# A Value-Driven Model for Data Manufacturing: An Application for Optimal Error-Correction Policy

Adir Even & G. Shankaranarayanan

Boston University School of Management, Information Systems Dept, 595 Commonwealth Ave., Boston, MA  02215,
P 617-353-2523, {adir, gshankar}@bu.edu

## 1.    INTRODUCTION

Data is an essential resource to the modern business. Corporate investments in data management solutions are increasing along with growing concerns regarding their cost and contribution to profitability. A data-related factor identified as having a significant effect on cost and profitability is the quality of the organizational data resources [Redman, 1996]. The Data Quality Management (DQM) literature addresses the quality issue at different levels – from the high-level paradigm of Total Data Quality Management (TDQM) [Wang, 1998], to specific monitoring and improvement methodologies [e.g., Redman, 1996, Ballou et al., 1998]. Research, however, has not examined to the same extent, the economic contribution of DQM initiatives to business-value and profitability. The high costs of DQM initiatives make their contribution important to understand. To what extent do organizations gain value from data quality improvements? Does the value gained offset costs? How should considerations of value and cost affect DQM decisions?

This study is a step towards better understanding the business value of DQM activities. It adopts the TDQM view of data management systems as a data manufacturing process (DMP) and their output as an information product (IP) [Wang, 1998, Ballou et al., 1998]. It introduces a value-driven framework for evaluation and optimization of DQM initiatives, based upon a DMP model that quantifies quality hazards and the value of associated managerial decisions. A new concept that this model introduces is the inclusion of business-value, cost, and profit-maximization considerations into process-optimization decisions. This allows assessing the overall cost of creating an IP, assessing the potential increase in IP profitability, and developing value-based optimal policies for managing quality. Section 2 introduces the DMP model that allows evaluation of decision alternatives, considering cost, value, and the accumulating stochastic effects. Section 3 demonstrates the use of the model in the development of a value-driven optimal error-correction policy. Section 4 offers conclusions and suggests directions for future research.

## 2.    A VALUE-DRIVEN MODEL FOR DMP/IP

The typical DMP has a complex setting of multiple-inputs and multiple-outputs [Ballou and Pazer, 1985b]. Literature has suggested techniques for modeling the DMP by mapping it onto a directed-network of processing stages [Ballou et al., 1998, Shankaranarayanan et al., 2003]. A multi-stage DMP model is typical in data warehouses (DW). Due to the complexity and the high volumes of data processed, the DW is typically vulnerable to data quality hazards. Parssian et al. (2004) examine the propagation of errors, originating at different data sources, through the DW processing stages and their effect on the IP outcome. Cui et al. (2000) develop a lineage-tracking mechanism for detecting the source of errors that are identified at final DW stages. Managing the complexity of the DW and maintaining high-level of quality requires substantial support from metadata – abstracted information about data and the systems that manage it [Shankaranarayanan and Even, 2004].

The proposed framework aims to support value-driven management of a complex DMP such as the DW. The processing-stage formulation (influenced by dynamic programming techniques [Bertsekas, 2000]) uses metadata characteristics as an input and incorporates an analytical representation of data transformations, random quality hazards, and the effect of managerial decision choices for preventing hazards or minimizing their damage. The staged-representation - a networked structure that defines the entire DMP and places the related technical decisions in an economical context - provides a powerful management tool for DMP. It allows analyzing data flows, detecting possible sources of quality errors, quantifying their accumulated effect on the profitability gained by the information product (IP), and evaluating decision alternatives for maximizing this profitability. The analytical development of such a model is complex. As a first step, this paper focuses on a sequential DMP with the final stage representing the IP. Analyzing the sequential (single input/output per-stage) case first, before addressing a more comprehensive staged-network, is a common analytical approach for developing error-correction policies (e.g., [Ballou and Pazer, 1985a], [Tayi and Ballou, 1988], [Chengalur et al., 1992]).

