

待ち行列理論を使った管理プロセスの リエンジニアリング

Nico. M. van Dijk (訳・山下英明, 逆瀬川 浩孝)

経営プロセスの再構築あるいはリエンジニアリング (BPR) は、今日の経営環境において、処理能力や技術を最大限に利用し効率的な運用を図るという目標を達成するためのきわめて現実的な考え方です。その重要な評価尺度の1つは、ある生産・仕事が終わるまでの総処理時間です。生産システムにおいてはこの尺度はかなり注目を集めてきましたが、それとは対照的に管理プロセスでは、どうしたら総処理時間を短縮できるかということについてあまり議論されていません。管理プロセスには、様々な仕事のタイプが混在し、個々の処理時間は一定ではなく、従業員には仕事に対する適性がある、というような特徴があります。

本稿では、簡単な待ち行列モデルを用いた考察の結果から、管理プロセスにおける総処理時間を短縮するための一般的なルールがあることを主張していきます。

1. はじめに

ビジネスロジスティクスは、効率的な運用や処理能力の有効利用を表す言葉として広く日常的に使われるようになりました。この言葉は、もともと製造業の分野で使われていましたが、今では様々な産業に広がっています。その適用分野や使われ方によっておおまかに分類すると

- 生産ロジスティクス
- 流通・輸送ロジスティクス

ファン・ダイク アムステルダム大学計量経済学部
Roetersstraat 11, 1018 WB Amsterdam, The Netherlands

やました ひであき 駒澤大学経営学部
〒154 世田谷区駒沢 1-23-1

さかせがわ ひろたか 早稲田大学理工学部
〒169 新宿区大久保 3-4-1

● 管理・情報ロジスティクス

のように分かれていますが、個々の相違についてこれ以上詳細な議論はしないことにします。むしろこれらすべてのロジスティクスは、客を満足させると同時に処理能力を有効に利用するために、仕事の遅れを出さず、総処理時間を短くしたいということから出発しているという共通点に注目して、これを実現するための一般的な考え方や方法が簡単な待ち行列理論を使って導かれることを指摘しておきましょう。管理プロセスは従来の文献においてあまり扱われていませんので、この解説では管理プロセス (ロジスティクス) を対象として分析することにします。他のロジスティクスについても同じように分析することができますから考えてみてください。

管理プロセスでは、通常ある決まった順番で遂行しなければならない一連の仕事を必要とします。最初にも述べましたが、管理プロセスの特徴を挙げると、次のようなものがあります。

1. 情報を全体に伝達しコミュニケーションを図ることの重要な役割
2. 処理時間と転送時間の著しいばらつき
3. 担当者毎に異なる適性
4. 異なった処理要求

情報工学の重要性は一般に広く認識されていますが、管理運営の分野でも必要とされる問題があるということはほとんど理解されていません。たとえば、適切な時と場所に十分な情報が与えられることによって、はるかに効率的な運用やプロセスの設定ができるというようなことです。つまり、情報工学の管理経営における総処理時間短縮への貢献度は、かなり過小評価されているようです。

この理由の1つは、生産や輸送などの他の領域では総処理時間の短縮が利潤に直接反映されるのに対し、

管理プロセスでは一般的に利潤(あるいは費用)に直接反映されないからです。しかし間接的には、管理プロセスにおいても処理時間を短くすることにより、処理能力が効率的に活用され人を削減することができ、また客により大きな満足を与えることができる、という2つの要因から、実質的な利益が得られるのです。

上述の理由のうち1つは、あまり分析されていないので支援する手法が少ない、ということがあげられます。生産・輸送ロジスティクスに関しては、解析的手法やコンピュータ化された手法が盛んに開発されています。たとえば、自動搬送車における保全・スケジューリング・最短ルート等のアルゴリズムの決定支援システムやMRPソフトウェアがそれです。このような手法の多くは、基本的な定量分析やオペレーションズ・リサーチの技法に基づいていますが、管理プロセスにおいては残念ながらこのような手法はほとんど利用されていませんし、存在すら知られていません。

したがって、本稿では、管理ロジスティクスにおける総処理時間を短縮する一般的な考え方や提案が、簡単な待ち行列モデルの分析から得られるということを目指していきます。また、本稿は2つの読者のグループ、すなわち、これらの方法を理解してほしい経営者の方々とこれらの方法を研究しているORの専門家の方々とを対象とします。具体的に言えば、この解説は管理プロセス(ロジスティクス)に関して

