

情報檢索시스템의 經濟評價

C. Warren Axelrod* 著
金 殷 植** 譯

譯者註: 이 글은 情報檢索시스템의 購入 如否의 意思決定을 위한 典型的인 標準節次를 提示한 것으로 C. Warren Axelrod, "The economic evaluation of information storage and retrieval systems", Information Processing & Management, v. 13, pp. 117~124, 1977을 옮겨 쓴 것이다.

1. 評價모델

一般的으로 情報檢索시스템의 評價모델은, "費用-時間-容量" 모델과 "性能" 모델의 2가지 카테고리로 나누어 생각할 수 있으며, ① 蓄積對象情報의 코우딩 및 索引 ② 質問內容의 코우딩 및 索引(①과 同一한 言語使用) ③ 質問과 蓄積情報의 對照·檢索 ④ 檢索情報의 評價 등의 4段階를 基本으로 한다.

1.1 費用-時間-容量 모델

이 모델에서는 ① 시스템의 活用量(容量), ② 月間 또는 單位檢索當의 料金(費用), ③ 質問處理에 所要되는 時間(時間) 등이 評價의 要素가 된다.

評價에 있어서 利用者와 購入者의 시스템性能에 대한 要求는 충분히 만족되며 시스템運用上의 모든 經費도 주어졌다고 가정한다. 시스템의 設置 및 稼動費用은 파일의 容量, 파일更新率, 하드웨어의 維持管理, 소프트웨어의 修正 및 活

用 등을 指標로 해서 算出하며, 블록線圖를 作成하여 시스템의 基本要素들을 나타내고 그 相互關係를 表示한다. 또 人力과 設備 등에 대한 所 要豫算과 그 內譯을 明示한다. 시스템의 壽命, 시스템과 파일의 更新 및 修正, 一般運營, 管理, 活用, 評價計算에 算入할 收益率, 시스템內의 여러 變化에 대한 利用者의 反應, 活用 또는 費用 및 時間에 대한 시스템의 依存度 등과 관련된 여러가지 가설도 설정한다. 費用은, 固定費用(間 接費), 準固定費用, 可變費用(直接費)으로 나누며 固定費用은 基本設備의 購入費, 研究開發費 및 기타 設置費를, 準固定費用은 研究開發의 持續과 같이 시스템의 規模나 活用과는 直接的인 關係없이 周期的으로 所要되는 費用, 可變費用은 시스템의 規模 活用 및 容量에 直接 關係되는 費用 및 파일의 更新과 管理에 所要되는 費用 등을 포함하는 것이다. 費用모델의 作成은 基準時點에 있어서의 初期現金支出과 시스템存續期間中의 部分的인 定規期間에 있어서의 負의 現金흐름(支出)을 가지고 한다. 이 部分的인 區間이 不定規的인 경우에는 全不定規期間의 "最小共通因子"를 取하여 이를 定規期間으로 看做하고 또 初期에 포함되지 않았던 기간중의 現金흐름은 零(Zero)으로 잡는다. 年間 基準의 資本費用(收益率)을 알면 等價期間率을 알 수 있으며 이것과 初期支出, 現金흐름, 現金出入間의 간격 및 시스템 壽命 등의 데이터로써 시스템의 "現在價值"를 算出할 수 있다(附錄 1參照). 數種의 購入對象시스템에 대하여 이러한 計算을 반복하여 "初期費用+割引費用"이 가장 적은 것을 採擇한다.

* Securities Industry Automation Corporation.

** KORSTIC 第 1 技術情報部 次長(技術士)

表 1. 2 시스템의 費用 比較

期間 (年)	0	1	2	3	4	5	合計
現金支出 (\$'000's)							
시스템 A	200	20	20	20	20	20	300
시스템 B	100	40	40	40	40	40	300
10% 割引要素	1.00	0.91	0.83	0.75	0.68	0.62	
割引支出							
시스템 A	200	18.2	16.5	15.0	13.7	12.4	275.8
시스템 B	100	36.4	33.0	30.0	27.4	24.8	251.6

1.2 實施例

有用期間 5年, 이 期間동안 시스템의 規模가 一定한 A, B 2시스템이 있다. 費用인플레이션은 없고, 시스템費用은 一定限度內에서의 活用에는 관계가 없으며 2시스템 모두 만족할만한 수준의 出力을 얻는다고 하자.

A시스템은 初期投資 200,000달러와 年間支出 (每年末에 支出) 20,000달러, B시스템은 100,000달러와 40,000달러를 각각 필요로 한다. 表 1에서와 같이 A, B시스템 모두 有用期間동안 300,000달러의 現金支出이 필요할때 購入者가 300,000달러線을 받아들인다면 2 시스템은 採擇의 對象이 될 수 있다.

