắc hơn. Các hình vẽ, bảng biểu nhằm mô tả một cách trực quan nhất các khái niệm và hoạt động của sơ đồ khối chức năng của các thiết bị cụ thể

Phần VI: (C***hương VI***) Trình bày cơ sở lý thuyết các hệ mật bao gồm các hệ mật khóa bí mật và các hệ mật khóa công khai. Do khuôn khổ có hạn của giáo trình, một số vấn đề quan trọng còn chưa được đề cập tới (như trao đổi và phân phối khóa, xác thực, đảm bảo tính toàn vẹn …)

Sau mỗi chương đều có các câu hỏi và bài tập nhằm giúp cho sinh viên củng cố được các kỹ năng tính toán cần thiết và hiểu sâu sắc hơn các khái niệm và các thuật toán quan trọng.

Phần phụ lục cung cấp một số kiến thức bổ xung cần thiết đối với một số khái niệm quan trọng về một số số liệu cần thiết giúp cho sinh viên làm được các bài tập được ra ở các chương.

Giáo trình được viết dựa trên cơ sở đề cương môn học Lý thuyết thông tin do Bộ Giáo dục và Đào tạo và được đúc kết sau nhiều năm giảng dạy và nghiên cứu của tác giả. Rất mong được sự đóng góp của bạn đọc.

Các đóng góp ý kiến xin gửi về

KHOA KỸ THUẬT ĐIỆN TỬ 1 - HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

KM 10. ĐƯỜNG NGUYỄN TRÃI - THỊ XÃ HÀ ĐÔNG

Email: KhoaDT1@hn.vnn.vn

Hoặc nguyenbinh1999@yahoo.com

Cuối cùng tôi xin chân thành cảm ơn GS. Huỳnh Hữu Tuệ đã cho tôi nhiều ý kiến quý báu trong các trao đổi học thuật có liên quan tới một số nội dung quan trọng trong giáo trình này.

NGƯỜI BIÊN SOẠN

CHƯƠNG I: NHỮNG VẤN ĐỀ CHUNG VÀ NHỮNG KHÁI NIỆM CƠ BẢN

1.1. VỊ TRÍ, VAI TRÒ VÀ SƠ LƯỢC LỊCH SỬ PHÁT TRIỂN CỦA "LÝ THUYẾT THÔNG TIN"

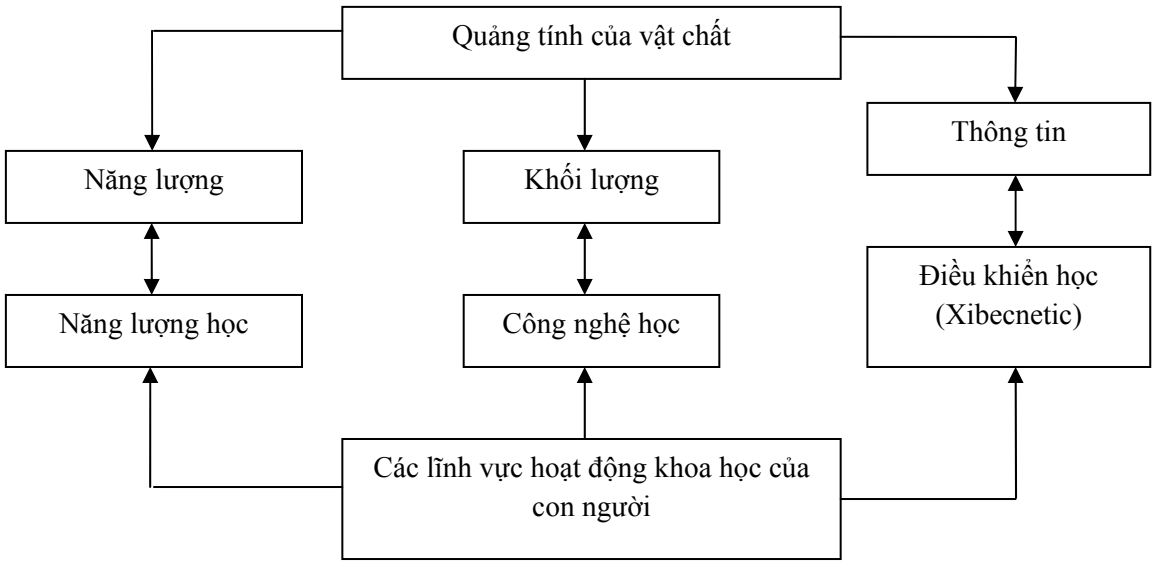
1.1.1. Vị trí, vai trò của Lý thuyết thông tin