- 処理能力の配分
- ばらつきの影響
- 仕事の配置

の3つの側面の重要性に焦点をあて、

- 一般的な洞察と法則
- 実際の事例研究

について述べて行きます。この事例研究の結果は、管理環境下におけるBPRを目的とした一般的な法則の有効性を立証することでしょう。

以下の第2節では、まず待ち行列理論の基本的な法則を概説します。これによって、いくつかの一般的な洞察や法則が導かれます。第3節では、これらの法則を実際のある保険会社に適用した事例を紹介し、それを評価してこの解説を締めくくります。

2. 待ち行列モデルにおける重要な洞察と法則

この節では、効率、処理能力、ばらつきという概念を考えることによって、いくつかの一般的な簡単な結果を導きます。最近、高橋幸雄氏の待ち行列理論に関するすぐれた初級向け解説が本誌に掲載されたので[4]、基本となる待ち行列理論に関する説明や表現は簡単に触れるだけに留めておきます。また、この節では仕事とかプロセスという代わりに、客とかサービスという言葉を用いることがあります。

2.1 客とシステムの関係と効率

待つという現象の分析に入る前に、待ち行列は大まかにいえば、**客の要求とシステムの要求の違い**から発生するという点を注意しておきましょう。客が望むことは、できるだけ遅れ時間や総処理時間が短く、しかもできるだけ待たないで済むことです。一方、一般的にシステムが目指すことは、できるだけ多くの処理量を可能な限り少ない処理能力で行うことです。

これら双方の要求が矛盾しているということは明らかです。BPRでは「効率」という考え方が重要ですが、それは普通

$$\text{効率} = \frac{\text{処理量}}{\text{処理能力}}$$

$$= \frac{\text{単位時間あたり完了したサービス数}}{\text{単位時間あたり完了可能な最大サービス数}}$$

と定義されます。しかし、この考え方には遅れ時間や総処理時間というものまったく考慮されていませんから、客の要求の満足度を測ることはできません。たとえば、以下の2つの状況を比べてみましょう。

(i) 1時間あたり $\mu = 30$ 人の客を処理できる単一設備に、1時間あたり $\lambda = 20$ 人の客が到着する場合(μ をサービス率、 λ を到着率という)

(ii) (i)と同じサービス率 μ をもつ2つの設備からなるシステムに、倍の客が到着する($2\lambda = 40$)場合

すぐ分かるようにこの2つのシステムの「効率」は

$$\rho = \frac{\lambda}{\mu} = \frac{2\lambda}{2\mu} = 66.7\%$$

となって、同じになります。しかし、待ち行列理論を用いると、総遅延時間あるいは客のシステム滞在時間は(i)では6分なのに対し、(ii)では3.6分となり違った結果になります。このように、効率が同じでも、滞在

時間を60%短縮できるシステムがあることがわかります。実際の遅れ時間や滞在時間は客からすれば大変重要な評価尺度ですので、この例は少なくとも効率だけを目標にすべきではないということを示しています。

2.2 待ち行列モデルにおける2つの重要な要因

待ちの現象に関して興味がある問題としては次のようなものがあります。

1. そもそもなぜ遅れが生じるのか?
2. どの要因が最も遅れに影響するのか?
3. 遅れを表す適切な尺度は何か?
4. それはどのように測れるか?

これらは次のような具体的な問題につながります。

- 何人の作業員を投入すべきか?
- 総処理時間はどれくらい改善(短縮)されるのか?
- どんな作業員配置(たとえば、直列や並列)を用いるべきか?

高橋幸雄氏の解説[4]でも説明されているように、遅れをもたらす、その時間に影響を与える最も重要な要素は、**処理能力とばらつき**です。処理能力が重要な要因であることはすぐに分かりますが、客の到着の仕方やその客が要求するサービス(処理)時間が変動する状況、つまりばらつきを考えた場合に、混雑してきたときの遅れ時間のふるまいはあまり理解されていないでしょう。[4]の例にあるように、客の到着やサービス時間が変動する場合には、システム利用率が100%に満たない場合でも、遅れ時間は実際の平均サービス時間の何倍にもなります。処理能力に余裕があるにもかかわらずこのように遅れ時間が激増するという現象はなかなか実感できませんし、そのようなことを想定して対策を立てることなどありません。管理プロセスでは、それを構成するそれぞれの仕事がこのように遅れの現象を起こしますので、管理プロセスにおける遅れが容易に発生するということが理解できるでしょう。単に**処理能力が十分であるだけでなく、注意深く処理能力を配分することがきわめて重要な**のです。

