

論文96-33A-10-4

# 통신에 있어서 서비스품질 측정방법

## (Methodology on measuring the quality of service in telecommunications)

李永煥\*, 權世赫\*\*

(Young Hwan Lee and Se Hyug Kwon)

### 요 약

통신에 있어서 사용자가 느끼는 서비스 품질을 설문분석을 통하여 측정하는 방법을 제시하였다. 설문분석을 통하여 사용자의 주관적인 평가를 측정한 후, 그 평가를 계량화하는 방법으로 MOS 방법을 이용하였다. 그러나 현재 널리 사용되는 일반적인 MOS 방법에는 문제가 있으므로 본 연구에서는 문제점을 언급하고 그 대안으로 제안되어 있는 누적 퍼센트 곡선을 통신에 있어서 서비스 품질 측정방법으로 제시하였으며, 그 응용 예로 국내 그룹3 팩시밀리의 화상 품질을 측정하였다.

### Abstract

The methodology is suggested to measure the quality of service in telecommunications. Through questionnaire, the subjective evaluation of users is surveyed, and quantified by MOS method. The MOS method has some problems in interpretation which are mentioned in this paper. The cumulative percent curve is proposed to evaluate the quality of service. As an example, the image quality of G3 facsimile is measured.

### I. 서 론

서비스 품질 (QOS: Quality Of Service)은 통신에 있어서 중요한 개념이다. 일반적으로 품질은 "fitness for use"나 "loss from expectation"로 정의되는데 전자는 공급자, 후자는 이용자 측면에서 본 품질로 해석되고 있다. 아직 통신에 있어서 선진국의 기술을 이전 받아 서비스 제공에 만족하고 있는 우리나라에서는 공급자에 의해 제공 가능한 서비스 수준, 국제통신기구에서 권장한 표준이나 그 표준을 근거로 제정된 기술

기준을 만족하는 정도의 서비스 수준만이 이용자에게 제공되고 있는 실정이다. 그러나 WTO 체제가 출범함에 따라 세계경제가 무한경쟁시대로 돌입함에 따라 통신 시장의 개방도 불가피하게 됨에 따라 통신 서비스 품질을 사용자가 만족할 수 있는 수준까지 높이지 않으면 시장 경쟁력을 상실하게 될 것이다. 서비스 품질에 대한 평가는 공급자와 이용자 사이에 이견을 보이는 것은 당연하다. 한 예로 현재 제공되는 있는 아날로그 이동전화 서비스를 살펴보면, 많은 이용자는 이동전화 서비스에 대한 불만을 호소하고 있으나 공급자는 제공되는 서비스가 국제기구 표준이나 기술기준에서 규정하고 있는 기술적 파라미터를 지키고 있으므로 문제가 없다고 한다. 이런 이견이 있는 것은 서비스를 아직 공급자 측면에서 바라보고 있으며, 통신의 발달로 인하여 이용자가 기대하는 서비스 수준은 높아졌으나 공급자는 아직도 서비스의 하드웨어에만 치중하고 있기 때문이다.

\* 正會員, 韓國電子通信研究所 無線通信標準研究室  
(Wireless Communication Standardization Section, Protocol Engineering Center, ETRI)

\*\* 正會員, 漢南大學校 應用統計學科  
(Dept. of applied statistics, Hannam University)  
接受日字: 1996年2月2日, 수정완료일: 1996年10月7日

통신 서비스 이용자가 평가하는 품질은 주관적이므로 객관적인 수치로 측정하는 것은 불가능하므로 설문 조사분석에 의해 측정될 수 있을 것이다. 조사분석에 의해 얻을 수 있는 정보는 사실, 의견, 행동으로 분류하고 있는데<sup>11)</sup>, 그 중 의견은 응답자의 선호도나 감정의 표현이다. 조사분석을 하는 이유는 다양하지만 어떤 특성의 분포를 발견하여 대상이 된 모집단에 대한 기술적인 논거를 주장하기 위하여 종종 시행된다<sup>12)</sup>. 조사분석에 의해 얻은 응답자의 의견은 기술적으로 측정 불가능하거나 어려운 경우 모집단의 파라미터(예를 들면 통신에 있어서 이용자가 평가하는 서비스품질)를 측정하는데 이용된다. 그러므로 설문을 통하여 서비스 이용자의 주관적인 평가를 조사한 후 그 평가를 수량화함으로써 이용자가 느끼는 서비스 품질을 측정할 수 있다.