Do sự phát triển mạnh mẽ của kỹ thuật tính toán và các hệ tự động, một ngành khoa học mới ra đời và phát triển nhanh chóng, đó là: "Lý thuyết thông tin". Là một ngành khoa học nhưng nó không ngừng phát triển và thâm nhập vào nhiều ngành khoa học khác như: Toán; triết; hoá; Xibecnetic; lý thuyết hệ thống; lý thuyết và kỹ thuật thông tin liên lạc… và đã đạt được nhiều kết quả. Tuy vậy nó cũng còn nhiều vấn đề cần được giải quyết hoặc giải quyết hoàn chỉnh hơn.

Giáo trình " Lý thuyết thông tin" này (còn được gọi là "Cơ sở lý thuyết truyền tin") chỉ là một bộ phận của lý thuyết thông tin chung – Nó là phần áp dụng của "Lý thuyết thông tin" vào kỹ thuật thông tin liên lạc.

Trong các quan hệ của Lý thuyết thông tin chung với các ngành khoa học khác nhau, ta phải đặc biệt kể đến mối quan hệ của nó với ngành Xibecnetic.

Mối quan hệ giữa các hoạt động khoa học của con người và các quảng tính của vật chất được mô tả trên hình (1.1).



Hình 1.1. Quan hệ giữa hoạt động khoa học và quảng tính của vật chất

- Năng lượng học: Là một ngành khoa học chuyên nghiên cứu các vấn đề liên quan tới các khái niệm thuộc về năng lượng. Mục đích của năng lượng học là làm giảm sự nặng nhọc của lao động chân tay và nâng cao hiệu suất lao động chân tay. Nhiệm vụ trung tâm của nó là tạo, truyền, thụ, biến đổi, tích luỹ và xử lý năng lượng.

- Xibecnetic: Bao gồm các ngành khoa học chuyên nghiên cứu các vấn đề có liên quan đến khái niệm thông tin và tín hiệu. Mục đích của Xibecnetic là làm giảm sự nặng nhọc của trí óc và nâng cao hiệu suất lao động trí óc. Ngoài những vấn đề được xét trong Xibecnetic như đối tượng, mục đích, tối ưu hoá việc điều khiển, liên hệ ngược. Việc nghiên cứu các quá trình thông tin (như chọn, truyền, xử lý, lưu trữ và hiển thị thông tin) cũng là một vấn đề trung tâm của Xibecnetic. Chính vì vậy, lý thuyết và kỹ thuật thông tin chiếm vai trò rất quan trọng trong Xibecnetic.

- Công nghệ học: gồm các ngành khoa học tạo, biến đổi và xử lý các vật liệu mới. Công nghệ học phục vụ đắc lực cho Xibecnetic và năng lượng học. Không có công nghệ học hiện đại thì không thể có các ngành khoa học kỹ thuật hiện đại.

1.1.2. Sơ lược lịch sử phát triển

Người đặt viên gạch đầu tiên để xây dựng lý thuyết thông tin là Hartley R.V.L. Năm 1928, ông đã đưa ra số đo lượng thông tin là một khái niệm trung tâm của lý thuyết thông tin. Dựa vào khái niệm này, ta có thể so sánh định lượng các hệ truyền tin với nhau.