サービスに対する注文や要求の発生(到着)間隔のばらつき、サービス時間のばらつき、仕事のタイプのばらつき、実際に利用可能な処理能力のばらつき等のばらつきは、一般に遅れが発生する主な原因としてあまり認識されていません。初待ち行列理論でしばしば取り上げられる例ですが、ある停留所でのバスの平均待ち時間は、直感的にはバスがどのように到着しても変わらないと思われそうですが、実際はバスの到着間隔のばらつきが大きくなるほど長く待たされるようになります[4]。このばらつきのマイナス効果に気づくことが、**待ち時間と処理時間を減少させる現実的な提案に通じる**わけです。たとえば、一般にサービス要求時間が異なる様々な仕事を同一の作業ユニット(たとえば人間)に課すべきでなく、別々の作業ユニットを用いる方がよいということが出来ます。あるいは、全体の時間を分割し、仕事ごと(あるいは大体同じ時間要する仕事ごと)に遂行する時間帯を定めてもよいでしょう。おおまかにいえば、これは**処理能力の分割**または**時間の分割**のどちらかを実行すべきであるということです。これに関連して、異なった種類の仕事や客を組み合わせるべきかどうか、すなわち処理能力の共同利用の問題があります。

2.3 処理能力を共同利用すべきか

一般に、処理能力を共同利用することは効率的であると思われています。アンケートを行ってみますと、郵便局等において全ての窓口に対して一列になって順番を待つ方式の方が、それぞれの窓口に別々に並ぶ方式よりも好ましいと答える人が多いようです。心理学的にみてこの理由は明らかで、一列に並んで順番を待てば公平だと考えるからです。定量的な分析によってもこの一列に並ぶ方式の有利さがある程度説明できます。

例として、2つ窓口がある状況を考えてみましょう。2つの窓口に一列で並ぶ場合、その列に対するサービスの速さは各窓口に一列で並ぶ場合の2倍ですが、サービスの要求量も2倍になるわけですから、各窓口で別々に並ぶ場合と比べて平均待ち行列長はあまり変わらないように思えます。しかし、2つの窓口に一列で並べば、片方の窓口にだけ客が待っていてもう一方の窓口は空いているといった状況にはならないので、この方が待ち時間は少なくなります。この種のことは皆さんも実際によく経験することでしょう。

しかし、我々は待ち時間にもっとも影響がある重要

な要素であるばらつきをまだ考慮していません。待ち時間は到着とサービス時間のばらつきに非常に影響を受けるので、一列に並べてしまうと、異なるタイプの仕事が含まれている場合、特定の種類の仕事に対する平均待ち時間が長くなる可能性があります。

数値例

ここで簡単な数値例を挙げてみましょう。2つの窓口に、2種類の客が到着する場合を考えます。タイプ1の客の到着率は50(人/時間)で、サービス要求時間は1分(一定)であるとします。一方タイプ2の客の到着率は5(人/時間)で、サービス要求時間は10分(一定)であるとします。いま、各窓口の担当を分けて、1つの窓口にはタイプ1の客だけをサービスさせ、もう1つの窓口にはタイプ2の客だけをサービスさせてみます。このとき、[4]にあるような一定サービス時間の待ち行列理論の公式により、タイプ1の客とタイプ2の客の平均待ち時間 W_1 、 W_2 、そして全体の客の平均待ち時間 W は、それぞれ以下ようになります。

$$W_1 = 2.5 \text{ (分)}$$

$$W_2 = 25 \text{ (分)}$$

$$W = \frac{10}{11}W_1 + \frac{1}{11}W_2 = 4.54 \text{ (分)}$$

しかし、どちらのタイプの客も1つの待ち行列に並んで、両方の窓口がどちらのタイプの客もサービスできるようにすると、サービス要求時間にばらつきが発生します。実際、任意の客がタイプ1であるかタイプ2であるかはそれらの到着率の比で決まりますので、サービス要求時間の平均 S と分散 σ^2 は次のようになります。

$$S = \frac{10}{11} \times 1 + \frac{1}{11} \times 10 = 1.82 \text{ (分)}$$

$$\sigma^2 = \frac{10}{11}(1 - 1.82)^2 + \frac{1}{11}(10 - 1.82)^2 = 6.69$$

ばらつきがなければ、処理能力の共同利用によって平均待ち時間が短縮したかもしれませんが、客のタイプによるサービス要求時間のばらつきによって、全体の客の平均待ち時間は次のようになります。

$$W = 7.14 \text{ (分)}$$

このように、客全体として平均遅れ時間が長くなるだけでなく、大部分を占めるタイプ1の客の平均遅れ時間がかなり増加していることが分かります。このことから、客のサービス要求時間があまり変わらない場合は

