

AI기반 상수도시설 개량 의사결정 모델 분석

김기태¹, 민병원², 오용선^{2*}

¹목원대학교 정보통신융합공학부 박사과정, ²목원대학교 정보통신공학과 교수

Model Analysis of AI-Based Water Pipeline Improved Decision

Gi-Tae Kim¹, Byung-Won Min², Yong-Sun Oh^{2*}

¹Ph.D. Student, Department of Information Communication Engineering, Mokwon University

²Professor, Department of Information Communication Engineering, Mokwon University

요약 상수도분야 인공지능 기술개발 관심도가 증가함에 따라 상수도 관로에 대해서 노후관 상태평가 데이터 결과를 활용하여 반복적인 학습으로 개량 의사결정 등급을 예측할 수 있는 인공지능망 알고리즘을 개발하고 검증과정을 통하여 가장 신뢰성 있는 예측 모델을 제시하고자 한다. 2020년 한강유역의 노후관로 정비 기본계획에 의한 간접평가 데이터 12개 항목을 기반으로 데이터 전처리 하고 인공지능망 알고리즘을 적용하여 반복학습과 검증을 통해 계산된 결과값과 직접평가 결과값의 일치율이 90% 이상이 되도록 역전파 과정을 통해 가중치를 업데이트 하면서 최적화하여 관로 등급을 예측하는 알고리즘을 개발하였다. 알고리즘 정확도 검증결과 모든 관종 데이터가 고르게 분포되어 있고 학습 데이터가 많아야 예측평가 정확도가 높아지는 것을 확인할 수 있었다. 향후 전국의 다양한 데이터가 확보되면 인공지능망을 이용한 관로등급 예측의 신뢰도가 좀 더 향상되어 객관화된 노후관 상태평가 의사결정 지원 역할을 수행할 수 있을 것으로 기대된다.

주제어 : 인공지능, 데이터 분석, 노후관 상태평가, 알고리즘 모델, 간접평가, 직접평가

Abstract As an interest in the development of artificial intelligence(AI) technology in the water supply sector increases, we have developed an AI algorithm that can predict improvement decision-making ratings through repetitive learning using the data of pipe condition evaluation results, and present the most reliable prediction model through a verification process. We have developed the algorithm that can predict pipe ratings by pre-processing 12 indirect evaluation items based on the 2020 Han River Basin's basic plan and applying the AI algorithm to update weighting factors through backpropagation. This method ensured that the concordance rate between the direct evaluation result value and the calculated result value through repetitive learning and verification was more than 90%. As a result of the algorithm accuracy verification process, it was confirmed that all water pipe type data were evenly distributed, and the more learning data, the higher prediction accuracy. If data from all across the country is collected, the reliability of the prediction technique for pipe ratings using AI algorithm will be improved, and therefore, it is expected that the AI algorithm will play a role in supporting decision-making in the objective evaluation of the 한 condition of aging pipes.

Key Words : Artificial Intelligence Technology, Data Analysis, Aging Pipe Condition Evaluation, Algorithm Model, Indirect Evaluation, Direct Evaluation

*교신저자 : 오용선(ysunoh@mokwon.ac.kr)

접수일 2022년 7월 10일

수정일 2022년 8월 28일

심사완료일 2022년 8월 30일

1. 서론

1.1 연구의 배경 및 목적

우리나라의 상수도는 1908년 근대식 정수장인 독도 정수장이 완공된 이후 100년이 지난 지금까지 양적 및 질적으로 많은 발전을 이루고 있으며 최근 급속도로 발전하고 있는 정보통신기술을 접목함으로써 시설의 안정성을 더욱 제고시키고 있다.[1]

정수장에서 생산된 수돗물을 공급하는 과정에서 노후된 관로로 인한 수질 및 누수사고 등이 빈번하게 발생하고 있으므로 안정적 수돗물 공급에 대한 요구가 지속적으로 증대되고 있어, 이를 예방하기 위해서는 신뢰성 있는 관로 상태평가와 이에 따른 개량 의사결정이 매우 중요하다.