The suggested DMP/IP model (Figure 1), incorporates the following constructs:

***Processing Stages*** $(S^n_{n=0...N})$***:*** The DMP has $N+1$ processing stages. $S^0$ represents a data-source and $S^N$ is the final stage, where data is consumed by end-users. Stages $S^1$ to $S^{N-1}$, represent intermediary processing and manipulation (e.g., transfer, cleansing, or aggregation).

***Metadata Vector*** $(X^n_{,n=0...N})$***:*** The metadata vector is a collection of characteristics that describes the dataset (such as last update date/time, number of records, or quality level)

***Stage Transformation*** $(L^n_{,n=0...N})$***:*** the data transformation associated with stage $S^n$, is annotated $L^n$, such that $X^n=L^n(X^{n-1})$, where $X^{n-1}$ denotes metadata associated with the input-data entering the stage and $X^n$ denotes

*Figure 1. Value-Driven Sequential DMP/IP Model*
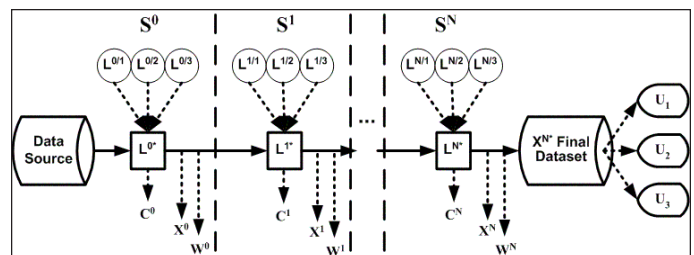
the metadata associated with the output-data after processing. In reality, the transformation may not be deterministic, but the model assumes that when needed, an equivalent deterministic approximation can be obtained.

***Random Quality Hazards ($W^n_{,n=0..N}$):*** Quality hazards might occur at any stage. For annotation purposes, that quality hazards at $S^n$ occur *after* the data has been processed at $S^n$ and before being transferred to $S^{n+1}$. Hazards are annotated as a random deviation vector $W^n$ of identical dimensionality as $X^n$, with a known probability distribution $F_n(W)$. Hence,

$X^{n*} = X^n - W^n$, and $X^{n+1} = L^{n+1}(X^{n*}) = L^{n+1}(X^n - W^n)$, where:

$X^n$ – The metadata vector of a dataset after it was processed at stage $S^n$

$W^n$ – Metadata deviation due to random quality hazards at stage $S^n$

$X^{n*}$ - The metadata representing the actual data that is forwarded to stage $S^{n+1}$

$L^{n+1}$ – The transformation associated with the following stage $S^{n+1}$

***Transformation Alternatives and the Optimal Transformation ($\{L^n\}$, $L^{n*}$, $_{n=1..N}$):*** Different transformation alternatives $\{L^n\}$ can be considered at stage $S^n$ to minimize quality damages. The data management decision is selecting the optimal transformation $L^{n*}$ among a set of feasible transformations to maximize overall profitability P (see definition below). The chosen optimal transformation is assumed to affect the quality hazards, and so does the metadata, hence the probability distribution function of quality hazards at stage *n* ($F_n(W)$), is a function of $X^n$ and $L^{n*}$.

***Cost ($C^n_{,n=0...N}$):*** Applying a transformation at $S^n$ introduces a cost $C^n$, expressed in monetary values, which can be influenced by the transformation choice, the input/output data, and the random quality hazards, hence $C^n = c^n(L^{n*}, X^{n-1}, W^{n-1}, X^n, W^n)$

***Utility ($U^i_{, i=1..P}$):*** The final information product (IP) output at $S^N$ can be used in multiple different business contexts, each associated with a monetary utility $U^i$, reflecting the business-value of its use in that context. The effect of the IP characteristics ($X^{N*}$) on the contextual utility $U^i$ is described by an utility function [Ballou et al., 1998], $U^i = u^i(X^{N*})$

***Profitability (P):*** The profitability gained by implementing the DMP is defined as the margin between the overall utility gained by the information product and the overall cost, or $P = Ó_i U^i - Ó_{n=0..N} C^n$. Alternatively, considering stochastic behavior, the expected profitability over time (denoted $E[.]$) is given by $E[P] = Ó_i E[U^i] - Ó_{n=0..N} E[C^n]$.