그러나 附錄 1의 現在 價値法을 적용하여 10%의 收益率을 計上한다면 周期的 現金흐름은 割引計算되어야 한다. 따라서 表 1에서와 같이 A시스템은 B시스템보다 費用支出이 약 24,000달러가 더 많으며 다른 比較要因이 없는 한 B시스템이 採擇되어야 할 것이다.

1.3 性能모델

이 모델에 있어서는 利用者, 利用者와 購入者 및 購入者가 情報檢索시스템에 대하여 바라는 性能要件을 評價의 對象으로 한다. Webb⁸⁾는 費用, 時間(서어비스, 滯貨: Backlog, Idle), 信賴度, 人力, 面積, 擴張機能 및 시스템간의 경쟁력 등 주로 管理者가 考慮해야 할 시스템의 效用面에 대하여 언급하고 있다. 그러나 利用者의 입장에서는 서어비스的인 면 즉, 대기시간, 再

現度, 適合度, 信賴度 등을 重視한다. 결국 性能要件을 충분히 分析하고 SMART⁴⁾와 같은 시뮬레이션 시스템을 利用하여 그러한 要件들에 가장 적합한 시스템을 設計하는 것이 바람직하다. 다음에 이 시스템에 任意의 質問을 주어 시스템의 性能을 實証한다. 그리고 그 性能特性들을 導出하여 性能要件들과 對照한다.

性能要件이 시뮬레이션 시스템에 의하여 만족되면 서어비스의 類型을 결정하는 하드웨어 및 소프트웨어의 要件을 導出하고 이 시뮬레이션을 實稼動시스템으로 轉換시킨다. 이렇게 해서 시스템의 基本構造가 결정되면 費用모델을 公式化하여 基本費用 및 시스템固有因子에 의존하는 事項들을 具體化한다.

2. 利用者-購入者-모델 相互作用

그림 1의 블록線圖는 性能시뮬레이션과 費用모델간의 관계로서 利用者集團, 購入者 및 시스템모델사이의 相互作用을 나타낸 것이다.