処理能力を共同利用し、サービス要求時間がかなり違う場合は処理能力を分割するとよいことがわかります。どの程度分割するかは、客のタイプの相違とばらつきの大きさ次第ですが、シミュレーションと待ち行列理論を組み合わせて解析することが、結論を得るために最も有効な手段になるでしょう。

また、処理能力をなるべく有効に使うために、サービス要求時間によって客に優先順位をつけて共同利用することも考えられます。たとえば、10分の客をサービスすることになっている窓口でも、10分の客がいなるときには1分の客を扱えるようにするという事です。もちろん、窓口は常に本来サービスすべき客(10分用の窓口は10分の客がいる限りはその客)を最優先でサービスすることになります。

2.4 柔軟な処理能力の運用

待ち行列モデルの基本的かつ重要な要素である処理能力とばらつきに話を戻しましょう。推奨される最後の一般的法則は、**処理能力を柔軟に運用すること**です。よく考えられた実際にある例として、とても興味深いのは次のオランダのスーパーマーケットのチェーン店の事例です。

スーパーマーケットの事例

このスーパーマーケットチェーンのHoogvlietでは、来店する客数を定期的に数えていて、それによって開けるレジの数を調整しています。つまり、処理能力を変化させることによってサービス要求のばらつきを吸収しているわけです。これは柔軟性と呼ぶことができます。このスーパーマーケットでは客へのおまけとして、レジ待ちの行列が2人以下のレジがなかったら、客は代金を払わずにレジを通過してよいことになっています。このおまけで損する費用を差し引いても、この新しい仕組みは大成功しているようです。その理由はおまけそのものというより、客にとってあまり待たなくて済むようになったことにあります。

管理プロセスにおける並列の作業ユニットに対して処理能力を柔軟に運用することは、上の例と同じように効果的であることは明らかです。

2.5 一般的法則

もちろん、管理プロセスにおいて遅れが生じる原因としくみの全ての定量的側面をここで論じたわけで

はありませんが、これらの基本的原則を考えることによって、管理プロセスにおける遅れを減少させ総処理時間を短縮させるのに有効な一般的法則を導くことができます。これらの一般的法則を簡単に箇条書きにしてみましょう。

総処理時間を短縮するための一般的法則

- サービスの要求の発生のばらつき(あるいは任意性)を減少させる。
- サービス時間(処理時間)のばらつき(あるいは任意性)を減少させる。
- 可能な場合には処理能力を柔軟に運用するよう心がける。
- サービス時間がほぼ等しい仕事は処理能力を共同利用する。
- サービス時間の異なる仕事は処理能力を分割する。
- 独立した仕事は並行して行う。
- 関連のある仕事は組み合わせて行う。
- 優先順位をつける。

3. 事例

この節では、上で述べた一般的法則をあてはめ、その有効性をシミュレーションによって評価した実際の事例を紹介します。これらの法則は待ち行列理論を応用して得られた結果得られたものですから、ここで用いる方法は、解析結果を利用した複合シミュレーションとすることができます。

3.1 事例の説明

Lavobはオランダの中堅の保険会社です。この会社の主な仕事は、個人向けと企業等団体向けの生命保険や年金計画の提供、契約とその維持です。以後単に製品と呼ぶことにします。各製品は、処理の所要時間と複雑さの程度に応じておおまかに次の3つのカテゴリに分類されています。

カテゴリ1: コンピューターで処理できる標準的な製品

カテゴリ2: 手作業で行わなければならない特別な製品

カテゴリ3: 特別な専門的能力を必要とする非常に特殊な製品

約90%の製品が、カテゴリ1に属します。この製品としては個人向けの保険がほとんどで、製品の実際の提供はコンピューターが30分~1時間程度で行い、完成までに正味約1週間かかります。2番目のカテゴリ(約7%)に属するのは、主に中小企業の団体年金プランの提供であり、個別に対応しなければなりません。したがって、これらの製品は通常、特殊能力のある数人の従業員の手作業によって処理され、完成までにかかる処理時間は、1週間のものから数カ月のものまで様々です。3番目のカテゴリ(約3%)はカテゴリ2よりもさらに特殊なもので、大企業の団体年金プランの提供です。これは経営陣の承認が必要なおえに、特別な能力や専門的知識も必要とされます。このような能力を持った従業員の数は限られているので、製品が完成するまでに1年以上かかることもあります。しかし、カテゴリ2・3の製品からは実益が上がっていますので、どのカテゴリの製品もこの企業にとって非常に重要な製品です。カテゴリ1はキャッシュフローのしっかりとした基盤を約束するもので、カテゴリ2と、とりわけカテゴリ3は、純利益確保のために欠かせないものになっています。以上のことより、次のような目標がたてられました。

