통신로 용량(Channel Capacity)의 문제

陳 庸 玉 〈경희대 교수·통신공학/본지 편집고문〉

통보문이나 기호가 생성되는 엔트로피(또는 정보)는 과정의 통계적 특성에 의해 결정되며 통보문의 통계특성은 전적으로 정보원의 특성에 의존한다.

전호(88년 5월호)에서는 정보량과 평균정보량(엔트로피)등에 대해 설명했다. 이번호에서는 통신로의 용량에 대해 알아보기로 한다.

88년 1월호의 〈그림 2〉와 같이 디지틀통신모델에서 통신로 용량이란 「주어진 정보원에서 발생되는 정보를 전송하는 능력」이라 표현할 수 있다.

만일 정보원의 모든 기호가 동일한 시간주기(예 : 전신)를 갖는 간단한 종류이고, 선택된 각 기호는 s비트로 표시되고(2개중 자유롭게 선택하는 2진 기호의 경우) 통신로가 초당 n기호를 전송할 수 있다면 채널 용량 C는 ns bps로 정의된다.

보다 일반적인 경우에는 여러기호의 다양한 길이를 고려해 넣어야 한다. 따라서 통신로 용량의 일반적 표현은 어떤 시간주기에 있어서 다루어진 기호의 갯수를 포함하게 된다. 그러므로 일반적인 경우의 채널용량은 초당 전송되는 기호의 갯수가 아니라 정보의 양으로 표시되며 단위는 bps이다. 예를 들어 1초에 5자를 보낼 수 있는 전송로가 있다면 1자에 5비트(88년 5월호 참조)씩 표현되므로 통신로 용량은 25bps이다.

1. 부호화(Encoding)와 통신로 용량의 관계

송신기는 우선 통보문(Message)을 통신로를 통해 수신기로 전송될 신호의 형태로 바꾼다.

전화의 경우 송신기는 가청신호, 즉 전류라는 형태로 전송하게 된다. 그러나 송신기는 통보문에 대한 신호를 발생시키기 위해 더 복잡하게 작동할 수도 있다. 예를 들어 통보문을 암호화 하기 위해 어떤 부호를 사용하여 특정한 숫자의 열로 바꾸기도 한다.

그리하여 보통 송신기의 기능은 부호화하는 것이고 수신기의 기능은 복호(Decode)화 하는 것이다.

불연속 기호의 경우 잡음이 없는 통신로에서의 전송 기초이론에 대해 알아보기로 하자. 이 이론은 매초당 C비트의 용량을 가지고 매초당 H비트의 엔트로피(또는 정보)를 발생하는 정보원으로 부터 신호를 받는 통신로에 관한 것이다. C/H의 평균비율에 따라 기호가 전송되도록 적당한 부호화과정을 고안해 그 부호화가 아무리 잘되어도 C/H를 초과할 수는 없다.

잡음이 존재하는 경우에는 이 용량이 변화하게 되는데 잡음이 없다고 가정할 때는 부호화의 역할에 대하여 알아야 한다.

통보문이나 기호가 생성되는 엔트로피(또는 정보)는 과정의 통계적 특성에 의해 결정되며 통보문의 통계특성은 전적으로 정보원의 특성에 의존한다. 그러나 통신로에 의해 실제로 전송되는 신호의 통계적 특성, 즉 통신로의 엔트로피는 채널에 보내고자 하는 기호나 신호의 성질에 따라 다르다. 예를 들어 전신에서는 점과 점 사이, 점과 선 사이, 선과 선 사이에 간격을 두어야 하며 그렇지 않을 경우 점과 선은 인지될 수 없다.

통신로가 신호의 완전한 자유성을 제한할 때 제약받지 않는 신호구조의 경우보다 더 큰 신호 엔트로피로 유도하는 신호의 통계적 특성이 있게 되고, 이때 신호 엔트로피는 통신로 용량과 똑같게 된다.