먼저 設問을 통하여 利用者集團으로부터 바라는 性能要件을 抽出한다. 이때 利用者가 購入決

情報檢索시스템의 經濟評價

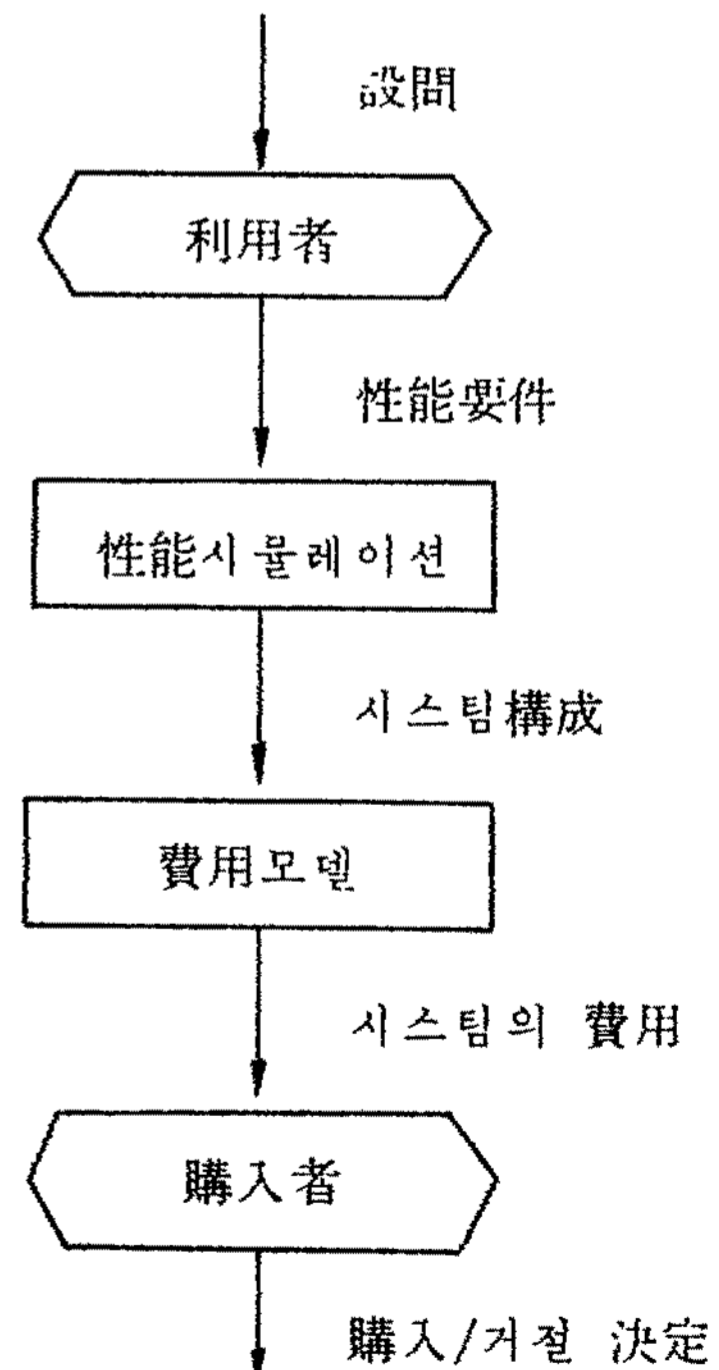


그림 1. 利用者, 性能모델, 費用모델 및 購入者간의 關係圖

定까지 調整하게 되면 (Webb⁸⁾의 경우) 모델은 한층 더 복잡해진다. 利用者が 바라는 性能要件은 시스템의 最適시물레이션 모델 開發에 應用한다. 시스템의 構成은 시물레이션으로부터 誘導하며 이를 파일의 크기, 入力과 出力의 形態 및 設備의 性能 등 購入者에게 提供할 實稼動시스템의 基本파라미터에 適用시킨다.

시스템構成이 具體化되면 各 項目에 대한 費用이 결정되어 費用모델이 作成된다. 이 費用모델에 의하여 시스템의 總費用이 결정되며 이것을 土台로 하여 購入決定을 하게 된다.

3. 피이드백과의 相互作用

여기서 修正提案할 것은, 시스템의 設計 및 購入決定에 있어서 利用者나 購入者の 機能은 그들사이의 情報피이드백을 手段으로 하여 費用모델과 性能시물레이션과의 관계가 확장될 수 있다는 것이다. 그림 2는 이 修正시스템으로서, 實線은 기본시스템을 나타낸 것이고 破線은 피이드백 링크를 追加한 것이다.

