

大型證券商 CRM 系統資料倉儲規劃與建置

Data Warehouse Planning and Construction for A CRM System of Large Scale Securities

曾守正*

Frank S.C. Tseng

國立高雄第一科技大學 資訊管理系

Dept. Information Management

National Kaohsiung First University of Science and Technology

Tel: 07-6011000 Ext. 4113

Fax: 07-6011042

E-mail: imfrank@ccms.nkfust.edu.tw

<http://www.nkfust.edu.tw/~imfrank>

摘 要

由於網際網路興起所帶動的資訊流通速度，使得現代化企業競爭的環境變化已經快到讓人難以喘息。在商業資料的應用領域方面，後續的挑戰將會是：「如何將以往的歷史資料歸納整理後，與外部相關的資料作一番完整的整合，以提供快速而準確的決策分析能力。」在這樣的需求之下，建置企業「資料倉儲」(Data Warehouses) 的觀念便應運而生。由於資料倉儲可以協助企業將所有過去的歷史營運資料與經由網際網路或產業公報上所收集到的精確資料做完美的整合，以提供多維度的線上資料分析，進一步輔助企業做出有效的決策，因此越來越受到企業的重視。而目前國內的金融證券業，其主要的業務是提供服務以獲得利潤。所以若要獲得更多的利潤，就要開發更廣大的客戶群。除此之外，如何了解客戶的組成結構，以及留住具有影響力的的大客戶，就成了證券業的當前要務。

國內的大型證券商爲了加強對客戶的服務品質，提升客戶對企業的忠誠度與向心力，目前正大力引進了「知識管理」(Knowledge Management) 的觀念，並結合目前在金融服務業普遍認同的「客戶關係管理」(Customer Relationship Management, CRM) 觀念，期待能透過完整的系統建置，讓該公司的服務品質在客戶口耳相傳的優良口碑中，永續維持國內證券業領導性地位。

雖然國內目前證券業希望建置 CRM 系統的公司相當多，但是在面對規模龐大的資料庫與資料倉儲規劃上，常常令業界人士產生無從下手的困境，尤其到目前爲止我們並沒有看到較完整的系統建置報告出現。因此，本論文將針對我們在協助國內某大型證券商規劃 CRM 系統資料庫與資料倉儲的過程提供說明，希望能提供整個系統規劃經驗的分享。相信以如此規模的大型券商做爲實務上的建置研究，當能提供想要建置資料倉儲系統的產、官、學界一個實務上的參考。

關鍵字：XML、異質性資料庫整合、倉料倉儲、倉料探勘、客戶關係管理。

* To whom all correspondence should be sent.

Data Warehouse Planning and Construction for A CRM System of Large Scale Securities

Frank S.C. Tseng*

Department of Information Management
National Kaohsiung First University of Science and Technology
Tel: 07-6011000 Ext. 4113, Fax: 07-6011042
E-mail: imfrank@ccms.nkfust.edu.tw
<http://www.nkfust.edu.tw/~imfrank>

Abstract

As the applications of World Wide Web in commercial activities grow tremendously, the next challenge of business data applications will be a process of integrating historical and external data to support the ability for efficient and accurate decision-making. Under such circumstances, it is indispensable to construct an enterprise-wide data warehouse. This is because the construction of data warehouses can help enterprises to integrate all historical operational data with industrial official announcement to provide multi-dimensional on-line analytical processing, which effectively promotes the quality of enterprise decision-making. In domestic financial circles, profits are gained mainly from customer services. Therefore, to increase the profit, it is undoubtedly that enterprises should develop new customers and decrease the churning rate more aggressively.

To upgrade the customer service quality, the concepts of knowledge management and customer relationship management (CRM) are adopted by many domestic securities. Hopefully, by establishing such a computer system, the quality of customer service will be improved and the leading position will be retained.

On constructing an integrated data warehouse, many companies are really frustrated due to the labor-intensive, time-consuming, and error-prone process. Therefore, up to now we have not seen a domestic technical report regarding the planning and construction of a financial enterprise data warehouse. In this paper, we intend to share our experience of planning and constructing a data warehouse for the CRM system of a large scale domestic securities. We believe this can provide a very practical research testbed for the domestic financial circles, the government, and the academy to reference.