따라서 가장 효율적인 부호화에 대해 규명할 수 있다. 실제로 최상의 송신기는 신호가 사용되는 채널에 가장 잘 맞는 적절한 통계적 특성을 갖도록 통보문을 부호화하는 데 있다. 즉 신호의 엔트로피를 최대화하고 채널용량 C와 같게 하는 것이다.

이상의 기초이론에 따른 부호화 방법은 기호전송에 대해 C/H를 최대비율이 되게 하는 것이다. 이 전송률에 도달하기 위해서는 그만한 댓가를 치루어야 한다. 부호화를 보다 이상적으로 하고자 할수록 부호화 과정에 있어서 더 많은 시간이 걸린다. 전자기기에서 「길다」라는 의미는 몇 분의 1초를 의미하지만 부호화로 인해 손실된 시간에 반해 전송에 대한 이익과 균형을 이루어야 하는 난점이 있다.

2. 잡음(Noise)의 영향과 통신로 용량의 변화

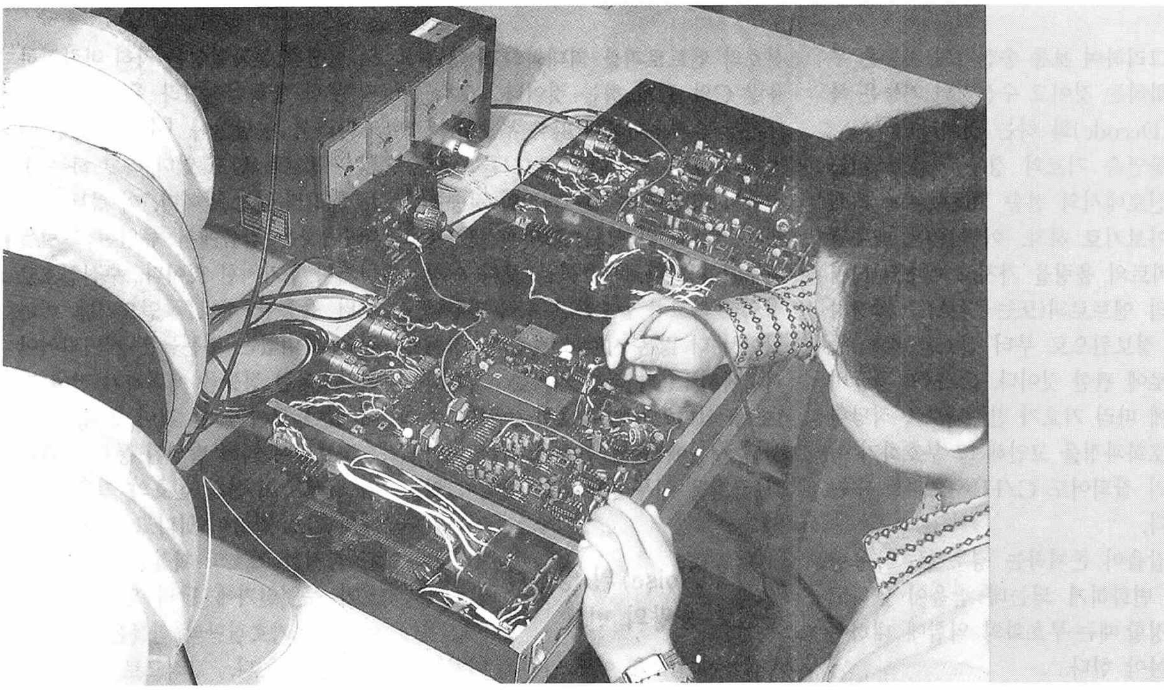
잡음이 정보에 어떤 영향을 미치는가? 되새겨 보면 정보란 통보문중에서 하나를 선택함에 있어서 선택의 자유 척도이다. 이 선택의 자유가 크면 클수록, 따라서 정보가 가치있게 되면 될수록 실제로 선택되는 통보문은 불확실성이 큰 상태였다. 선택의 자유가 증대되고 불확실성이 증가하고, 정보의 가치가 증대되는 것은 같은 맥락으로 간주할 수 있다.