국내에서 노후관에 대한 상태평가 방식에는 크게 간접평가와 직접평가가 있으며 간접평가는 예측수단으로 기존 관 제원, 운영 매설환경 인자를 중심으로 대개 점수평가 방식으로 노후수준은 추정할 수 있으나 개량방식을 결정하는 데에는 한계가 있어 직접평가가 필수적이다.[2]

직접평가는 굴착에 의한 직접 조사방식으로 많은 비용이 발생하고 단수, 교통통제, 현장 작업인력 필요 등 작업사항에 대한 애로사항이 발생하여 샘플링을 통한 등급 조사를 시행하고 있다.

따라서 기존 방안에 대한 한계를 극복하고 비용을 최소화하는 다양한 연구가 필요한 실정이다.

1.2 연구의 내용 및 범위

본 연구에서는 인공지능에 대한 기술개발 관심도가 증가함에 따라 노후 상수관로 정비 기본계획으로 도출된 상태평가 데이터 결과를 활용하여 반복적인 학습으로 개량 의사결정 등급을 예측할 수 있는 인공지능망 알고리즘 모델을 개발하고 다양한 검증과정을 통해 신뢰성 있는 예측모델을 제시하고자 한다.

이를 위해 2020년 한강유역 7개 지자체의 노후관로 간접 상태평가 데이터 및 직접 상태평가 데이터를 대상으로 연구를 진행하였다.

본 연구의 예측모델 개발 절차는 Fig. 1과 같다. 상수관로의 간접평가 및 직접평가 데이터를 수집하고 수집된 데이터를 전처리하여 반복 학습을 시행하고 정확도 높은 인공지능망 알고리즘 모델을 구축한다.



[Fig. 1] Prediction model development procedure

개발된 인공지능망 알고리즘 모델을 4가지 방법으로 데이터를 적용하여 반복 학습하고 예측 결과(정확도)를 검증하여 가장 신뢰성 있는 개량 의사결정 알고리즘 모델을 제시하고자 한다.

- A도 상태평가 데이터 학습 모델을 통한 A도 평가
- B도 상태평가 데이터 학습 모델을 통한 B도 평가
- 통합(A+B도) 데이터 학습 모델을 통한 통합 평가
- 관종별 데이터 학습 모델을 활용한 관종별 평가

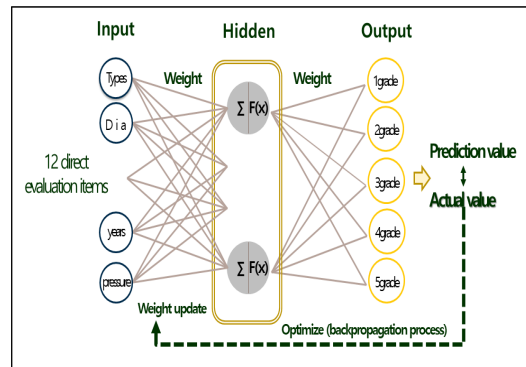
2. 인공지능망 알고리즘 개발

2.1 인공지능망 알고리즘 모델 개요

인공지능망은 퍼셉트론(인간의 신경계 모방)을 여러 계층으로 조합한 분석모델로 입력과 출력의 관계를 잘 설명할 수 있는 복잡한 함수식의 가중치를 찾는 과정이다.[3][4]

본 연구에서는 관로데이터 입력변수의 선형 합을 활성화 함수(ReLU, Softmax 등)에 넣고 계산된 결과값과 실제값(직접평가 결과값)의 일치율이 90% 이상이 되도록 반복 학습을 통해 최적화하였으며 이때 역전파 과정을 통해 가중치를 업데이트 하면서 오차를 줄이고 최종적으로 관로 등급을 예측한다.[5][6]

본 연구의 예측 인공지능망 구조는 Fig. 2과 같다.



[Fig. 2] Prediction model artificial neural network structure

인공지능망은 입력값이 들어가는 입력층, 중간층, 그리고 출력값이 나오는 출력층, 입력층과 중간층, 중간층과 출력층 사이를 연결하는 각각의 가중치로 구성되어 있다.[7][8]

상수관로 정밀조사 매뉴얼(안(2020, 환경부)에 따라

간접평가 항목 12개를 Input(입력층)으로 설정하고 직접 평가에 의한 개량등급은 5개의 등급으로 결과를 도출함으로써 Output(출력층)은 5개 등급으로 설정한다.

