

データウェアハウスの応用研究

2AD-3

千葉幸司 東北情報ネットワークサービス(株)

宮内隆 宇部興産(株) 石井孝子 (株)ソーティス

中川裕道 三菱マテリアル(株) 杉崎信彦 富士通(株)

1. はじめに

データウェアハウス[1]（以下DWHと記す）の提唱者ビル・インモン氏は、その著書において構築や活用の具体的な方法論を述べていない。そこで、我々富士通LS研「データウェアハウスの構築と活用」分科会では、実在の企業をモデルに構築と活用の両面からインモン説DWHの実践的な検証を行った。モデル企業は、幅広い適用範囲の成果が得られるように、業態・前提条件の異なる流通業A社と製造業B社を選定した。この論文は、上記2社の現状と今後の方針をもとにしてDWHによる情報系システムの設計を行う過程で発生した様々な課題の中から主要なものについてまとめたものである。

2. 流通業A社におけるDWH適用

A社は、全国に20店舗を有する大型小売業である。同社は、近年の流通業界動向に先駆けて、メンバーズ・カードによる顧客データとPOS売上データとの連動から、パーソナル・マーケティングを目指している。しかしながら、現在のシステムは業務遂行を優先した構成・運用となっているため、情報系システムとして有効に機能していない。そこで、顧客主導型経営への展開の第一歩として、顧客サービス部と商品企画部の分析担当者を対象に、DWHによるPOSデータ分析システムを構築することとした。新システムへの要望を以下にまとめる。

- ・高精度の傾向分析や予測分析を実現するために、データを10年分保存して欲しい。
- ・使いやすく覚えやすい操作性、かつ、分析の思考を妨げないレスポンスを実現して欲しい。

上記の要望に沿ってシステム化を検討したとき、

A Practical Study of Building the Data Warehouse
Koji Chiba

Tohoku Information Network Service Co.,Inc.

付随する問題点も発生した。下記に問題点とその対策を述べる。

まず、10年という長期にわたりPOS分析データを時系列保存した場合、データとマスタとの間に不整合が生じると考えられた。対策としては、分析データに合わせて各マスタ情報も履歴保存することがあげられるが、単にマスタの履歴情報を保存しただけでは、店舗や商品の統廃合などがあった場合に、過去データを最新マスタ情報で読み替えて分析したいなどといったニーズに対応できない。そこで、マスタ情報の履歴保存に加えて、マスタ情報の変遷過程もメタデータとして保存することとした。これにより、いかなる分析においても整合性を保証できるDWHの実現を可能とした。

ところで、A社のDWHは将来の拡張性を考慮して汎用的な項目を多岐にわたって網羅するので、操作性やレスポンスの面では必ずしもユーザの要件を満たし得ないと考えられる。そこで、各ユーザ毎のニーズに特化したデータマート[1]（以下DMと記す）を切出すこととした。DMは一定の条件で抽出、要約して保管できるので、DWHに比してコンパクトなものとなり、レスポンスが格段に向上する。加えて、DBMSに多次元DBを採用すれば、多種多様かつスピーディな分析が提供可能となる。DMの設計にあたっては、DWHに対してR-OLAPツールを使用して仮説・検証の分析エミュレーションを繰り返し、必要とするユーザビューを精査、統合して実装する。この結果、A社では顧客サービス部と商品企画部の分析視点に明確な違いがあることがわかり、それぞれに最適なDMを導き出した。

3. 製造業B社におけるDWH適用

B社は、製造から販売までを一体化した製品別事業部制を敷いている。同社では、情報活用の観点か

ら昭和63年より「汎用情報検索システム」をメインフレーム上に構築し、営業部門のユーザを中心に、定型・定常業務の合理化・省力化の点で一定の成果をあげている。しかしながら、同社は、商品の競争力を高めて消費者のニーズをつかむために、現行システムをベースとしながらも、全ホワイトカラーの生産性を向上し、迅速で正確な意思決定を支援する、より優れた新情報系システムを求めるようになった。下記に新システムに対するニーズの一例を示す。

- ・売上や未納などの即時性が高く、かつ内容が時々刻々変化するデータを、リアルに近いタイミングで提供して欲しい。
- ・住宅着工状況（WWWに存在）などの外部データも充実して欲しい。
- ・集約から明細へドリルダウンする際に、テーブル選択などの作業を自動化・簡略化して欲しい。

上記のニーズの解決方法を、インモン説にもとづいて模索した結果、インモン説だけでは解決できない事象があり、下記のような検討を行った。

まず、B社の売上や未納といったデータは、オペレーショナル・データストア[1]（以下ODSと記す）に格納することとした。ODSは、インモン説ではDWHにデータを投入する前の統合化の手段とされているが、同社ではそれと異なる活用方法を見出したのである。ODSには業務系システムからトランザクション発生タイミングで、オンラインディレイド処理によりデータを格納し、DWH本体には業務系システム上の確定情報のみを夜間バッチで格納する。このようにすれば、即時性の高い不確定情報と、確定情報の双方が提供可能となる。

また、同社では社内のPCやサーバ上にあるデータや、インターネット上のHTMLファイルなども対象データとして扱う。そのため、新システムでは、ユーザの作成したデータはDMに、インターネット上のファイルはWWWサーバ上に格納したうえで、データ格納先の違いをユーザに意識させないように、アクセスにはWWW技術を生かしたWebブラウザによる検索を採用することとした。

最後に、メタデータの有効な活用方法について述

べる。もちろん、メタデータは構築時からテーブル設計に利用するが、ここでは、活用時における利用としてドリルダウン検索への適用を考察した。すなわち、検索されるデータに関する集計の手順を、メタデータとして蓄積するのである。具体的には、個々の集計データ項目について、集計条件とその集計元データが保存されている場所（テーブルやサーバ）を記述しておく。これにより、ユーザは項目を選択するだけで、ドリルダウン検索が可能となる。

4. まとめ

上記に代表される検討から、A社においては大量データの蓄積と特定のニーズにポイントを絞った効率のよいデータ分析によりパーソナルマーケティングを実現するシステムを、また、B社においては様々なニーズを大局的に整理して各部門が必要な情報を横断的に共有化するシステムを創案した。

その結果、インモン説が実践面でも有益であることや、概念レベルでは共通だったDWH像も導入する企業毎にかなり明確に形を変えたものになることが確認できた。さらに、日本の企業風土においては、インモン説をやや拡大解釈する必要のあることもわかった。このことは、実践面ではDWHに共通の理想像はなく、各企業それぞれが自社に最適な構築と活用の形態を探求していかなければならないことを意味する。言い換えれば、DWHはその構築と活用の仕方次第で、同業他社に差をつけるための強力な武器になり得るということである。

これからDWHの構築を予定している企業は、B社のような恵まれた環境にあるところよりも、A社のようなゼロから構築せねばならないところがほとんどではないだろうか。そのような企業において、この論文が、効果的なDWHを構築する一助となれば幸いである。

（参考文献）

- [1] W. H. Inmon & R. D. Hackathorn : よくわかるデータウェアハウス活用法, インターナショナル・トムソン・パブリッシング・ジャパン, 96年7月