만일 잡음이 유입되면 수신된 통보문이 왜곡이나 착오 또는 이물질을 함유한 것과 같게 됨으로써 잡음으로 인해 불확실성이 증가되었다고 말할 수 있다. 그러나 불확실성이 증가하게 되면 정보량도 증가하게 된다. 이것은 또한 마치 잡음이 유용하다는 것처럼 이해될 수 있을 것이다. 그러나 일반적으로 잡음이 존재할 경우 수신신호는 송신신호보다 더 많이 변화된 집합에서 선택된다는 사실이다. 송신자의 편에서 볼 때 선택의 자유로 발생될 불확실성은 바람직한 불확실성이지만, 잡음이나 착오의 영향으로 발생될 불확실성은 가히 바람직하지 못한 불확실성이라 할 수 있다.

따라서 농담삼아 「수신신호가 더 많은 정보를 담고 있다」라고 하는 것도 틀리지 않는 말이다. 이 정보 중에 어떤 것은 잡음에서 유입된 그럴듯하거나 무가치한 것이다. 수신신호로부터 유용한 정보를 얻으려면 이와 같은 의미없는 부분들은 제거되어야 한다. 이런 점들을 제거하기 전에 먼저 알아 놓아야 할 것이 있다.

정보원에 의해 발생된 통보문 기호와 실제로 수신된 신호의 기호 등 2개의 기호 집합이 있다고 하자. 수신된 특정한 기호의 확률은 보내진 기호가 무엇인가에 달려 있기 때문에 이들 두 기호집단의 확률은 상호 연관성을 갖는다. 그러므로 잡음이나 별다른 착오가 없을 경우 수신된 기호가 정확하게 송신된 신호에 대응할 것이고 가능한 착오의 존재 여건 하에서는 수신된 기호가 송신된 신호에 거의 완전하게 대응하거나 밀접한 관련을 갖기 위해 일정한 확률적인 값이 존재하게 된다.

이제 이 시점에서 다른 기호들과 관련이 있는 하나의 기호 집합의 엔트로피라는 것을 계산할 수 있을 것이다. 예를 들어 신호와 관련된 통보문의 엔트로피에 대하여 알아보자. 우선 실제로 받아들여진 어떤 신호의 기호가 있다고 가정하자. 각각의 통보문 기호는 일정한 확률값을 갖는데 그 확률값은 수신된 신호와 기호 값보다는 비교적 동등하거나 크지만 다른 것에 비하여는 비교적 작다. 이러한 확률들의 집합을 이용하여 시험적인 엔트로피값을 구해보자. 이것은 미리 알려져 수신된 기호나 신호 기호의 가정하에 이루어진 통보문 엔트로피이다. 아무리 좋은 조건 일지라



도 이 값은 작은 값을 갖게 될 것이다. 왜냐하면 내포된 확률이 여러 가지 경우로 넓게 퍼져 있는 것이 아니라 하나 또는 몇 가지 경우에만 한정된 것이기 때문이다. 잡음이 완전히 없는 경우라면 그 값은 0이 되며 따라서 신호의 기호가 이미 알려진 경우 1이라는 확률을 갖는 한가지 기호(즉 수신된 한 가지 기호)를 제외한 메시지의 확률은 모두 0이 될 수밖에 없다.

이때 신호 기호가 수신된다면 각각의 가정에 따라 통보문 엔트로피 중 하나의 값을 계산할수 있다. 전부 계산한 후 추측한 신호 기호의 확률에 따라 그들의 평균을 구한다. 계산하고자 하는 기호의 집합이 2일 경우 이런 식으로 산출된 엔트로피를 상대 엔트로피라 한다 (88.5월에서 규정한 의미와 동일). Shannon은 이와 같은 특별한 경우로 신호와 관련된 통보문의 엔트로피를 모호성(Equivocation)이라 칭했다.