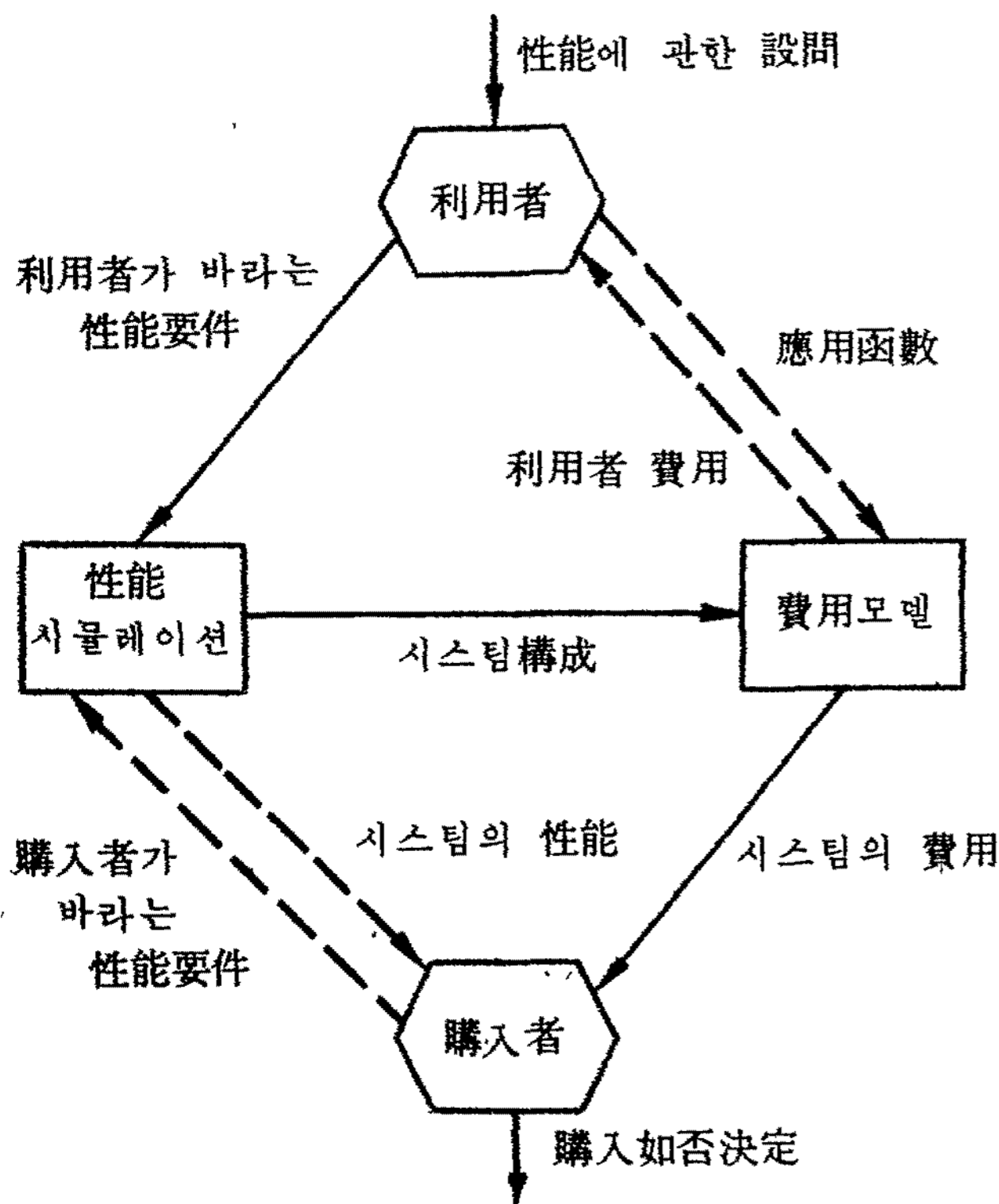
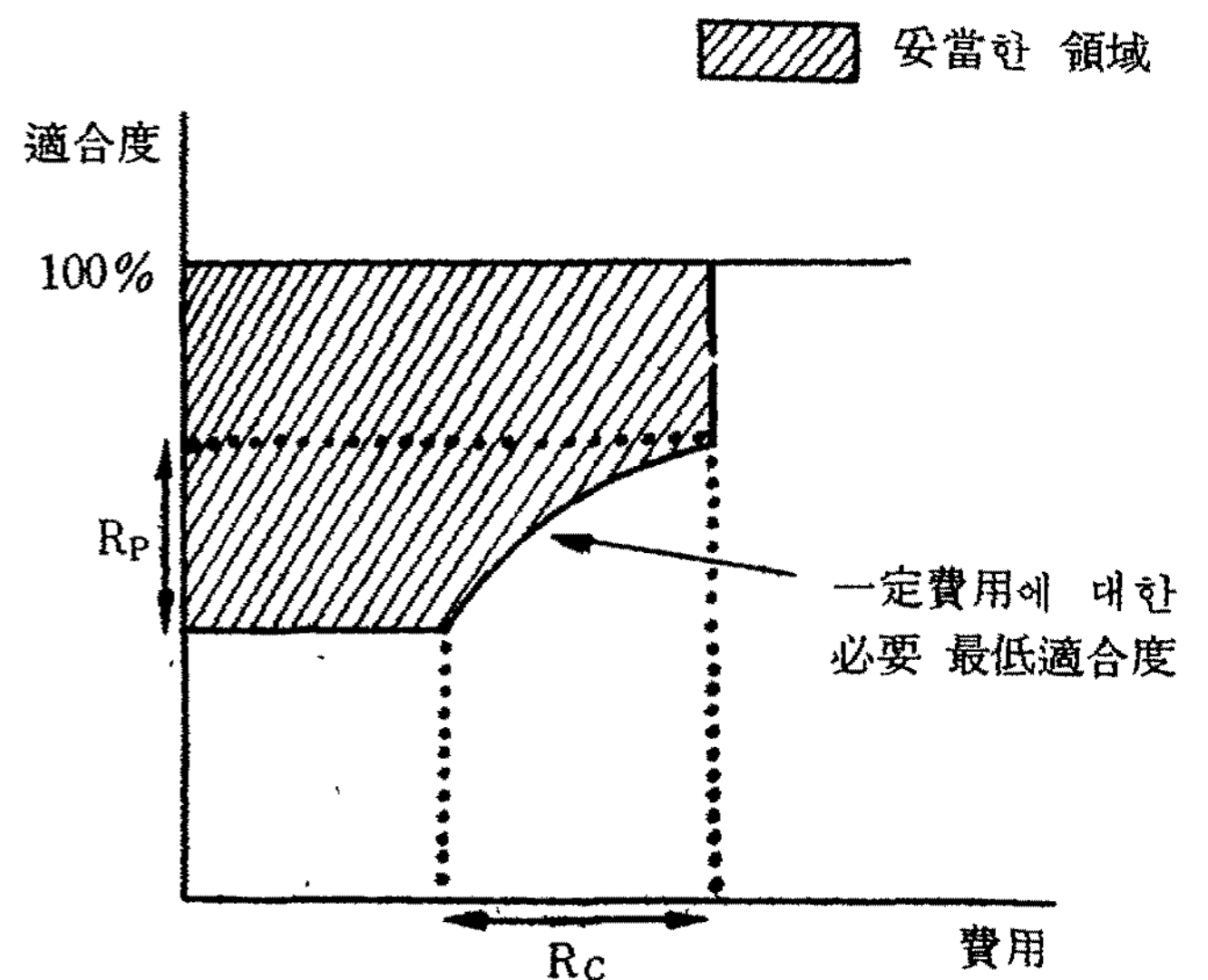