Keywords: *XML, Heterogeneous Database Integration, Data Warehouse, Data Mining, Customer Relationship Management.*

* To whom all correspondence should be sent.

一、導 論

「資料庫管理系統」(Data Base Management Systems, 簡稱 DBMS) 以及架在其上所開發出來的各類型「管理資訊系統」(Management Information Systems), 在過去的幾十年間已經大量深入各行各業的商業應用中, 也大多能成功地協助企業妥善管理其資料, 並產出深具價值的各類資訊。但是, 正如 Kimball [12] 所指出: 過去受限於儲存空間的大小, 以及電腦對於大量磁碟運算無法提升速度的影響, 使得這些應用大多只侷限於每天的商業交易或是短期應用的運算, 很難加以進一步提昇, 例如: POS 系統 (Point of Sale Systems)、庫存管理系統、或會計資訊系統等。

近年來, 受惠於儲存空間的大幅提升與價格下跌, 還有電腦速度進一步提升與平行處理技術的互相搭配, 想要將過去大量累積的歷史資料帶入線上分析的應用, 以提供更精準而有效決策的夢想已經可以實現了。同時, 拜 Internet 蓬勃發展之勢, 這些線上分析甚至於可以在任何時間、任何地點透過網際網路來達成。網際網路的興起帶起了一片透過網路加值現有產業價值之風氣。在軟體業界所引發之風潮主導了電子商務、網路安全、數位認證、通路整合等的進一步發展。諸多在美國那斯達克 (NASDAQ) 上櫃之網路概念股如: AOL、Amazon、Yahoo 等公司所傳達的一項理念即為—未來之網路發展已無國界之分, 亦無產業之分, 未能充分應用網路科技之公司極有可能在不久的將來於此波工業革命衝擊之下消失。

在此一波網路革命中, 有人專注於網路軟、硬體技術之開發與傳播, 如: Microsoft、IBM、Cisco 等, 也有人專注於網路通路/電子商務之建立, 如: Amazon、eBay、eTrade 等, 或有人專注於電信與網路之服務, 如: AOL 與 Hinet 等。而進一步整合這些技術與通路所創造出來之網路與行動電子商務則是兵家必爭之地。但是, 今日網路行銷之客戶不但數量龐大, 而且大多數客戶很難加以區隔, 已經無法像過去經營傳統商店, 僅靠人際關係或簡易的促銷活動即可達成。雖然目前可以透過電子郵件帳號或其他簡易登錄訊息來進行行銷活動, 但是以大量廣告信所產生之效果, 實在很難估算。因此, 對目前所有具備電子商務營運的公司來看, 雖然現有網站技術已日趨成熟, 然而對現有客戶之管理與行銷模式的掌握則完全付之闕如。如何而能結合最新網站技術與加強現有客戶滿意度, 提高客戶忠誠度, 從而在了解現有客戶之消費習性與其他資訊之後, 擬定最佳行銷策略獲致公司最大之利益, 殊為現有電子商務之盲點。

1. 現代化企業與資料倉儲

現代化企業競爭的環境變化已經快到讓人難以喘息，因此在商業資料的應用領域方面，我們認為後續的挑戰將會是：「如何將以往的歷史資料歸納整理後，與外部相關的資料作一番完整的整合，以提供快速而準確的決策分析能力。」在這樣的需求之下，「資料倉儲」(Data Warehouses) 的觀念便應運而生，而且越來越受到企業的重視。其原因不外乎：它可以協助企業將所有過去的歷史營運資料與透過網際網路、或產業公報上所收集到的精確資料做完美的整合，提供多維度的線上資料分析，進一步輔助企業做出有效的決策 [12] [13] 。