이런 방법으로 계산된 모호성에서 의미를 파악할 수 있다. 이는 신호가 알려졌을 때 메시지에 담긴 평균적 불확실성을 측정하게 된다는 것이다. 알고 있는 신호에 대하여 아무런 잡음도 유입되지 않는다면 그 통보문에 관해서는 불확실성이 있을 수 없다. 신호가 알려진 후에도 어떤 잔여분의 불확실성이 정보원 내에 있을 경우 이것은 잡음으로 인한 불필요한 불확실성임이 틀림없다.

이를 요약하면 「수신된 기호가 알려졌을 때 통보문의 정보원 내의 평균 불확실성」이라 할 수 있으며 비슷한 표현이지만 「알려진 통보문이 보내졌을 때의 수신된 신호에 대한 평균적 불확실성」이라고도 할 수 있다. 물론 잡음이 없을 경우에는 후자의 불확실성 또한 0이 된다.

이러한 양들 사이의 관계인 공식을 증명하는 것은 용이하다.

$$H(x)-H_y(x)=H(Y)-H_x(Y) \cdots\cdots (1)$$

여기서 $H(x)$는 정보원의 엔트로피이거나 정보량이고, $H(y)$는 수신된 신호의 정보 또는 엔트로피이다. $H_y(x)$이 알려진 신호라면 정보원 내에서의 불확실성이라 할 수 있고, $H_x(y)$는 보내진 통보문이 알려졌을 경우 수신된 신호 내에서의 불확실성이거나 잡음때문에 생긴 수신된 신호 정보라고 할 수 있다. 이때 이 등식의 우변은 잡음의 악영향에도 불구하고 전송된 유용한 정보를 나타낸다.

이제 잡음 채널의 용량 C가 의미하는 것을 설명할 수 있을 것이다. 사실 이것은 채널을 통하여 전송될 수 있는 유용한 정보(예 : 총불확실성——잡음의 불확실성)의 최대 비율과 동등하다고 정의된다.

다시 말해서 유용한 정보전송비는

적당한 부호화 작업을 사용함으로써 최대화 할 수 있는 의미이다.

마지막으로 잡음채널에 대한 기초 이론에 대하여 알아 보자. 위에서 설명했듯이 잡음 채널의 용량은 C라고 가정하고 엔트로피 H(x)bps, H(y)bps로 수신되었다고 하자. 채널 용량 C가 H(x)보다 크거나 같을 경우에 적당한 부호화 시스팀에 의하여 고안된 정보원의 출력은 각자가 원하는 만큼 작은 착오만을 갖고 통신로를 통하여 전송될 수 있다. 그러나 특정한 착오가 아무리 작아도 요구에 만족하는 부호가 나올 수 있다. 만일 C가 통보문을 수용하는 정보원의 엔트로피인 H(x)보다 작다면 착오빈도를 원하는 수준으로 줄일 수 있는 부호화 방안을 고안한다는 것은 불가능하다.

부호화 과정에 아무리 익숙하더라도 신호가 수신되고 난 뒤에는 실질적인 통보문에 대하여 불필요한(잡음) 불확실성이 존재하는 것은 어쩔 수 없다. 또한 이 불필요한 불확실성 ―모호성(Equivocation)은 H(x)−C보다 크거나 같을 것이다. 더 나아가서 통보문에 대하여 작은 양이지만 H(x)−C의 값을 초과하지 않도록 불확실성을 감소시킬 수 있는 부호가 적어도 하나는 항상 존재하게 된다.

3. 연속적인 통보문 (Continuous Message)의 경우 통신로 용량

이제까지 이상적인 기호의 형태인 철자로 구성된 단어나 단어로 구성된 문장, 음표(음)로 구성된 멜로디, 한정된 불연속 점들로 이루어진 망형태의 판화등과 같은 메시지에 관하여 언급했다. 그런데 음의 고저나 강도 등도 계속해서 바뀌는 말과 같은 연속적인 메시지의 경우 어떤 이론이 형성될까?