은닉층과 은닉노드는 많을수록 무조건 예측력이 좋아지는 것은 아니므로, 분석을 통해서 최적의 설정값을 찾아내어 정확성을 판단해야 한다. 따라서 다수의 테스트 중에서 모형의 형태와 정확성을 확인하며 테스트를 진행한 결과 은닉층 한 개에 은닉노드를 100개로 설정하고 학습과 검증을 진행하였다.

기존의 노후관로 상태평가의 간접평가를 적용하여 평가하는 방식과 개발한 인공지능망 알고리즘 모델에 의한 결과로 평가하는 방식을 비교하면 Table 1과 같이 나타낼 수 있다.[9]

[Table 1] Comparison of traditional and AI methods

| | Existing Indirect Evaluation | AI Evaluation |
|-------------------|-------------------------------|---|
| Weighted value | Score evaluation table | Initial weight(He) |
| Score calculation | Score calculation method | Update weighting factors through backpropagation |
| Learning | Update Score evaluation table | Backpropagation algorithm (Update initial weight) |
| Verification | Status evaluation grade table | Predict ratings by reducing error values |

2.2 전처리 및 예측등급 산정 방법

인공지능망 알고리즘 모델을 개발하기 위해 한강유역 7개 지자체의 노후관로 간접 상태평가 데이터 및 직접 상태평가 데이터 수집하여 데이터 전처리 작업을 먼저 시행하였으며 Table 2와 같이 관로 등급예측에 필요한 간접평가 12개 항목의 데이터를 추출, 재조합, 입력값 형식 표준화로 데이터 변형을 수행한다.

[Table 2] Direct evaluation items

| | items |
|-------------------------|--|
| Pipeline data | Types, Diameter, years |
| Corrosion environment | Soil Types, Water corrosiveness, Soil resistivity, CP system |
| Inside and outside load | Maximum pressure, Depth of burying, Road Types |
| Accident history | Number of breakages |
| Joint quantity | Number of breakages |

관종의 경우 CIP, DCIP, SP 이외의 관종은 비금속관으로 관종을 구분하고 데이터 이상치 및 누락에 대해서

는 제거하였다.[10]

이상치 제거 방법으로는 IQR(Inter Quantile Range) 방식을 이용하였고 결과년도 120년 및 매설깊이 0m인 데이터는 이상치 값이라 확인되어 제거하였다.[11][12]

인공신경망(ANN)의 입력값은 숫자 데이터만 가능하므로 Table 3과 같이 문자데이터를 모두 숫자 데이터로 변경하고 문자데이터의 공백 제거와 지역마다 의미가 같은 데이터 값은 하나로 통일시킨다.

[Table 3] Pipeline data standardization

| Pipe types | CIP | DCIP | SP | Non-metallic |
|------------|-----------------|---------------|----------|-----------------|
| Conversion | 0 | 1 | 2 | 3 |
| Road types | Industrial road | Walking road | One-lane | Two-lane |
| Conversion | 0 | 1 | 2 | 3 |
| Soil Types | Sand | Sand and soil | Gravel | Gravel and soil |
| Conversion | 0 | 1 | 2 | 3 |

Table 4는 인공지능망 알고리즘에 입력되는 12개 관로 데이터의 실제 입력값을 일부 표시하였다.[13]

[Table 4] Data actual inputs

| Types | Diameter | years | Soil Types | Water corrosiveness | Soil resistivity |
|-------|----------|-------|------------|---------------------|------------------|
| DCIP | 100 | 24 | Sand | -1.1 | 11638 |
| DCIP | 200 | 21 | Sand | -1.1 | 11638 |
| DCIP | 100 | 21 | Sand | -1.1 | 11638 |
| DCIP | 200 | 21 | Sand | -1.1 | 11638 |
| DCIP | 100 | 24 | Sand | -1.1 | 11638 |
| DCIP | 200 | 22 | Sand | -1.1 | 11638 |

관로의 개량 의사결정을 위한 등급평가는 5등급으로 분류하며 1등급에서 5등급까지 등급 신뢰도를 계산하여 하나의 배열에 나열하고 배열 안에 신뢰도 중에서 가장 높은 값을 찾아 최종 예측등급 결과값으로 선정한다.