MOS (Mean Opinion Score) 방법은 설문분석에 의하여 응답자의 주관적인 평가를 측정하는데 널리 사용되고 있다<sup>13)</sup>. 그러므로 통신에 있어서 이용자가 느끼는 서비스 품질을 측정하는데 MOS 방법을 사용할 수 있는데, 이 방법에 의해 서비스 품질을 올바르게 측정하기 위해서는 신뢰할 만한 설문지 작성, 모집단을 대표할 만한 응답자 선택을 위한 표본 설계, 의미있고 알맞은 통계기법의 적용이 매우 중요하다. 설문지 작성에 있어서 서비스의 품질을 제대로 측정할 수 있고 의미있는 항목설정이 가장 중요하므로 통신 전문가와 심리학자 그리고 통계 분석가의 공동연구는 필연적이다. 또한 설문 응답자를 선택하기 위한 모집단의 정의와 표본설계도 품질을 정확하게 측정하는데 중요한 역할을 한다. MOS 방법은 응답자의 주관적인 평가를 5점 척도를 사용하여 계량화한 후 산술 평균을 비교함으로써 두 집단을 비교하고 있으나 이런 단순비교에는 문제점이 있어 해결책으로 누적 퍼센트 곡선을 이용한 분석 방법이 제안되었다. 본 연구에서는 이용자가 평가하는 그룹3 팩시밀리 화상 품질을 측정하는데 있어서 설문지 작성 및 분석, MOS 방법과 누적 퍼센트 곡선의 적용을 살펴봄으로써 통신에 있어서 서비스 품질을 측정하는 방법을 제시하고자 한다.

## II. MOS 방법과 누적 퍼센트 곡선

MOS 방법은 평가자의 주관적인 평가를 5점 척도를 사용하여 계량화한다. 각 설문 문항에 따라 어휘 선택

의 차이가 있을 수 있으나 일반적으로 문항의 보기를 매우 좋다, 좋다, 보통이다, 나쁘다, 매우 나쁘다의 순으로 배열하여 응답자의 평가를 조사한 후 각각 5부터 1까지의 점수로 환산하여 그들의 주관적인 평가를 계량화시킨다. 우리나라에서는 응답자가 대체로 극단적인 평가는 (매우 나쁘다 혹은 매우 좋다) 회피하는 경향이 있어 4점 척도를 사용하기도 한다. 5 점 척도를 사용하는 경우 응답자의 평가점수의 평균과 분산은  $i$ 를 점수,  $P_i$ 를 응답자가  $i$ 번째 문항을 ( $i$ 점수) 선택한 확률이라고 하면 아래와 같이 정의된다.

$$\text{Mean} = \sum_{i=1}^5 i P_i$$

$$\text{Std} = \left( \sum_{i=1}^5 (i - \text{MOS})^2 P_i \right)^{1/2}$$

설문분석을 통해 주관적인 의견을 계량화시킴으로써 서비스 품질을 측정할 수 있으나, 산술평균만으로는 온전한 해석이 불가능하다. 예를 들어 서비스 품질을 측정하기 위하여 5 점 척도를 사용한 설문분석 결과 평가점수의 평균이 3.2였다고 하면, 단지 서비스 품질은 보통보다는 조금 높다고 말할 수 있을 뿐이다. 그러면 3.3인 경우에는 어떻게 해석할 것인가? 어떤 서비스를 A회사와 B회사가 제공하고 있고, 설문조사 결과 A회사를 이용하는 사람 중 설문에 응답한 사람이 평가한 점수의 평균을 3.7, B회사의 경우 응답자의 평가점수의 평균을 4.1이었다고 하자. B회사가 A회사보다 좋은 서비스를 제공할 수 있다고 할 수 있는가? B회사가 제공하는 서비스를 제공받고 있는 이용자가 평가를 후하게 하였다면 평균의 절대적인 비교에 의한 결론은 의미가 없을 것이다.

이런 문제점을 해결하기 위하여 누적 퍼센트 곡선에 의하여 평가점수의 평균을 해석하고 비교하는 방법이 제안되었다. 설문조사 결과 구해지는 점수에 대한 히스토그램을 정규분포에 근사시키고 평균과 표준편차를 조정하면 후 그것에 의해 누적 퍼센트 곡선을 도출할 수 있다.