1. 3つのカテゴリ間の仕事量のバランスを良くすること
2. 非常に長くかかる処理時間を実質的に短くすること
3. すべての製品、とりわけカテゴリ1の製品にかかる総処理時間を減らすこと

またここでは、さらに次の2つの社会的、心理的条件を考慮することが必要でした。

- 北オランダや南オランダの献身的な従業員を指名するような客の好み
- 異なった専門知識または技術を身につけようとする従業員の希望

3.2 適用方法と評価

第2節で導かれた一般的法則を用いて、いくつかの異なった業務割り当ての形態を提案し、これらの形態の処理効率を大規模なシミュレーションを用いて評価

することによって、比較検討しました。これに先立って、様々な仕事の処理時間や需要分布をモデル化するために、統計的調査を行なっています。さらに全従業員が持っている専門的能力のリストも使用しました。種々の業務割り当ての形態を表1に簡単に示します。形態1はおおむねこの保険会社の従来の状況を表しています。

形態1と形態7についての解析結果を表2に示します。全ての製品の平均総処理時間は少なくとも20%短縮され、製品の大部分を占める製品1については、平

表 1: 業務割り当ての形態

業務形態	
1. 厳密な分割	3つのカテゴリーを厳密に区別する。従業員は各自、最も自分の能力を発揮できる仕事のみに専念する。これは、非効率的な処理能力の使用につながる。
2. 完全な共同利用	どの従業員も3つのカテゴリーの仕事を行う。
3. 優先順位をつけた共同利用	従業員はどのカテゴリーの仕事も行うが、自分の能力を最大限に発揮できる仕事を最優先で行う。
4. 業務内容による分業(並列)	製品カテゴリーではなく、業務内容によって3つのグループに分ける。
5. 客を満足させるための北オランダと南オランダの区別	3と同じ条件で、客が満足するように地域による区別を付加する。かなり処理能力が低下する。
6. 関連する仕事を統合する(組み合わせ)	たとえば、計算を担当する従業員は、必要な入力や計画書の作成も行う。これにより、内部的な遅れが回避される。
7. 4,5,6の組み合わせ	従業員の適性能力、製品の種類、業務の種類に応じた区別を行う。割り当てられた仕事がないときのみ、他の仕事にも従事する。同時に、南北オランダの区別も行う。この場合、処理能力の低下はあまり生じない。

表 2: 平均総処理時間

カテゴリー	1	2	3
形態1	6.24	9.03	646.15
形態7	5.05	5.05	5.05

均総処理時間が6日から5日に減少しました。とりわけ、完成まで非常に長くかかる製品がなくなったことは注目されます。この結果より、現実には1年以上に及んでいた非常に長い処理時間も、臨機応変な調整により回避できることがわかりますが、この結果はあくまでシミュレーションによって得られたものであることを確認しておきます。それでもなお、この結果は達成可能な改良点を十分に示していると言えるでしょう。

形態7を実施することは業務の根本的な再構築とみるべきで、運用の仕方や職場配置を徹底的に替えるのは当然として、従業員の意識をも思い切って改革する必要があります。したがって、この試みは1995年の始めから2年計画で、徐々に、しかるべき段階を踏んで実行するよう計画しました。

以上の結果をまとめると、管理プロセスのリエンジニアリングにとって、簡単な待ち行列理論に基づく結果、特に処理能力と処理時間のばらつきが果たす役割を認識することによって得るものは大きいことが分かります。

また簡単な待ち行列理論を使って、総処理時間を実質的に短縮させる一般性のある実用的法則を得ることができます。

これらの法則を実際に現場で運用するには、解析結果を利用した複合シミュレーションが必要になることが多いでしょう。

参考文献

- [1] van Dijk, N.M., (1992) "Why waiting always takes too long," *Volkskrant* (Dutch National Newspaper), 21 April, 1992.
- [2] van Dijk, N.M., (1994) "No more waiting", *HELPWARE Magazine*.
- [3] van Dijk, N.M., (1993) *Queueing Networks and Product Forms*, John Wiley and Sons, New York.
- [4] 高橋幸雄, (1995,1996) "やさしい待ち行列 (1)-(4)", *オペレーションズリサーチ*, 40.11, 40.12, 41.1, 41.2.