그림 2. 利用者と 費用모델間 및 購入者と 性能시물레이션間의 情報交換을 考慮한 利用者, 費用모델, 性能모델 및 購入者間의 關係圖

利用者は 시스템의 性能測定에 있어서 性能-費用間의 效用函數를 提示해 준다. 예를 들면, 利用者は 適合도에 대한 요구와 이 適合度の 範圍內에서 자기가 지불할 費用을 미리 提示할 수 있다. 그림 3은 이것을 나타낸 것이다. 그림의 頂上部는 利用者が 支拂코자 하는 最大費用과 그 費用限度內에서 받아들일 수 있는 最低適合率을 表示한 것이고 下低部는 받아들일 수 있는 最低適合率과 그 適合率에 대해서 支拂하고자 하는 費用을 나타낸 것이다. 물론 주어진 料金에 대해서 利用者は 最低適合率을 提示했지만, 그림 3의 斜線部分과 같이 더 높은 適合率의 서비스를 받을 수도 있다.

위의 방법은 이용자의 選好를 性能시물레이션과 관련지우는 것이며, 시물레이션으로부터 費用모델을 誘導하는 것은 前述한 評價모델에서와 같이 한다.

費用모델로부터 시스템의 總費用과 利用者 費用을 알 수 있으며 (附錄 2參照), 利用者費用만 알면 利用者集團에 대하여 어떠한 費用水準으로 시스템을 이용할 것인가의 設問을 통하여 利用者の 數와 利用회수를 알 수 있게 된다. 이 修正活用函數를 費用모델에 피이드백 시키므로써 새로운 利用者費用을 算出할 수 있다. 이러한 技法을 반복하여 시스템의 活用函數와 시스템總費



R_p = 利用者が 수금하는 適合度 範圍

R_c = R_p 에 대한 費用 範圍

그림 3. 시스템의 費用과 適合度와의 關係

用이 平衡이 되게 한다. 여기서 알아두어야 할 것은 시스템의 活用函數는 시간에 따라 變化할 수 있는 것이므로 設問時에 장차의 利用行態에 관한 내용도 밝혀 두어야 한다는 점이다. 또 新規利用者의 要求도 미리 豫測하여 活用函數속에 포함시켜 두어야 한다. 이러한 일을 편리하게 하기 위해서는 임의로 선택한 活用函數에 대한 費用모델을 設定하여 시스템의 活用函數에 대한 利用者費用을 利用者가 提示한 費用과 신속히 對照할 수 있도록 해주는 것이 좋다. 表 2는 4개시스템의 4개 活用函數에 대한 利用者費用을 表示한 것으로 각 시스템의 壽命은 5年, 月割引率은 1%로 잡은 것이다. 또 初期支出(固定費)은 100,000달러이고 月1,000달러의 定期的인 支出이 있다. 活用函數는 時間에 대한 2次方程式 즉, 月數를 $i(i=1, 2, \dots, 60)$ 라고 할 때 函數 U 는,

$$U = UA + (UB \times i) + (UC \times i^2) \text{로 된다.}$$

또 可變費用은 시간에 대하여

$$V = VA + (VB \times i)$$

즉 線形增加를 나타낸다. 附錄 2는 4個 活用函數와 4個 可變費用에 대하여 分析한 것이며, 表 2는 시스템의 現在價値, 總費用 및 利用者費用을 나타낸 것이다.

表 2. 시스템(1, 2, 3, 4)의 應用函數(A, B, C, D)

	UA	UB	UC	VA	VB	總費用	費用/利用者(\$)
A.1	100	5	0.1	1000	10	202,289	13.12
A.2	100	5	0.1	1000	20	214,667	13.92
A.3	100	5	0.1	2000	10	247,244	16.04
A.4	100	5	0.1	2000	10	259,622	16.84
B.1	100	10	0.1	1000	10	202,289	9.36
B.2	100	10	0.1	1000	20	214,667	9.94
B.3	100	10	0.1	2000	10	247,244	11.44
B.4	100	10	0.1	2000	20	259,622	12.02
C.1	200	5	0.1	1000	10	202,289	10.16
C.2	200	5	0.1	1000	20	214,667	10.78
C.3	200	5	0.1	2000	10	247,244	12.42
C.4	200	5	0.1	2000	20	259,622	13.04
D.1	200	10	0.1	1000	10	202,289	7.75
D.2	200	10	0.1	1000	20	214,667	8.22
D.3	200	10	0.1	2000	20	247,244	9.47
D.4	200	10	0.1	2000	20	259,622	9.94

表 2에서 各列에는 文字와 數字의 코우드가 적혀 있는데 文字는 活用函數를, 數字는 可變費用函數를 나타낸다. 시스템의 總費用은 可變費用函數만에 左右되며 活用函數에는 영향을 받지 않음을 알 수 있다. 그러나 活用函數의 變化는 可變費用에 多少의 영향을 미칠 것이므로 실제와는 다를 수도 있다. 예컨대, 이것을 表 2에 적용시켜 活用函數가 A에서 B로 變化할 때 可變費用이 1에서 4로 變化한다고 가정하고 A.1과 B.4를 비교해 보면 알 수 있다.