The objective of using such a model is to select a set of optimal stage-transformations $\{L^{n*}\}$ to maximize profitability ($P$, or $E[P]$). Such a value-driven evaluation and optimization approach can address design choices (such as the design of a tabular dataset in [Even et al., 2005]) as well as DMP maintenance decisions. We address the latter by using the proposed model to develop an optimal error correction policy.

## 3. OPTIMAL ONLINE ERROR-CORRECTION POLICY

Error detection and correction is an archetype DQM approach (the others being process-control and process-design) addressing short-term "cures" to immediate needs rather than long-term solutions that target root-causes [Redman, 1996]. Error-detection compares data to a baseline that is perceived as being correct (e.g., the "real world", a set of business rules, or another dataset). Error-correction considers alternatives such as manual edits, automated fixes, data retransmission, or even providing

data "as is" when fixing it is too costly. Optimal inspection and correction policies that target maximal data quality level have been analyzed [Tayi and Ballou, 1988], [Chengalur et al., 1992]. However, approaching the highest quality might be sub-optimal from a broader economic perspective. Inspection and correction alternatives introduce trade-offs between the implementation difficulty and the level of quality obtained, i.e., between cost and value. Considering such factors, the goal of maximal quality might come at too-high a cost, and/or insignificant improvements.

The suggested framework incorporates cost and utility, hence, can better reflect the economic perspective. To demonstrate this argument, the sequential DMP/IP model is used here to develop an optimal error-correction policy. The dynamic programming algorithm used fundamentally assumes:

1. Optimal transformations $\{L^{n*}\}$ are applied online, after the damage done by the random quality hazard has been assessed.
2. The outcome of stage $S^n$ is independent of information at stages prior to stage $S^{n-1}$
3. Stochastic behavior at stage $S^n$ does not depend on previous or following stages
4. Costs and utilities are sum-additive.

Within these assumptions the algorithm for choosing stage-transformations is:

1. At $S^n$ we define the "Forward-Profitability" function $J^n = Max_{[Ln]}E[J^{n+1} - C^n]$
2. The boundary condition for forward-profitability at the final stage $S^N$ is $J^{N+1} = Ó_i U^i$
3. The transformation is chosen according to an optimal policy $L^{*k} = argmax \{E[J^{n+1} - C^n]\}$

It has been proven that within the given assumption, such a policy maximizes the expected profitability over time [Bertsekas, 2000]. However, the fundamental algorithm is insufficient as is and making it useful requires further quantification of model constructs.

In the scenario analyzed, the primary factor that affects utility and cost is the intrinsic value [Even and Shankaranarayanan, 2005], a sum-additive measure of business-activity that is associated with the processed data (e.g., sale amounts, revenues, or customer lifetime value). The process is assumed to be value-preserving – all stages are designed to maintain the intrinsic value, and losses are attributed to random quality failures. All stages have a built-in capability to reprocess the data such that the intrinsic-value loss can be recovered (e.g., reprocessing a damaged data-subset or manually correcting it). Value-loss recovery, however, comes at a cost that is linearly proportional to the actual loss. The goal is to develop an online error-correction policy that helps decide whether or not to recover the value-loss at each stage, such that the overall expected profitability is maximized. This scenario can be formulated as follows:

- The DMP has $N+1$ processing stages $S^n_{, n=0..N}$
- Expected profitability, as suggested by the general model, is the margin between the overall expected utility and the overall expected cost, or $E[P] = \Sigma_i E[U^i] - \Sigma_{n=0..N} E[C^n]$.
- The primary factor affecting utility and cost is the intrinsic-value. Therefore, the metadata variable at $S^n$ is modeled as a scalar $X^n \geq 0$ and the actual intrinsic-value is denoted $x^n$
- $W^n$ is the random intrinsic-value loss at $S^n$, $w^n$ is the actual value loss at $S^n$, $0 < w^n < x^n$
- $F_w^n(w,x) = P(W^n < w/X^n = x)$ is the conditional loss distribution function at $S^n$. We assume that $F_w^n(w, x)$ can be obtained once $x^n$ is known
- Once $x^n$ is known, the *loss ratio* at $S^n$ is defined as $q^n(x^n) = E[W^n/X^n = x^n]/x^n$, where $q^n(x^n)$ is within $[0,1]$. We assume that the expected loss at $S^n$ is approximately linear with the value, $E[W^n/X^n = x]H''q^n*x$. and that the set of $\{q^n\}$ can be approximated in advance.