예를 들어 현재의 전화선은 약 4,000Hz의 주파수까지 전송할 수 있다면 만족스러운 통신을 할 수 있다. 또 10,000~12,000정도의 주파수로는 교향악을 고충실도로 무선전송할 수 있다.

이 연속신호에 대한 매우 편리한 수학적 정리로서 주기시간이 T초이고 주파수 대역이 0에서 W까지의 대역으로 한정된 연속신호 2TW로 이산화시킬 수 있다(이 비율을 나이퀴스트율이라 한다). 보통 연속곡선은 유한개의 이산점을 연결한 것으로 근사화될 수 있으나 그 곡선에 대한 완전한 정보를 얻으려면 무한개의 수가 필요하다.

제한된 주파수 내에서는 유한 개의 매개변수만이 있어도 가능하며, 이는 무한개의 변수를 취급해야 하는 복잡한 상황에서 유한 개의 변수만을 다룸으로써 상당히 단순한 상황에 이르기까지의 연속신호로 인한 통신의 문제를 줄일 수 있다는 이점이 있다.

연속신호에 대한 이론에 있어서 주파수 대역폭이 W일 때 통신로의 최대 용량을 C라 하고, 전송에 필요한 평균전력을 P라고 하면, 섀논이 규정한 바에 따라 통신로는 잡음전력 N에 영향을 받는다. 이 잡음은 특별한 「백열잡음(White Thermal Noise)」으로서 이 백열잡음 자체는 가우스성 확률분포를 갖는다. 이러한 조건하에서 섀논에 의하면 최상으로 부호화했을 때 전송이 가능한 이진 전송률은

$$C=W\log_2\frac{S+N}{N}=W\log\left(1+\frac{S}{N}\right)\cdots(2)$$

로 전송할 수 있다. 그러나 아무리 부호화를 잘 시켜도 일정한 주파수를 올리지 않고서는 이 이상으로 전송하는 것은 불가능하다. 위에서 가정한 특별한 경우 「백열잡음」이 아닌 임의의 잡음인 경우 섀논은 통신로 용량에 대해 명백한 공식을 유도하지는 못했으나 통신로 용량의 상한과 하한을 구했다. 또 그는 송신기의 평균전력이 아니라 최대 순간전력만을 명시하여 채널 용량의 한계를 유도해 냈다.

섀논이 구한 결과는 다소 상세하지 못한 감이 있으나 매우 중요한 것임에는 틀림없다. 즉 일반적인 종류의 연속적인 통보문이나 신호에 대하여 수신된 통보문의 충실도, 정보원이 만들어 내는 정보 비율의 개념, 전송비율, 통신로 용량 등을 규명했다.

4. 결　　어

이상은 W. Weaver의 「최근통신의 수학적 이론에 대한 고찰(1949)」 2장에서 통신로 용량 부분을 발췌요약한 것이다. 상당한 수학적 개념이 요구되므로 어려운 감이 있으나 이에 대한 명확한 이해없이는 컴퓨터나 통신 특히 정보흐름에 대한 이해가 불가능하다고 할 수 있다. 왜냐하면 정보기기를 다루는 부분에서 정보의 본질적 이해가 선행되는 것은 필수적인 것이기 때문이다.

특히 연속전송로의 통신로 용량식은 애널로그와 디지틀 통신로를 이어주는 가교적 역할을 하는 수식으로 평가되므로 이에 대한 명확한 이해가 필요하다.

(2)식에서보면 연속통신로의 통신로 용량은 대역폭 W에 비례하지만 신호대 잡음비(수식에서는 S/N)의 항은 대수비에 비례함을 알 수가 있다. 따라서 대역폭이 고정되었을 때는 S/N만이 변수임을 보여주고 있다. ■