## III. 설문분석

사용자가 그룹 3 팩시밀리를 사용하여 문서를 보내는 경우 보내고자 하는 문자나 그림 등의 화상이 팩시밀리에 입력되면 그 정보는 디지털 신호로 워치된다. 그러나 그룹 3 팩시밀리는 아날로그 공중전화망을 사

용하므로 전송시 모든 정보는 아날로그 신호로 변조한다. 팩시밀리 안에는 모뎀이 내장되어 있어 디지털 신호를 아날로그 신호로 변조된 후 송신하며, 수신된 신호는 다시 디지털 신호로 복조되어 수신된 화상이 감열방식에 의해 인쇄된다. 화상을 디지털 신호화하기 위해 주사선 밀도는 8 pel/mm를 사용하며 각 주사선은 하나의 프레임으로 인식되어 전송된다. 부주사선 밀도도 3.85로 사용하고 있지만 프레임에 포함된 정보의 내용에 따라 각 프레임은 서로 다른 수의 비트로 인식되어 전송된다. 주사선과 부주사선으로 등분된 한 펄(pel)은 하나의 정보로 인식되어 디지털 부호화 하지만 같은 정보가 반복되면 송신시간 절약을 위해 독자적 부호 방식을 가지므로 같은 길이의 프레임이라도 디지털 부호화한 후 비트의 개수는 상이하게 된다.

아날로그 전화와는 달리 팩시밀리에 있어서는 비트의 에러보다는 프레임에서의 에러 즉 FER(Frame Error Ratio)에 의해 서비스 품질을 측정한다. 전송된 프레임 내에서 하나의 비트에 에러가 생겨도 그 프레임은 에러를 가진 프레임으로 간주되지만 (기술적 측면에서 에러), 비트의 에러 발생이나 프레임 에러가 발생하더라도 사용자는 화상이나 문자를 판독하는 데는 문제가 없을 수 있다. 한 프레임 내의 다수의 비트 에러 발생 혹은 다수의 프레임 에러가 발생하면 판독자는 화상의 차이를 느끼게 된다. 현재 ITU-T 권고에 의하면 연속적으로 4개 이상 프레임 에러의 발생이 한 번 이상인 경우, 12개 이상의 프레임 에러가 발생한 경우, 혹은 연속적으로 2개나 3개의 프레임 에러들의 발생이 3번 이상인 경우 전송된 페이지는 판독이 어려운 화상으로 판정하고 있다<sup>[4]</sup>.

그러나 실제 전송된 화상에 있어서 FER을 측정하는데 어려움이 있으며 프레임 에러의 형태에 따라 화상 품질이 달라진다. 팩시밀리 화상 품질을 측정하기 위해 기술적으로 FER을 측정하기보다는 사용자의 주관적 평가 방법에 의해 팩시밀리가 제공하는 서비스 품질을 측정할 수 있다. 설문조사를 통하여 화상 품질을 측정하기 위하여 ITU-T No.2 시험 화상을 사용하여 실제 선로에서 화상을 송수신 한 후 응답자들이 수신된 시험 화상을 보고 설문지에 답함으로써 그들이 느끼는 품질을 조사할 수 있다.

설문지 작성을 위하여 시험 도표 No.2를 살펴보면 다음과 같다. 시험도표 No.2는 길이 297mm이고 폭 210mm로 문자의 판독성에 대한 일반적인 평가용으로

국제표준기구에 의해 권고되고 있으며, 화상 품질 측정을 위해 여러가지 사항들을 고려하여 작성되었는데 이것을 정리하면 표1과 같다<sup>[5]</sup>.

표 1. ITU-T 표준 시험 도표 No. 2의 측정 항목표

Table 1. List of Items in test chart No. 2 of ITU-T.

	도표 항목	측정되는 사항
1	화살표와 선의 4가지 그룹	프레임의 왜곡 정도
2	0.5mm의 교대되는 흑백선	최소 허용 해상도의 시험
3	흑·백색 대역	수평 해상도
4	대각선	시험 패턴 표면의 균일성
5	흑색원·동심원	송수신 장치간의 협동지수 차
6	수평대역	수직해상도
7	문자	문자 판독의 한계

설문지의 문항은  $\mu\text{m}$  단위로 응답되는 3번의 흑·백색 대역의 항목을 제외하고는 각 항목의 평가가 5점 척도로 작성될 수 있다. 실제 설문지에서는 설문 문항이 많았으나 “당신은 이 시험도표에 대한 종합적인 평가는 무엇입니까?”라는 설문문항 하나로 화상 품질을 측정할 수 있을 것이다<sup>[6]</sup>. 하나의 문항만으로 측정할 수 있으나 상관분석과 요인분석을 통하여 유의한 문항들을 이용하여 같은 방법으로 화상 품질을 측정할 수 있다. 여러 개의 문항을 고려하는 경우는 그 문항의 점수들을 적절한 방법에 의해 합치는 과정만 더 필요하게 된다. 본 연구에서는 총체적인 화상 품질을 묻는 하나의 문항 평가 점수를 이용하여 팩시밀리 화상 품질을 측정하였다.