3. 인공지능망 알고리즘 모델 적용

3.1 알고리즘 모델 학습 검증

본 연구에서는 2020년도 노후 상수관로 정비비용액 A

도 3개 시와 B도 4개 시·군의 도수·송수·배수관로 간접 평가 및 직접평가 데이터를 이용하였으며 사용된 데이터 현황은 Table 5와 같다

[Table 5] Indirect and direct status evaluation data

| | | | |
|------------|------|------------|------------|
| | Sum | Province A | Province B |
| Data lines | 1626 | 736 | 890 |

관종별 데이터는 Table 6와 같이 비금속 관종(COP, FP, HI-3P, HI-VP, PE, PE-3, PIP, PVC, SPOL-PIPE 등)이 가장 많고 SP, DCIP 순으로 사용하고 있다.

[Table 6] Pipeline data

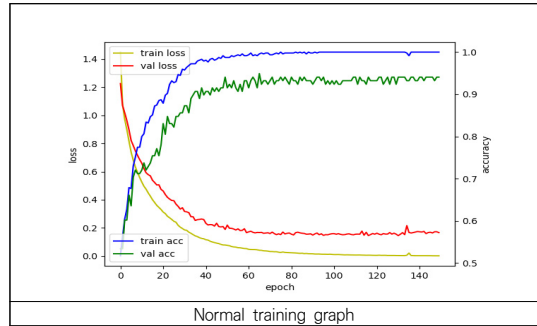
| Types | DCIP | SP | Non-metallic | etc |
|------------|------|-----|--------------|-----|
| Province A | 357 | 282 | 69 | 28 |
| Province B | 11 | 106 | 759 | 14 |
| Sum | 368 | 388 | 828 | 42 |

데이터 취득형식은 엑셀파일 형태로 라인수(Row) 당 한 개 관로 평가데이터를 적용하고 Scikit-learn 라이브러리를 사용하여 데이터가 고르게 분포되어 정상적인 모형이 형성되도록 랜덤으로 데이터를 섞어주었으며 전체 데이터를 인공신경망 알고리즘 모델의 학습과 검증 그리고 평가를 위해서 Table 7과 같이 각각 60%(학습):20%(검증):20%(평가)의 비율로 분할을 진행하였다.

[Table 7] Split data ratio

| | Ratio | Purpose |
|------------|-------|--------------------|
| Learning | 60% | Model Learning |
| Validation | 20% | Model Check |
| Test | 20% | Model Verification |

학습을 위해 60%의 데이터를 적용하여 반복 학습을 수행하였고 데이터의 모형 검증은 Fig. 3와 같이 Val Loss(검증데이터 손실값)가 점차적으로 감소하고 Val Acc(검증데이터 정확도)가 점차 증가하는 그래프를 확인하며 알고리즘 모델로 예측한 결과값과 실제값(직접평가 결과값)의 일치율이 90% 이상이 되도록 반복 학습훈련을 진행하였다.[14][15]



[Fig. 3] Training graphs for model verification

3.2 알고리즘 모델 평가

한강유역 7개 지자체의 상수관로 상태평가 자료를 적용하여 Table 8과 같이 4가지 방법으로 반복 학습과 검증을 완료한 인공신경망 알고리즘 모델에 단계별로 평가용 데이터(20%)를 적용하여 예측 결과(정확도)를 검증하고자 하였다.

[Table 8] Configuring data set of model

| Modeling | Data | Data set structure |
|----------|--------------|--|
| 1 | Province A | Learning(60%) + Validation(20%) + Test(20%) |
| 2 | Province B | Learning(60%) + Validation(20%) + Test(20%) |
| 3 | Province A+B | Learning(60%) + Validation(20%) + Test(20%) |
| 4 | Province A+B | Split of data by type of pipe Learning(60%) + Validation(20%) + Test(20%) |

3.2.1 A도 및 B도 데이터 학습 모델을 통한 A도, B도 평가

A도 3개 시의 상수관로 상태평가 데이터를 반복 학습과 검증 후 평가데이터를 가지고 정확도 예측결과 Table 9과 같이 94.9%의 높은 정확도가 나타났으며 B도는 모델 예측결과 81.9%의 정확도가 나타났다.

[Table 9] Accuracy prediction result

| | Province A | Province B |
|----------|---------------------|---------------------|
| Accuracy | 129ea/136ea (94.9%) | 145ea/177ea (81.9%) |

정확도 면에서는 A도 보다 B도가 다소 낮게 확인되었지만 이는 B도의 데이터가 A도 보다 지역적으로 넓고 관종의 분포가 고르지 못하고 기존의 노후도 상태평가에

취약한 비금속 관종이 대부분을 차지하고 있어 이로 인한 학습 부족으로 판단된다.