- $L^n$, the transformation at $S^n$, is coded either as 1, when value-loss is recovered or as 0 when no action is taken. Hence, it can be formulated as $X^n = X^{n-1} - (1-L^n)*W^{n-1}$
- $C^n$, the cost at $S^n$ is zero if no action is taken $(L^n = 0)$, or increasing with the value-loss if corrective measures are taken $(L^n = 1)$. This model assumes a correction cost that increases linearly with value loss, hence, $C^n = c^n *L^n*W^{n-1}$.
- The metadata vector of the final IP is denoted $X^{N*}$. We assume no value-loss at $S^N$, hence $X^{N*}=X^N$. Contextual utilities are assumed to be linear with the intrinsic value: $U^i = a^i*X^N$
- The forward-profitability at $S^N$ is defined $J^{N+1}=\Sigma_i U=\Sigma_i a^i*X^N=a*X^N$, where $a=\Sigma_i a^i$

The suggested DP algorithm for this setup is:

1. Obtain the forward-profitability function $J^n$ defined as $J^n=Max_{L_n} E[J^{n+1} -C^n]$
2. The optimal transformation-choice policy is $L^{n*} = argmax\{E[J^n]\} = argmax\{E[J^{n+1}-C^n]\}$
3. The boundary forward-profitability function at the final stage is $J^{N+1} = a*X^N$

***Proposition (proof by induction in Appendix A):*** The process is optimal, in terms of maximizing profitability, when the following backward-recursion decision-rules are applied at stage $S^{N-i}_{\ i=0..N}$:

1. Obtain the marginal value $a^{N-i} = a^{N-i+1} - q^{N-i}*Min\{c^{N-i+1}, a^{N-i+1}\}$
2. Recover the value-loss $(L^{N-i}=1)$ if $c^{N-i}<a^{N-i}$, or take no action $(L^{N-i}=0)$ otherwise
3. The forward profitability function is given by $J^{N-i}=a^{N-i}*X^{N-i-1} - Min\{a^{N-i}, c^{N-i}\}*W^{N-i-1}$
4. For boundary condition, consider $a^{N+1}=á$, $c^{N+1}=0$ and $q^N=0$.

A special case is when the marginal costs are fixed at all stages $(c^n=c, _{n=0..N})$. Two possible scenarios requiring further analysis are:

1. The marginal fixing cost is larger than the marginal utility $(c>á)$:
- Since $c>a^N=á$, no action is taken at the final stage $S^N$ (hence, $L^N=0$)
- $Min\{c,a^N\}=a^N$ and $a^{N-1}=a^N-q^{N-1}*a^N=(1-q^{N-1})*a^N <c$, no action is taken at stage $S^{N-1}$
- Similarly, since $a^{N-i}$ keep decreasing with [i], no action is taken at any stage

2. The marginal fixing cost is equal to or smaller than the marginal utility $(c > a)$:
- Since $c\leq\alpha=a^N$, value-loss is recovered at stage $S^N$ (hence, $L^N=1$)
- $Min\{c,a^N\}=c$ and $a^{N-1}=\alpha-q^{N-1}*c<\alpha$
- As $a^{N-i}$ decreases, when $a^{N-i}\leq c$, no action is taken for all stages before and including stage $S^{N-i}$ or value-loss is recovered for all stages subsequent to $S^{N-i}$

The scenarios are illustrated in Figure 2. In (1) the marginal utility (increasing line) is always higher than the marginal cost (fixed) and therefore, value-loss is recovered at all stages. In (3) the marginal utility is always lower and therefore no action is taken at any stage. In (2), the initial marginal utility is low and no action is taken, but at a certain stage it becomes higher than the cost, hence, from that stage forward value-loss is recovered.