表 2의 모든 시스템이 利用者와 購入者의 性能要件을 만족한다고 가정하면 아래와 같이 費用分析을 할 수 있다. 즉 購入者가 總費用이 250,000달러를 넘는 것을 바라지 않는다면 A.4, B.4, C.4 및 D.4시스템은 除外된다. 또 活用函數A와 B가 利用者費用 10~12달러, C와 D가 8~10달러를 必要로 한다면 A와 C그룹은 역시 除外되고 B.3, D.2 및 D.3은 위의 條件과 平衡된다. B.1과 B.2는 活用函數 C와 D를 包含하고 있지만(利用者費用의 妥當範圍가 8~10달러이므로) C는 이미 除外되었기 때문에 D만이 考慮의 대상이 된다. 시스템 2는 B.2와 D.2가 平衡上에 있으며, 시스템 1은 B.1과 平衡이 되나 D.1은 그렇지 못하며 그 利用者費用은 그 活用函數범위보다 적은데 이는 D.1은 여기서 明示되지 않은 다른 活用函數와 平衡을 이루고 있음을 의미한다.

活用函數가 安定平衡値와 일치하지 않을 수도 있는데 이것은 그 시스템이 대상으로 하는 利用者集團中 特定그룹에 대해서는 使用價値가 없다는 것을 의미한다. 그림 4는 特定시스템의 適合率/再現率의 關係와 2個 利用者集團의 效用函數를 圖示한 것이다. 그림에서 맨 위의 曲線은 시스템의 性能函數曲線과 교차하지 않는데 이것은 利用者의 관심밖의 性能을 나타내는 것이다. 만일 利用者集團 1이 U_1 으로 表示한 最低性能을 要求하고 또 더 싼 費用일지라도 그 이하의 性能을 원치 않는다면 이 시스템은 어떠한 경우에 있어서도 그의 要求를 만족시킬 수 없다. 그러나 利用者 2의 效用函數인 U_2 는 點A에서 適合率/再現率曲線과 接하고 있다. 이것은 利用者 2가 適合率과 再現率이 각각 P_A, R_A 인 條件下