그러나 기존의 노후도 상태평가 방법으로 시행한 간접평가 결과와 직접평가 결과의 일치도 보다 높은 정확도를 확인할 수 있었다.

3.2.2 통합(A+B도) 데이터 학습 모델을 통한 평가

A도와 B도를 고루 섞이게 한 후 상태평가 데이터를 반복 학습과 검증 후 평가데이터(20%)를 가지고 정확도 예측결과 직접평가 등급과의 일치도가 86%의 정확도를 확인하였다.

이는 A도만 적용한 모델보다 A+B도 통합 모델이 낮게 정확도가 나왔지만 B도만 적용한 모델보다는 높게 정확도가 나타난 결과로 보아 A도가 B도의 자료보다 좀 더 폭넓은 범위의 데이터가 고르게 분포되어 있음을 확인하였으며 이를 통해 좀 더 폭넓은 데이터의 자료와 지속적인 데이터의 반복 학습을 통해 인공지능망 알고리즘 모델을 구축한다면 좀 더 신뢰성이 높은 모델이 구축됨을 확인할 수 있었다.

3.2.3 관종별 데이터 학습 모델을 통한 평가

관종별 등급을 예측하는 모델을 구축하기 위해 A도와 B도를 데이터 통합 후 관종별로 나누는 전처리 작업을 진행하였다.

CIP관 경우 노후도평가에 상관없이 전면 교체대상으로 평가 결과값 5등급을 일괄적용하여 제외하였다.

A도와 B도 관종별 상태평가 데이터를 고루 섞이게 한 후 반복 학습과 검증 후 평가데이터(20%)를 가지고 관종별 예측 정확도 분석결과 Table 10와 같이 DCIP관, SP관, 비금속관 순서로 정확도가 높았다.

[Table 10] Accuracy prediction result

| | DCIP | SP | Non-metallic |
|----------|------|-----|--------------|
| Accuracy | 96% | 94% | 80% |

관종별 정확도가 차이 나는 이유를 분석한 결과 입력 데이터 분포가 DCIP관은 1~5등급이 고루 분포가 되어 정확도가 높은 것으로 판단되었으며 비금속관의 경우 입력데이터 등급 분포가 다른관에 비해 좋지 않고 비금속관의 관종이 다양하게 섞여 있어 예측 정확도가 가장 낮은 것으로 판단된다.

4. 결론

본 연구에서는 2020년 노후관로 정비 기본계획에 의한 직접 및 간접평가 데이터를 활용하여 지자체 관로의 노후도 평가등급을 예측하고 직접 평가할 지역을 선정 및 의견을 제시하기 위한 보조 수단으로 사용하고자 노후도 상태평가 의사결정 인공지능망 알고리즘 모델을 개발하고 적용하였다.

학습 및 검증 방법으로 각 지자체의 자료를 활용하여 그 지자체에 적용시 결과예측, 지자체 데이터 통합 후 학습과 결과예측, 관종별 지자체 데이터 통합 후 학습과 결과예측의 순서로 인공지능망 알고리즘 모델의 신뢰도 평가 연구를 진행하였다.

본 연구 결과분석을 통해 A도와 B도 비교시 관종이 고루 분포된 A도가 정확도가 높았는데 이는 B도의 데이터가 A도 보다 지역적으로 넓고 기존의 노후도 상태평가에 취약한 비금속 관종이 대부분을 차지하고 있어 이로 인한 학습 부족으로 판단된다.

관종별 비교시에서도 노후관 개량등급 자료가 5개 등급으로 골고루 분포되어 있는 DCIP관이나 SP관이 정확도가 높게 나왔다.

따라서 전체적으로 데이터가 골고루 분포되고 데이터 양이 많을 때 정확도가 높은 것으로 평가되므로 향후 전국의 충분한 데이터가 확보되고 지속적으로 반복 학습을 시행한다면 신뢰성 있는 노후도 평가데이터를 통한 개량 의사결정 인공지능망 알고리즘 모델을 구축할 수 있을 것으로 기대된다.