*Figure 2. Scenarios of Fixed Marginal Cost*
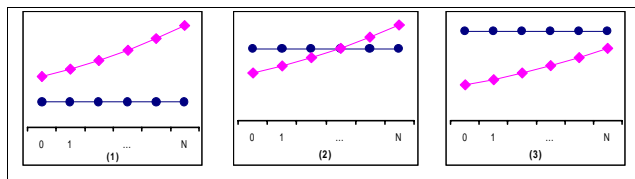


It is important to note that this error-loss policy addresses recovery of value-loss at the previous stage only. Had recovery been applied to overall loss, going backwards to all the previous stages, costs at later stages would have been significantly higher, resulting in a different optimal policy – concurring with DQM "good practices" that recommend correcting errors as close to the source as possible.

## 4. CONCLUSIONS & DIRECTIONS FOR FUTURE RESEARCH

Data management today is typically technology-driven and economical issues are not sufficiently addressed. To better understand the economic factors that affect data management decisions, this study looks inside the "black-box" of the DMP - although *business-value* is associated with the process output (the IP), *the costs* are attributed to the process that creates it. Value and cost are linked through the dataset characteristics at each stage, represented as a metadata vector. Design and maintenance decisions affect these characteristics, consequently affecting value, cost and profitability. Unifying these pieces into a single framework allows better assessment of alternatives where the goal is to maximize profitability. Furthermore, the directed-network representation of the DMP highlights the impact of local decisions on the cost at subsequent stages, the value of the final output, and on overall profitability.

The optimal error-correction policy demonstrates an application of the framework for data maintenance. The result underscores the argument that considering economical aspects may lead to different data management decisions. Different from the technically-driven approach, economically, it may be sensible to allow data imperfections if the value obtained can not justify the overhead cost. However, this study examined a specific case of process models (sequential) and considered a simplified decision scenario with a limiting set of assumptions and conditions. The focus was deliberately limited to better illustrate and easily formulate the value-driven model and to obtain a closed-form analytical solution.

Further development of the process model ought to reflect operational DMP settings, addressing complex scenarios such as source integration, repeated processing, and multiple IP's. Model extension can also address alternatives to the simplifying assumptions and formulations that underlie the suggested optimal policy. Obtaining empirical support is also necessary. This can be done by examining the suggested value and cost formulations, identifying data and system characteristics that influence them, and looking into the dynamics of data processing within a real-life (or well-simulated) DMP. Such empirical testing is challenging: business processes that integrate data are complex and often involve complementary resources. Therefore, exploring a valuation framework may require developing techniques for mapping and attributing value within business processes.

This study reinforces the view that IS implementation ought to be driven by business needs. Organizational IS, particularly data management systems, are fundamental building blocks in the modern business infrastructure. The approach described integrates business *and* technology into the process of IS design and maintenance and hence can contribute to better aligning the two.

**REFERENCES**
1. Ballou D.P. and Pazer, H.L. (1985a), "Process Improvement vs. Enhanced Inspection in Optimized Systems," International Journal of Production Research, 23(6), 1985, pp. 1233-1245.
2. Ballou D.P. and Pazer, H.L. (1985b), "Modeling Data and Process Quality in Multi-Input, Multi-Output Information Systems," Management Science, 31(2), 1985, pp. 150-162.
3. Ballou, D. P., Wang, R., Pazer, H. and Tayi, G. K. (1998), "Modeling Information Manufacturing Systems to Determine Information Product Quality," Management Science, 44(4), 1998, pp. 462-484
4. Bertsekas, D. P. (2000), "Dynamic Programming," 2nd Ed., Athena Scientific

5.  Chengalur I. N., Ballou D.P. and Pazer, H.L. (1992), "Dynamically Determined Optimal Inspection Strategies for Serial Production Processes," International Journal of Production Research, 30(1), 1992, pp. 169-187.