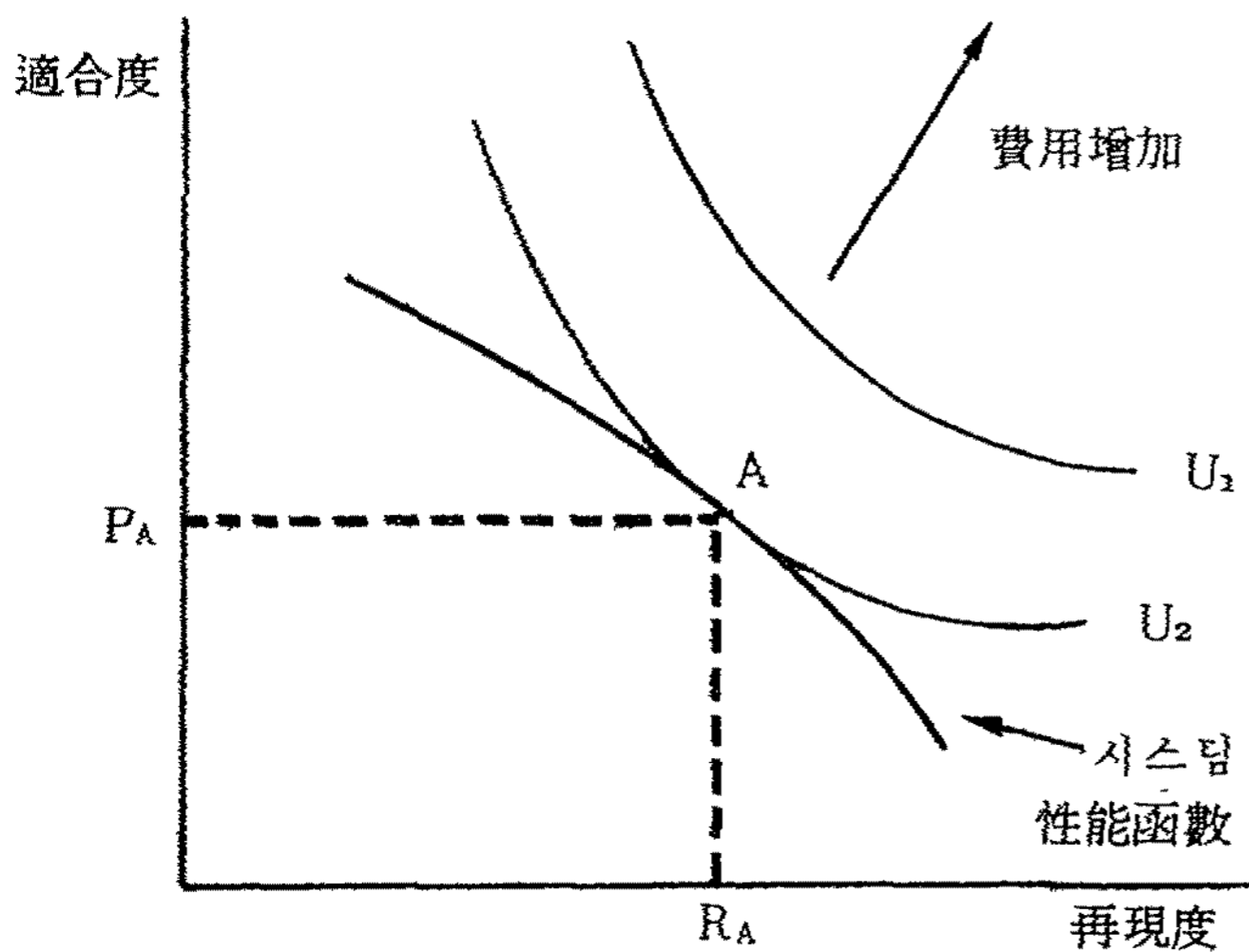


그림 4. 시스템의 性能函數와 利用者의 効用函數와 의 관계

에서 이 시스템을 이용하겠다는 것을 나타내는 것이다. 즉, 점 A는 利用者 2에 대해서는 이 시스템의 最適活用點이 된다, 왜냐하면, 이보다 높은 効用曲線은 性能函數曲線을 교차나 접촉하지 않을 것이고 또 이보다 낮은 効用曲線은 비록 2個 이상의 點에서 性能曲線과 교차하겠지만 効用性은 더 적기 때문에 利用者 2는 最大効用(이 경우는 點 A)을 얻기 위해서 더 많은 費用을 支拂코자 할 것이다.

4. 限界와 論評

効用函數의 組合이나 比較는 사실상 不可能한 일이므로 單純化를 위한 여러 가지 假說을 설정해 둘 필요가 있다. 가령, 特定 利用者集團의 構成員들이 모두 具體的인 同一性能要件을 가지고 있다면 이 集團은 個人으로 看做할 수 있으며, 特定性能要件과 集團을 代表하는 個人의 單一効用函數를 통한 費用과의 관계만 알면 받아들여질 수 있는 利用者費用 범위를 알 수 있다. 따라서 利用者集團이 시스템을 이용할 것인가의 판단은 費用만으로도 가능해진다.

Lesk³⁾는 시스템의 評價와 比較에 상당한 영향을 미칠 것임에도 불구하고 일반적으로 費用 모델에서 除外되는 다음과 같은 因子들을 지적하고 있다. 즉, ① 機械使用費用은 時間이 경과함에 따라서 減少된다. ② 技術의 發達에 의하

여 費用計算에 豫期치 않은 변화가 생길 수 있다. ③ 전반적인 費用이 인플레이션의 영향으로 증가한다.

5. 購入者-性能모델 相互作用

시스템의 購入者는 이용자에 비하여 시스템의 성능에 대한 요구정도가 낮다고 볼 수 있다. 즉, 구입자는 단지 그의 資金사정과 시스템總費用을 따져서 받아들여든가 거절하면 되기 때문이다. 그러나 일반적으로 購入者는 特定性能要件에 대해서는 이용자보다 더 意圖的으로 이를 具體化하고자 한다. 예컨대, 購入者는 特定部分에 대해서는 이용자가 지정한 性能限界에 대해서 同意하지 않을 수도 있다. 따라서 性能시뮬레이션에 있어서는 購入者나 利用者 쌍방의 性能基準을 반드시 計算에 넣어야 하며 가능하면 미리 정해진 尺度로써 그 무게를 實測하여 計算하되 이용자의 요구에 더 비중을 두어야 한다. 결국, 購入者는 費用/性能 評價의 組合에 의하여 購入決定을 하게 된다.

6. 結 論

利用者와 購入者의 시스템評價에 있어서의 役割을 擴張하고 더 많은 複合要素들을 費用모델에 導入함으로써 情報檢索시스템의 評價 및 比較에 보다 實際的인 接近이 可能하리라고 본다.

參 考 文 獻

- 1) C. P. BOURNE et al., *Requirements, Criteria, and Measures of Performance of Information Storage and Retrieval Systems*, Final Report to the National Science Foundation on Stanford Research Institute Project 3741, Menlo Park, California, S. R. I., AD 270942 (Dec. 1961).
- 2) *Research Study of the Criteria and Procedures for Evaluating Scientific Information Retrieval Systems*, Final Report to the National Science Foundation, Contract NSF-C218, New York, A. Anderson & Co., AD 273 115 (Mar.