REFERENCES

- [1] "Waterworks Standards", Korea Water and Wastewater Association, p2, 2010.
- [2] S M Lee and D S Kang, "Water pipe deterioration assessment using ANN-Clustering", Journal of the Korean Water Resources Association, pp959-960, 2018
- [3] Y R Choi and K H Kim, "Artificial Intelligence Overview and Application Cases", ie Magazine , p23, 2016.
- [4] S B Cho and J H Kim, "Computational Complexity of Artificial Neural Networks", Korean Institute of Information Scientists and Engineers, pp.315-318, 1989.
- [5] Y I Cho1, M J Oh, Y S Seok, S J Lee and M I Roh, "Resistance Estimation of a Ship in the Initial Hull

Design Using Deep Learning”, Korean Journal of Computational Design and Engineering, pp.203-210, 2019.

- [6] Cybenko, George. "Approximation by Superpositions of a Sigmoidal Function." Mathematics of Control, Signals, and Systems, pp.303-314, 1989.
- [7] S R Jo, H N Sung and B H Ahn, "A Comparative Study on the Performance of SVM and an Artificial Neural Network in Intrusion Detection", Journal of the Korea Academia-Industrial cooperation Society, pp.703-711, 2016.
- [8] C W Lee, S J Hong, J S Park and D G Yoo, "Assessment of Water Supply Pipes Deterioration Based on Deep Learning", Journal of the Korean Society of Civil Engineers, pp.62-63, 2020.
- [9] "Water Pipeline Precision Investigation Manual(Draft)", Ministry of Environment, pp.10-24, 2020.
- [10] K H Han, Y J Ryu, T S Kim and J H Heo, "Input Variables Selection of Artificial Neural Network Using Mutual Information", Korea Water Resources Association, pp.85-87, 2010.
- [11] H K Park, J S Lee, W H Song and S S Park, "Technology Forecasting using Patent Outlier Remova", Journal of the Korean Institute of Intelligent Systems, pp.177-184, 2018.
- [12] Y M Go, S W Go, "Alleviation of Vanishing Gradient Problem Using Parametric Activation Functions" Journal of the Korea Information Processing Society, pp.407-420, 2021.
- [13] "A Report on the Schematic Plan for the Improvement of the Aged Water Pipeline", Korea Water Resources Corporation, 2021.
- [14] J Y Seo, "Deep Learning TensorFlow Textbook", Gilbut, p.25, 2021.
- [15] S M An, "Deep learning architectures and applications" Journal of the Korea Intelligent Information Systems Society, pp.127-142, 2016.

김 기 태(Gi-Tae Kim)

[정회원]



- 1994년 2월 : 울산대학교 전기공학(공학사)
- 2007년 2월 : 충남대학교 전기공학(공학석사)
- 1994년 4월 ~ K-water 근무 중
- 2020년 3월 ~ 현재 : 목원대학교 IT공학과(박사과정), 한국콘텐츠학회 정회원.

<관심분야>

데이터 분석, 인공지능, 자동제어

민 병 원(Byung-Won Min)

[정회원]



- 2005년 2월 : 중앙대학교 컴퓨터 소프트웨어학과 (공학석사)
- 2010년 2월 : 목원대학교 대학원 정보통신공학과(공학박사)
- 2011년 3월 ~ 현재 : 목원대학교 정보통신공학과 교수, 한국콘텐츠학회, 한국통신학회, 대한전자공학회, 한국정보과학회 정회원.

<관심분야>

데이터베이스, 인공지능, 빅데이터 등

오 용 선(Yong-Sun Oh)

[정회원]



- 1983년 2월 : 연세대학교 공과대학 전자공학과(공학사)
- 1985년 2월 : 연세대학교 대학원 전자공학과(공학석사)
- 1992년 2월 : 연세대학교 대학원 전자공학과(공학박사)
- 1984년 3월 ~ 1986년 7월 : 삼성전자(주) 시스템개발실 연구원
- 1987년 1월 ~ 1988년 2월 : 3J TECH. INC. 선임연구원
- 1998년 9월 ~ 1999년 8월 : 한국해양대학교 객원교수
- 1988년 3월 ~ 현재 : 목원대학교 정보통신학과 교수, 한국해양정보통신학회, 한국통신학회, 대한전자공학회 IEEE 정회원.

<관심분야>

디지털 커뮤니케이션 시스템, 정보이론, 멀티미디어콘텐츠