6.  Cui, Y., Widon, J., and Wiener, J. L. (2000), "Tracing the Lineage of View Data in a Warehousing Environment," ACM Transactions on Database Systems, 25(2), 2000, pp. 179-227.

7.  Even, A., and Shankaranarayanan, G. (2005), "Value-Driven Data Quality Assessment," In the Proceedings of the Tenth MIT International Conference on Information Quality (IQ 2005), 2005, Boston, MA, pp. 22-236.

8.  Even, A., Shankaranarayanan, G., and Berger, P.D. (2005), "Profit Maximization with Data Management Systems," In the Proceedings of Twenty-Sixth International Conference on Information Systems (ICIS), 2005, Las-Vegas, NV, pp. 11-25.

9.  Parssian, A, Sarkar, S., and Jacob, V.S. (2004), "Assessing Data Quality for Information Products – Impact of Selection, Projection, and Cartesian Product," Management Science, 50(7), 2004, pp. 967-982.

10.  Redman, T.C. (1996), "Data Quality for the Information Age," Artech, Boston, MA

11.  Shankaranarayanan, G., and Even, A. (2004), "Managing Metadata in Data Warehouses: Pitfalls and Possibilities," Communications of the AIS, 14(13), 2004, pp. 247-274

12.  Shankaranarayanan, G., Ziad M., and Wang R. Y. (2003), "Managing Data Quality in Dynamic Decision Making Environments: An Information Product Approach," Journal of Database Management, 14(4), 2003, pp. 14-32.

13.  Tayi G.K., and Ballou D.P. (1988), "An Integrated Production-Inventory Model with Reprocessing and Inspection," Int. Journal of Production Research, 26(8), 1988, pp. 1299-1315.

14.  Wang, R.Y. (1998), "A Product Perspective on Total Quality Management," Communications of the ACM, 41(2), 1998, pp. 58-65

**APPENDIX A: PROOF BY INDUCTION**

First observe that for $i=0$: $X^N=X^{N-1}-(1-L^N)*W^{N-1}$, and

$J^N = Max_L^N\{E[J^{N+1}-C^N]\} = Max_L^N\{\alpha*E[X^N]-c^N*L^N*E[W^{N-1}]\} =$

$Max_L^N\{\alpha*E[X^{N-1}]-\alpha*E[W^{N-1}]+\alpha*L^{N-1}*E[W^{N-1}]-c^N*L^N*E[W^{N-1}]\}$

The transformation $L^N$ is chosen after $X^{N-1}$ and $W^{N-1}$ are known, hence the expected values can be replaced with deterministic expressions: $E[X^{N-1}]=X^{N-1}$ and $E[W^{N-1}]=W^{N-1}$, Therefore

$J^N = Max_L^N\{\alpha*X^{N-1} - \alpha*W^{N-1} + \alpha*L^N*W^{N-1} - c^N*L^N*W^{N-1}\}$

For $L^N=0$, $J^N=\alpha*X^{N-1}-\alpha*W^{N-1}$ and for $L^N = 1$, $J^N = \alpha*X^{N-1} - c^N*W^{N-1}$

To maximize, if $\alpha>c^N$ the value-loss is recovered $(L^N=1)$, otherwise no action is taken $(L^N=0)$. Therefore, $J^N=\alpha*X^{N-1}-Min\{\alpha,c^N\}*W^{N-1}$, hence, the formulation is confirmed for $i=0$.