「資料倉儲」的觀念首先是由 W.H. Inmon [11] 所提出來的。他強調資料倉儲是一個「主題導向」(Subject-Oriented)、 「完全整合」(Integrated)、 「隨時間變動」(Time-Variant)、 以及「非暫存性」(Non-Volatile) 的資料集合所構成，用於提供企業在決策過程的輔助支援。成功的資料倉儲建置可以協助企業找出以前未知的，或是無法很容易取得的資訊，因此可以幫助企業更了解內部的狀態變化、客戶的需求、經營趨勢，以及發現過去的缺失等。根據 IDC (International Data Corporation) 1996 年的研究報告指出 [6]：過去成功建置資料倉儲的企業當中，其三年的平均 ROI (Return on Investment) 高達 401%，其中有 15% 超過 600%、50% 超過 160%，高達 90% 超過 40%。可見企業建置資料倉儲所可能產生的優勢的確是相當驚人的。

目前國內的金融證券業，其主要的業務是提供服務以獲得利潤，例如：幫助客戶完成證券買賣的交易以賺取手續費。所以若要獲得更多的利潤，就要擁有廣大的客戶群。根據行銷學上的 80/20 法則：銷售額的百分之八十是由消費排名前百分之二十的顧客所達成的。加上 Peppers & Rogers [16] 的研究指出：開發新顧客的成本要比保有舊顧客的成本高出六到九倍。因此如何留住這些具有影響力的的大戶，以及開發新的大戶，就成了證券業的當前要務。而留住這些客戶、開發新大戶的不二法門，就是提供更好的服務。

2. 證券商的 CRM 系統設計

本論文將探討國內某大型證券與金融集團的客戶關係管理系統設計與規劃，該集團旗下主要分為四個體系：商業銀行體系、投資銀行體系、周邊事業體系，以及公益基金會。而其中又以投資銀行體系所涵蓋的子公司最多，包含有證券、投信、期貨、投顧、創投、國際資產管理、網路證券以及四家海外分公司或辦事處

等。該證券機構成立甚早，也相當具代表性。早在 1987 年政府解除券商管制以來，新券商如雨後春筍般出現之前，該公司便已具備豐富的營運經驗，而且頗富信譽。自上櫃以來，經紀、承銷、債券、自營、衍生商品、國外等各項業務皆穩定均衡成長。2000 年 6 月底時，其淨值已經高達新台幣 178.8 億元（約合美金 \$5.77 億元），總資產達 492.3 億元新台幣。如果以當時淨值總額計算，該證券商可以名列台灣第二大券商。因此，我們覺得以如此規模的業界廠商來討論其客戶關係管理系統的規劃過程，當可以提供國內各金融服務業一個可供參考的案例。

目前國內外 CRM 的廠商有：Siebel、PeopleSoft、Clarify、J.D. Edwards 等大廠。由於電子商務的市場相當大，加上 CRM 的觀念正在轉型：已逐漸由過去產品導向之 SFA (Sales Force Automation) 轉為以客戶為導向之真正 CRM 系統。所以許多資料庫軟體廠商：如：IBM, Oracle, Sybase 等，也紛紛加入此一戰役，搶佔市場。

有鑑於此，該集團在 2000 年底正式由資訊處召開會議並充分討論了整體企業在內、外環境的各種競爭優、劣勢，以及其中所潛藏的機會與威脅之後，整理出表 1 的 SWOT 分析結論 [8]。最後，會議決議邀請國內、外各知名大廠提供企業知識管理解決方案的簡報。但是經過深入了解，並與國內學者專家及業者多次交流，經過多方審慎評估之後，考量到國外顧問對國內企業文化與運作模式間的落差，往往會產成效率不彰的問題。因此，決定採用本土化公司某科技公司（以下簡稱 N 科技公司）的「知識工作者系統」(Knowledge Worker Portal System) 知識管理解決方案，在 Microsoft Windows 2000 Server 系列的平台上規劃並建置整個集團的知識管理資訊架構，並在 2001 年底進行導入，開始進行整合測試。

N 科技公司所提供的知識工作者系統 (Knowledge Worker Portal System)，其知識管理的對象主要分成以下三類：

1. 隱藏在人與人之間 (People-to-People) 的「隱含知識」(Tacit Knowledge)：透過人與人之間彼此的協調 (Collaboration) 運作來產生並加以管理。整個軟體平臺是採用了微軟的 Exchange Server 2000 與 Conferencing Server 2000 提供證券商內部社群、工作小組與專家顧問間的即時溝通 (Real Time Collaboration) 機制。