송·수신간에 거쳐야 하는 교환기의 수가 많을수록, 통화량이 많을수록 BER이 커지므로 FER도 증가할 것이므로 화상 품질에 영향을 미치는 요인으로 링크 수와 통화량을 고려하였다. 통화량이 많은 시간대와 적은 시간대로 나누었는데, 일반적으로 통화량이 많은 시간은 오전 9시부터 오전 12시 그리고 오후 6시부터 오후 9시까지로 간주되므로 본 연구에서도 같은 방법으로 시간대를 나누고 링크 수는 1(대전-대전:국번이 같음), 2((대전-대전:국번이 다름), 3(대전-충남 조치원), 그리고 4(대전-서울)로 나누어 요인들의 8가지 다른 결합 조건에 따라 시험도표 No.2를 전송하여 응답자에 의해 평가될 화상을 얻었다. 이것을 응답자는 “당신은 이 시험도표에 대한 종합적인 평가는 무엇입니까?”라는 설문문항에 응답한 결과를 분석함으로써 팩시밀리 화상

품질을 측정하였다. 설문조사 방법은 대학생(충남대 심리학과학생)을 대상으로 65명의 응답자를 무작위 추출하였다. 그러나 조사 후 문제점으로는 실제 팩시밀리를 한 번도 사용하지 않은 응답자가 상당히 많았었다. 각 응답자들은 8가지 결합 조건에 의해 전송된 시험도표를 각 3장씩 총 48장을 각각 주관적인 성향에 의해 "매우 좋다, 좋다, 보통이다, 나쁘다, 매우 나쁘다" 중 하나로 평가하였다. 각 문항의 보기는 5점부터 1점까지 등간 점수로 계량화되었다.

고전적인 방법(MOS 방법)은 다음과 같이 분석한다. 각 문항 (매우 좋다, 좋다, 보통이다, 나쁘다, 매우 나쁘다)의 도수(frequency)를 구하여 팩시밀리 화상품질을 평가하거나 65명에 의해 평가(설문응답)된 48장에 대한 응답점수의 평균에 의해 평가한다. 본 설문분석 결과 총 평가점수 평균은 3.154였다. 그러므로 고전적인 방법에 의하면 팩시밀리 화상품질은 보통보다 약간 높다는 결론을 내리게 된다. 앞에서 언급하였듯이 응답자의 성향에 따라 절대적인 팩시밀리 화상품질과는 달리 평가점수들은 달라지게 되므로 3.154의 값은 의미를 갖지 못한다.

누적퍼센트곡선은 다음과 같이 구해진다. (1) 각 결합조건(본 연구에서는 링크 수와 통화량에 의해 이루어진 8가지 조건)에 따라 점수의 상대도수를 나타내는 히스토그램들을 구한다. (2) 각 히스토그램들은 다음 식에 의해 평균  $\mu$ 와 표준편차  $\sigma$ 를 갖는 정규분포로 근사된다.  $\Phi$ 는 정규분포의 누적확률분포함수이다.

$$\text{Mean} = 5 - \sum_{i=1}^4 \Phi\left(\frac{i+0.5-\mu}{\sigma}\right)$$

$$\text{Std} = \left(25 - \sum_{i=1}^4 (2i+1)\Phi\left(\frac{i+0.5-\mu}{\sigma}\right) - (\mu)^2\right)^{1/2}$$

(3) 근사된 정규분포들의 분산을 이용하여 다음 식에 의해 통합표준편차(overall variance) S를 구하고 위의 첫 번째 식에서  $\sigma$ 에 S를 대입하여 각 결합조건의 평균을 조정한다.  $g$ 는 결합조건(본 연구에서는 G는 8)을 나타내며,  $S_g$ 는  $g$  결합조건의 표준편차이고  $n_g$ 는 그 결합조건의 관측치 수이다.

$$S = \sum_{i=1}^G \left(\frac{(n_g-1) S_g^2}{N}\right) \text{ where } N = \sum_{g=1}^G (n_g-1)$$

(4) 각 결합조건의 평균평가점수들(x축)과 각 문항의 누적도수(cumulative frequency)에 의해 누적퍼센트곡선을 도출한다. 누적퍼센트곡선은 다항식에 의해 근

사된다. 본 연구에서는 최적다항식이 일차다항식이었으며 도출된 곡선들은 <그림 1>이다. 누적퍼센트곡선은 각 문항 (매우 좋다, 좋다, 보통이다, 나쁘다, 매우 나쁘다)에 대해 도출되므로 누적퍼센트곡선은 5개가 되어야 하지만 매우 나쁘다를 선택한 사람은 한 명도 없었으므로 실제 도출된 것은 4개이다.