Observing that $a^{N+1}=\alpha$, $C^{N+1}=0$ and $q^N=0$:

$a^N = a^{N+1}+q^N*Min\{c^{N+1}, a^N\} = \alpha+0 = \alpha$, and $J^N = a^N*X^{N-1} - Min\{a^N,c^N\}*W^{N-1}$

Following, show that if the formulation holds for $[i-1]$, it holds for $[i]$. At stage $S^{N-i}$:

$X^{N-i} = X^{N-i-1} - (1 - L^{N-i})*W^{N-i-1}$, and $J^{N-i} = Max_L^{N-i}\{E[J^{N-i+1}-G^{N-i}]\} =$

$Max_L^{N-i}\{a^{N-i+1}*E[X^{N-i}]-Min\{a^{N-i+1},c^{N-i+1}\}*E[W^{N-i}]-c^{N-i}*L^{N-i}*E[W^{N-i-1}]\} =$

$Max_L^{N-i}\{a^{N-i+1}*E[(X^{N-i-1}-(1-L^{N-i})*W^{N-i-1})] - Min\{a^{N-i+1}, c^{N-i+1}\}*E[W^{N-i}] - c^{N-i}*L^{N-i}*E[W^{N-i-1}]\}$

The transformation $L^{N-i}$ is chosen after $X^{N-i-1}$ and $W^{N-i-1}$ are known, hence the expected values can be replaced with deterministic expressions: $E[X^{N-i-1}]=X^{N-i-1}$ and $E[W^{N-i-1}]=W^{N-i-1}$.

Also replace $E[W^{N-i}] = E[W^{N-i}|X^{N-i} = x^{N-i}] = q^{N-i}*(X^{N-i-1} - (1 - L^{N-i})*W^{N-i-1})$ and as result:

$J^{N-i} = Max_L^{N-i}\{a^{N-i+1}*X^{N-i-1} - a^{N-i+1}*W^{N-i-1} + a^{N-i+1}*L^{N-i}*W^{N-i-1} - Min\{a^{N-i+1}, c^{N-i+1}\}*q^{N-i}*(X^{N-i-1} - (1 - L^{N-i})*W^{N-i-1}) - c^{N-i}*L^{N-i}*W^{N-i-1}\} =$

$Max_L^{N-i}\{ (a^{N-i+1} - Min\{a^{N-i+1}, c^{N-i+1}\}*q^{N-i})*X^{N-i-1} - (a^{N-i+1} - Min\{a^{N-i+1}, c^{N-i+1}\}*q^{N-i})*W^{N-i-1} + (a^{N-i+1} - Min\{a^{N-i+1}, c^{N-i+1}\}*q^{N-i} - c^{N-i})*L^{N-i}*W^{N-i-1}\}$

Now denote $a^{N-i} = a^{N-i+1} + q^{N-i}*Min\{a^{N-i+1}, c^{N-i+1}\}$ and obtain

$J^{N-i} = Max_L^{N-i}\{ a^{N-i}*X^{N-i-1} - a^{N-i}*W^{N-i-1} + (a^{N-i} - c^{N-i})*L^{N-i}*W^{N-i-1}\}$

# Related Content

Cost Estimation and Security Investment of Security Projects

Yosra Miaoui, Boutheina A. Fessiand Noureddine Boudriga (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 4849-4861).*

www.irma-international.org/chapter/cost-estimation-and-security-investment-of-security-projects/184189

Viterbi Decoder in Hardware

Mário Pereira Véstias (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 6307-6318).*

www.irma-international.org/chapter/viterbi-decoder-in-hardware/184328

An Approach to Distinguish Between the Severity of Bullying in Messages in Social Media

Geetika Sarnaand M.P.S. Bhatia (2016). *International Journal of Rough Sets and Data Analysis (pp. 1-20).*

www.irma-international.org/article/an-approach-to-distinguish-between-the-severity-of-bullying-in-messages-in-social-media/163100

Tradeoffs Between Forensics and Anti-Forensics of Digital Images

Priya Makarand Shelkeand Rajesh Shardanand Prasad (2017). *International Journal of Rough Sets and Data Analysis (pp. 92-105).*

www.irma-international.org/article/tradeoffs-between-forensics-and-anti-forensics-of-digital-images/178165

A Grounded Theory Study of Enterprise Systems Implementation: Lessons Learned from the Irish Health Services

John Loonamand Joe McDonagh (2009). *Information Systems Research Methods, Epistemology, and Applications (pp. 58-72).*

www.irma-international.org/chapter/grounded-theory-study-enterprise-systems/23468