表 1 (a)：證券商 SWOT 分析 (內部環境部分)

內部環境	
競爭優勢 (Strength)	競爭弱勢 (Weakness)
<p>制度：</p> <ul style="list-style-type: none"> ◆ 帳務系統內部專案設計及系統標準制度文件均已建立完成，有利於傳承。 ◆ 雙軌制薪資已漸形成。 <p>人力：</p> <ul style="list-style-type: none"> ◆ 金融帳務流程設計之實務與技術 (Tandem 技術) 兼具之人材充裕。 <p>服務、產品：</p> <ul style="list-style-type: none"> ◆ 與國外軟體公司合作擁有較同業先進之網際網路系統架構及高連動性、低故障糾紛、高資源共享的網路下單系統。 ◆ 電子商務系統架構及下單速度、即時 Information (帳戶管理、即時行情)、8 條 Internet T1 頻寬、Internet T1 頻寬、均屬同業之冠。 ◆ 客戶庫存成本損益領先競爭對手。 ◆ e-Agent 700 位營業員 E 化之客服系統已經全面上線完成，技術整合門檻高，同業不易追上。 ◆ 客戶介面 (電話、現場、網際網路、語音按鍵) 完全透通。 ◆ 針對庫藏股及員工獎勵方案之員工選擇權雛型系統已經完成。 	<p>制度：</p> <ul style="list-style-type: none"> ◆ 現行最新技術之開發能力及內部程式開發標準均有待加強。 ◆ 「管理」、「開發」、「研發」、「規劃」等高速競爭組織機能，尚屬形成期。 <p>人力：</p> <ul style="list-style-type: none"> ◆ 網際網路管理及應用開發，規模量產人力未全面普及。 ◆ 資料倉儲、知識管理等開發人才欠缺。 ◆ 資訊基礎環境規劃人才欠缺 (安全「防駭、防毒、災害復原」、網路「Internet、Intranet、Extranet」、作業環境「效能調整、資料庫管理」)。 <p>服務、產品：</p> <ul style="list-style-type: none"> ◆ 網站商品交易功能欠缺。 ◆ 即時管理決策、幕僚情報系統尚未建立。 ◆ 集團資訊流通、知識分享、流程管理欠缺資訊系統以提昇行政效率。 ◆ 行銷分析、銷售執行、客戶服務等三項循環處理流程，系統尚未建置完成。

表 1 (b)：證券商 SWOT 分析 (外部環境部分)

外部環境	
外在機會 (Opportunity)	外在威脅 (Threat)
<p>機器設備：</p> <ul style="list-style-type: none"> ◆ PC Server 主機擴充能力提昇，可提供至 32 顆 CPU，有利 Windows 2000 用戶主機規劃。 <p>服務、產品：</p> <ul style="list-style-type: none"> ◆ 微軟新產品上市有助於企業使用較低的成本發展資料倉儲、OA、知識管理。 ◆ 無線上網業務興起 (第三代行動電話)，IA、手機 (PDA) 下單。 ◆ 代客操作法規開放，有助於資產管理業務成長，並建立起台資海外投資法人機構。 ◆ 電子簽章法、寬頻網路及家用電腦普及利於電子商務市場普遍發展。 ◆ 台灣投資風險意識提高，利於「國外投資品」、「套利商品」、「避險操作」、「固定收益」、「期貨商品」商品興起。 ◆ 購併券商成本較低。 	<p>災害處理：</p> <ul style="list-style-type: none"> ◆ 聖嬰現象天候異常與台灣天然災害，考驗企業本身災害復原能力。 <p>人力：</p> <ul style="list-style-type: none"> ◆ 資訊技術應用商務興起，技術人才高度競爭。 <p>機器設備：</p> <ul style="list-style-type: none"> ◆ 資訊產品更新快速已延伸至大型主機設備，Tandem 投資之設備更新加劇，設備使用延續性低、易增加成本。 <p>服務、產品：</p> <ul style="list-style-type: none"> ◆ WTO 開放國外券商加入競爭、國內券商價格競爭，商品設計及行銷能力提昇益形重要。 ◆ 投資風險意識提高，客戶交易意願降低，不利傳統經記業務。 ◆ 同業對電子商務及資料倉儲發展的共識性極高，應用系統開發速度競爭激烈，委外設計比率增加。