Năm 1933, V.A Kachenhicov chứng minh một loạt những luận điểm quan trọng của lý thuyết thông tin trong bài báo "Về khả năng thông qua của không trung và dây dẫn trong hệ thống liên lạc điện".

Năm 1935, D.V Ageev đưa ra công trình "Lý thuyết tách tuyến tính", trong đó ông phát biểu những nguyên tắc cơ bản về lý thuyết tách các tín hiệu.

Năm 1946, V.A Kachenhicov thông báo công trình "Lý thuyết thế chống nhiễu' đánh dấu một bước phát triển rất quan trọng của lý thuyết thông tin.

Trong hai năm 1948 – 1949, Shanon C.E công bố một loạt các công trình vĩ đại, đưa sự phát triển của lý thuyết thông tin lên một bước tiến mới chưa từng có. Trong các công trình này, nhờ việc đưa vào khái niệm lượng thông tin và tính đến cấu trúc thống kê của tin, ông đã chứng minh một loạt định lý về khả năng thông qua của kênh truyền tin khi có nhiễu và các định lý mã hoá. Những công trình này là nền tảng vững chắc của lý thuyết thông tin.

Ngày nay, lý thuyết thông tin phát triển theo hai hướng chủ yếu sau:

Lý thuyết thông tin toán học: Xây dựng những luận điểm thuần tuý toán học và những cơ sở toán học chặt chẽ của lý thuyết thông tin. Cống hiến chủ yếu trong lĩnh vực này thuộc về các nhà bác học lỗi lạc như: N.Wiener, A. Feinstain, C.E Shanon, A.N. Kanmôgorov, A.JA Khintrin.

Lý thuyết thông tin ứng dụng: (lý thuyết truyền tin)

Chuyên nghiên cứu các bài toán thực tế quan trọng do kỹ thuật liên lạc đặt ra có liên quan đến vấn đề chống nhiễu và nâng cao độ tin cậy của việc truyền tin. Các bác học C.E Shanon, S.O RiCe, D. Midleton, W. Peterson, A.A Khakevich, V. Kachenhicov đã có những công trình quý báu trong lĩnh vực này.

1.2. NHỮNG KHÁI NIỆM CƠ BẢN - SƠ ĐỒ HỆ TRUYỀN TIN VÀ NHIỆM VỤ CỦA NÓ

1.2.1. Các định nghĩa cơ bản

1.2.1.1. Thông tin

Định nghĩa: Thông tin là những tính chất xác định của vật chất mà con người (hoặc hệ thống kỹ thuật) nhận được từ thế giới vật chất bên ngoài hoặc từ những quá trình xảy ra trong bản thân nó.

Với định nghĩa này, mọi ngành khoa học là khám phá ra các cấu trúc thông qua việc thu thập, chế biến, xử lý thông tin. ở đây "thông tin" là một danh từ chứ không phải là động từ để chỉ một hành vi tác động giữa hai đối tượng (người, máy) liên lạc với nhau.

Theo quan điểm triết học, thông tin là một quảng tính của thế giới vật chất (tương tự như năng lượng, khối lượng). Thông tin không được tạo ra mà chỉ được sử dụng bởi hệ thụ cảm. Thông tin tồn tại một cách khách quan, không phụ thuộc vào hệ thụ cảm. Trong nghĩa khái quát nhất, thông tin là sự đa dạng. Sự đa dạng ở đây có thể hiểu theo nhiều nghĩa khác nhau: Tính ngẫu nhiên, trình độ tổ chức,…

1.2.1.2. Tin

Tin là dạng vật chất cụ thể để biểu diễn hoặc thể hiện thông tin. Có hai dạng: tin rời rạc và tin liên tục.

Ví dụ: Tấm ảnh, bản nhạc, bảng số liệu, bài nói,… là các tin.