- 1962).
- 3) M. LESK, *Man-machine Relations in Future Information Centers*, (unpublished paper).
 - 4) G. SALTON, *Automatic Information Organization and Retrieval*. McGraw-Hill, New York (1968).
 - 5) D. H. ROTHENBERG, *An Efficiency Model and a Performance Function for an Information Retrieval System*, Comparative Systems Laboratory Report No. 13, Cleveland, Ohio, (Nov. 1967).
 - 6) H. BIERMAN and S. SMIDT, *The Capital Budgeting Decision*, 4th Edn. MacMillan, New York (1975).
 - 7) C. P. BOURNE and D. F. FORD, Cost analysis and simulation procedures for the evaluation of large information systems. *Am. Docum.* 1964, 142.
 - 8) K. W. WEBB et al., *Evaluation Models for Information Retrieval and Command Control Systems*. IBM Corporation, Bethesda, June 1964.

附錄 1.

○ 投資評價의 現在價值法

投資의 現在價值란, 投資者가 投資를 마련하는 機會에 대하여 支拂하는 最高額을 말한다. 가령 2年안에 支拂할 수 있는 金額이 100달러라고 하면 그 現在價值는 2年후에 100달러를 얻기 위해서 複利로 計算해서 지금 당장 投資해야 할 금액을 말하며 이것은 金錢이 불어나는 利子率과 複利計算빈도에 좌우되는 것이다.

計算例 ; 지금 資金은 每年 複利로 計算되고 利子率은 13%라고 하자. 이 投資는 2年후에는 100달러의 回收가 보장되나 初年度末에는 回收가 전연 안된다. 오늘 3%의 利子率로 投資한 1.00달러는 每年 複利로 增加하여 2年후에는 1.0609달러가 된다. 따라서 2年후 100달러의 現在價值는 $100 \div 1.0609 = 94.26$ 달러, 즉 複利로 計算되는 年 3%의 利子收益이 있는 現在의 94.26달러는 2年후의 100달러와 同一한 價値를 가진다.

一般式 ; 投資의 現在價值를 PV , 期間을 n , 利子率을 r , 各年の 現金흐름을 $Q_i (i=1, 2, \dots, n)$ 라고 하면,

$$PV = \frac{Q_1}{(1+r)} + \frac{Q_2}{(1+r)^2} + \dots + \frac{Q_n}{(1+r)^n}$$

즉, $PV = \sum_{i=1}^n Q_i (1+r)^{-i}$ 이 成立된다.

또 Q_0 를 初期現金支出(創業投資)이라고 하면 投資의 純現在價值는 $NPV = Q_0 + PV$ 이다.

Bourne과 Ford⁷⁾는, 投資評價에 있어서 現在價值法과 等價年間費用法을 併用하고 있다. 等價年間費用은 現在價值를 投資年數로 나눈 값을 말한다.

附錄 2.

○ 시스템의 現在價值 算出을 위한 費用모델

이 모델은 시스템의 모든 費用을 주고서 現在價值를 計算하는 것이다. 이때 可變費用은 時間에 따라 線形增加한다고 가정하고 또 시스템의 壽命과 收益率도 미리 정하여 둔다. 純現在價值는 附錄1의 一般式을 이용하여 算出할 수 있으며, 全現金흐름은 支出로 본다. 따라서 純現在價值는 그 시스템의 總費用을 나타내게 된다.

다음에는 利用者費用을 計算한다. 이때 活用函數는 시간에 대한 2次方程式으로 表示되며, 單位利用者費用은 各 利用者에게 費用을 부과했을 때 純現在價值가 零(Zero)이 되는 價格 즉, 시스템의 採算點이라고 가정하고 각 利用者は 시스템의 사용에 대하여 定額을 지불한다고 본다. 지금 U_k 를 K 기간 동안의 利用者數, 純現在價值가 零으로 되게 各 利用者에게 부과하는 價格을 P 달러라고 하면, K 기간 동안의 收益은 $U_k \times P$ 달러가 되며 이것을 $(1+r)^k$ 로 나누어서 割引한다. 여기서 r 는 期間收益率이다. 따라서 總現在價值가 零일 때, 시스템總費用 = $\sum_{k=1}^n U_k \times \$P \times (1+r)^{-k}$ 이다. 단, n 은 시스템壽命동안의 期間數이다. 만약 P 가 時間의 影響을 받지 않는다면 式은 $P = \text{總費用} / \sum_{k=1}^n U_k (1+r)^{-k}$ 로 된다.

이 式을 사용하여 表 1의 P 를 計算할 수 있다.