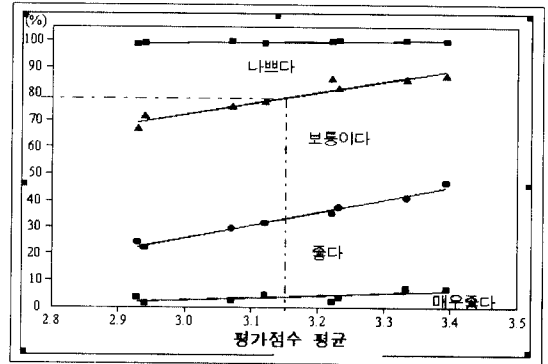


그림 1. 평가점수평균에 의한 누적퍼센트곡선  
Fig. 1. Cumulative percent curve at MOS.

설문조사 결과 응답자 65명에 의해 평가된 48장의 총 평가점수 평균은 3.154였다. 그러므로 그림1에 의해 응답자의 30%는 현재 제공되고 있는 팩시밀리 화상품질을 좋거나 매우 좋다고 생각하고 있으며, 80% 정도는 화상품질이 보통 이상이라고 평가하고 있다는 결론을 내릴 수 있다.

#### IV. 결론

통신에 있어서 이용자가 느끼는 서비스품질을 높은 수준까지 향상시켜야 통신시장을 개방해야 하는 우리로서는 국제경쟁력을 갖출 수 있을 것이다. 본 논문에서는 이용자가 느끼는 주관적인 평가를 설문 분석을 통하여 조사하고, MOS 방법에 의해 계량화하는 방법에 있어서 문제점을 언급하였다. 그 문제점을 해결하는 방안으로 누적퍼센트곡선을 제안하여 서비스품질을 측정하였다.

향후 연구에서는 주관적인 평가점수와 품질에 영향을 미친다고 생각되는 기술적 파라미터들과의 관계를 분석함으로써 이용자가 느끼는 서비스품질을 극대화하기 위하여 로짓모형을 이용한분석방법으로 서비스품질에 영향을 미치는 기술적 파라미터와 그들의 최적 수준을 제시할 것이다.

## 참 고 문 헌

- [1] F. C. Dane, "Research Methods", Thompson Information/Publishing Group, pp 128-140, 1960.
- [2] E. R. Babbie, "Survey Research Methods", Wadsworth Publishing Company Inc., pp 227-248, 1973.
- [3] J. Guilford, "Psychometric Method", 2nd edition, McGraw-Hill, 1954.
- [4] ITU-T, "Quality as corrupted by transmission-induced scan errors", Draft Rec., E.45y, 1993.
- [5] ITU-T, "Terminal Equipment and Protocols for Telematic Services", BlueBook, Vol.VII, T.21, pp 71-76, 1988.
- [6] 권세혁, 황진, "팩시밀리 화상품질 측정 연구", 전자통신동향분석, 제 8권 제4호, pp 157-162, 1994.1

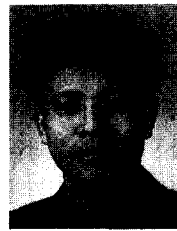
## 저 자 소 개



李永煥(正會員)

1962년 8월 5일생. 1984년 2월 숭실대학교 전자공학과 졸업(공학사). 1986년 2월 광운대학교 대학원 전자공학과 졸업(공학석사). 1989년 7월 ~ 현재 한국전자통신연구소 무선통신표준연구실 선임연구원.

주관심분야는 서비스품질, 이동통신, 정보통신 표준화 등임



權世赫(正會員)

1961년 1월 3일생. 1983년 2월 성균관대학교 통계학과 졸업(학사). 1985년 2월 성균관대학교 대학원 통계학과 졸업(석사). 1992년 12월 North Carolina State University(통계학 박사). 1993

년 5월 ~ 1995년 2월 한국전자통신연구소 선임연구원. 1995년 3월 ~ 현재 한남대학교 응용통계학과 조교수. 주관심분야는 서비스품질, 자료분석, 전산통계 등임