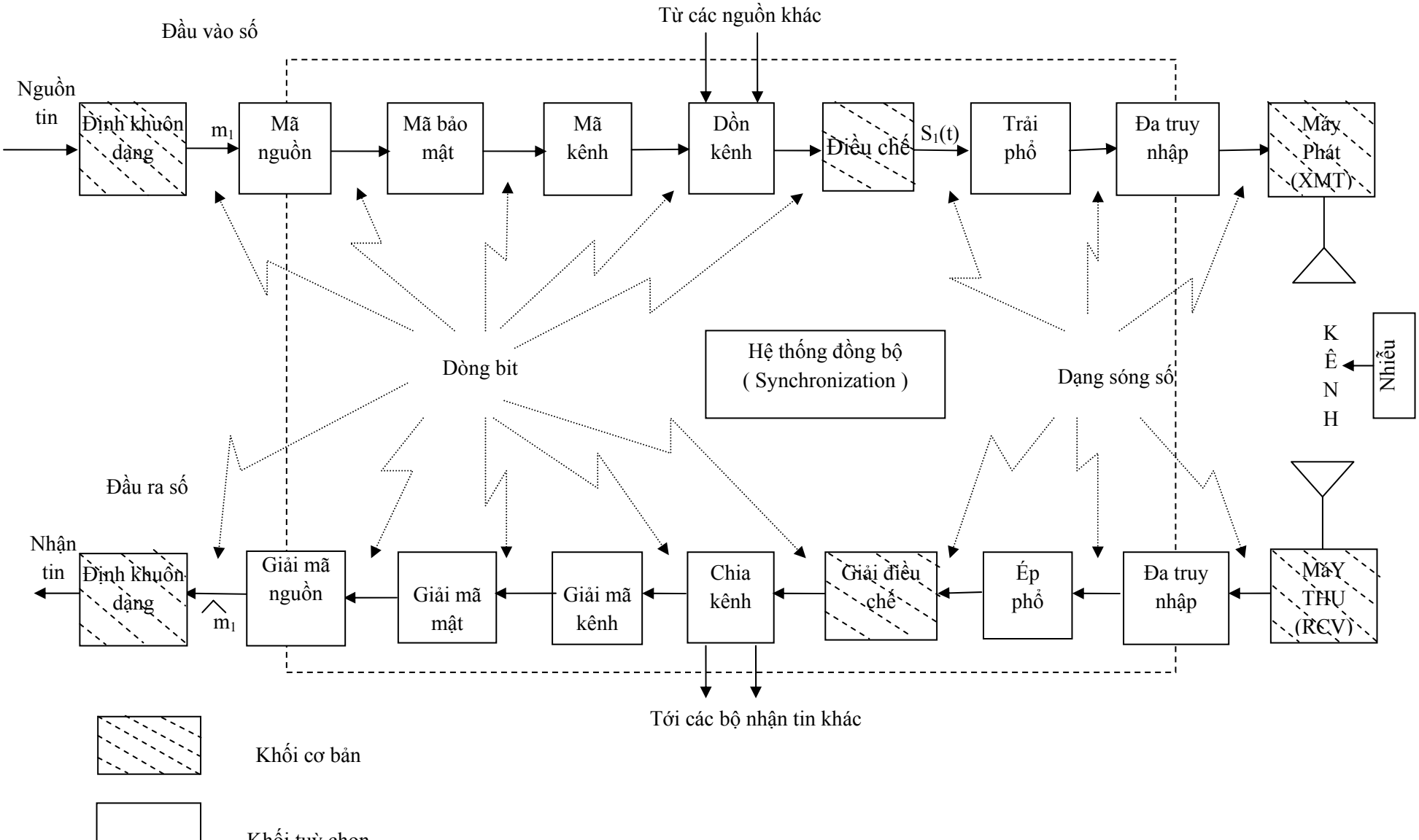
1.2.1.3. Tín hiệu

Tín hiệu là các đại lượng vật lý biến thiên, phản ánh tin cần truyền.

Chú ý: Không phải bản thân quá trình vật lý là tín hiệu, mà sự biến đổi các tham số riêng của quá trình vật lý mới là tín hiệu.

Các đặc trưng vật lý có thể là dòng điện, điện áp, ánh sáng, âm thanh, trường điện từ

1.2.2. Sơ đồ khối của hệ thống truyền tin số (Hình 1.2)



Hình 1.2. Sơ đồ khối hệ thống truyền tin số.

1.2.2.1. Nguồn tin

Nơi sản ra tin:

- Nếu tập tin là hữu hạn thì nguồn sinh ra nó được gọi là nguồn rời rạc.

- Nếu tập tin là vô hạn thì nguồn sinh ra nó được gọi là nguồn liên tục.

Nguồn tin có hai tính chất: Tính thống kê và tính hàm ý.

Với nguồn rời rạc, tính thống kê biểu hiện ở chỗ xác suất xuất hiện các tin là khác nhau.

Tính hàm ý biểu hiện ở chỗ xác suất xuất hiện của một tin nào đó sau một dãy tin khác nhau nào đó là khác nhau.

Ví dụ: $P(y/ta) \neq P(y/ba)$

1.2.2.2. Máy phát

Là thiết bị biến đổi tập tin thành tập tín hiệu tương ứng. Phép biến đổi này phải là đơn trị hai chiều (thì bên thu mới có thể "sao lại" được đúng tin gửi đi). Trong trường hợp tổng quát, máy phát gồm hai khối chính.

- Thiết bị mã hoá: Làm ứng mỗi tin với một tổ hợp các ký hiệu đã chọn nhằm tăng mật độ, tăng khả năng chống nhiễu, tăng tốc độ truyền tin.

- Khối điều chế: Là thiết bị biến tập tin (đã hoặc không mã hoá) thành các tín hiệu để bức xạ vào không gian dưới dạng sóng điện từ cao tần. Về nguyên tắc, bất kỳ một máy phát nào cũng có khối này.

1.2.2.3. Đường truyền tin

Là môi trường vật lý, trong đó tín hiệu truyền đi từ máy phát sang máy thu. Trên đường truyền có những tác động làm mất năng lượng, làm mất thông tin của tín hiệu.

1.2.2.4. Máy thu

Là thiết bị lập lại (sao lại) thông tin từ tín hiệu nhận được. Máy thu thực hiện phép biến đổi ngược lại với phép biến đổi ở máy phát: Biến tập tín hiệu thu được thành tập tin tương ứng.

Máy thu gồm hai khối:

- Giải điều chế: Biến đổi tín hiệu nhận được thành tin đã mã hoá.

- Giải mã: Biến đổi các tin đã mã hoá thành các tin tương ứng ban đầu (các tin của nguồn gửi đi).

1.2.2.5. Nhận tin

Có ba chức năng:

- Ghi giữ tin (ví dụ bộ nhớ của máy tính, băng ghi âm, ghi hình,…)

- Biểu thị tin: Làm cho các giác quan của con người hoặc các bộ cảm biến của máy thụ cảm được để xử lý tin (ví dụ băng âm thanh, chữ số, hình ảnh,…)

- Xử lý tin: Biến đổi tin để đưa nó về dạng dễ sử dụng. Chức năng này có thể thực hiện bằng con người hoặc bằng máy.

1.2.2.6. Kênh truyền tin

Là tập hợp các thiết bị kỹ thuật phục vụ cho việc truyền tin từ nguồn đến nơi nhận tin.

1.2.2.7. Nhiễu

Là mọi yếu tố ngẫu nhiên có ảnh hưởng xấu đến việc thu tin. Những yếu tố này tác động xấu đến tin truyền đi từ bên phát đến bên thu. Để cho gọn, ta gộp các yếu tố tác động đó vào một ô trên hình 1.2.

Hình 1.2 là sơ đồ khối tổng quát nhất của một hệ truyền tin số. Nó có thể là: hệ thống vô tuyến điện thoại, vô tuyến điện báo, rađa, vô tuyến truyền hình, hệ thống thông tin truyền số liệu, vô tuyến điều khiển từ xa.

1.2.2.8. Các phương pháp biến đổi thông tin số trong các khối chức năng của hệ thống