2. 隱藏在人與文件資訊 (People-to-Information) 之間的「明確知識」(Explicit Knowledge)：透過人對於文件與資訊內容 (Content Management) 的互動模式來加以管理。整個軟體平臺採用了 Microsoft 的 Exchange Server 2000 搭配 Web Storage System 提供內部文件的內容管理 (Content Management)，以及企業入口網站與搜尋 (Portals and Search) 機制。
3. 隱藏在人與資料 (People-to-Data) 之間的「潛在知識」(Potential Knowledge)：透過企業資料倉儲的建置來加以應用與管理。軟體平臺採用的是 Microsoft 的 MS SQL Server 2000 搭配其 Analysis Service 提供結構化資料整合，以及資料倉儲建置。

本論文所專注探討的也正是上述的第三項：潛在知識管理中的資料倉儲建置工作，並在初期專注於客戶關係管理的應用部分。該系統的涵蓋範圍如表 2 所示，其中的 E-Agent 與 E-Trader 乃該證券商自行開發的應用程式，都是用於對客戶的服務方面，陰影部分則是本論文所討論的重點。

表 2：證券商 CRM 系統的涵蓋範圍

	互動面	營運面			決策分析面
		服務	銷售	行銷	
內容	<ul style="list-style-type: none"> ◆ Call Center ◆ Web-Centric Service ◆ E-mail 	<ul style="list-style-type: none"> ◆ 客戶互動記錄 ◆ 客戶交易記錄 ◆ 問題解決 	<ul style="list-style-type: none"> ◆ 產品目錄 ◆ 銷售術語 ◆ 篩選銷售對象 ◆ 佣金計算 	<ul style="list-style-type: none"> ◆ 行銷專案規劃 ◆ 行銷成效檢討 	<ul style="list-style-type: none"> ◆ 客戶貢獻度、喜好 ◆ 產品貢獻度 ◆ 營業員貢獻度 ◆ 行銷效益分析
解決方案	<ul style="list-style-type: none"> ◆ Web 網站 ◆ E-Trader 	<ul style="list-style-type: none"> ◆ E-Agent 	<ul style="list-style-type: none"> ◆ 知識入口網站 ◆ 研究報告上網 ◆ 業務手冊上網 ◆ 銷售執行結果 	<ul style="list-style-type: none"> ◆ 知識入口網站 ◆ Web 化專案管理 	<ul style="list-style-type: none"> ◆ 資料倉儲與線上分析

N 科技公司透過與本校的創新育成中心 (Innovation and Incubator Center) 合作，由我們負責設計與規劃整個資料倉儲的系統架構，以及資料庫端使用者自訂函數 (User-Defined Functions)、預儲程序 (Stored Procedures) 與觸發程序 (Triggers) 的撰寫，並由該公司的工程人員進行前端程式的開發，整個規劃過程歷時約一年，程式開發部分則約花費十二個月。本文的目的在於提供整個規劃過程的經驗分

享，相信以如此規模的大型券商做為實務上的建置研究，當能提供想要建置資料倉儲系統的產、官、學界一個實務上的參考。

3. 客戶關係管理簡介

「客戶關係管理」(Customer Relationship Management, CRM) 的觀念其實早已在商場上行之有年，只是早期所仰賴的是業務員與客戶間的人際關係維繫，並未大量透過電腦的輔助來加以管理。在 1980 年代初期則有初步形成了「接觸管理」(Contact Management) 的觀念，專門收集客戶與公司聯繫的相關資訊，並加以電腦化；到了 1990 年代初期又演變成爲包括電話服務中心 (Call Center) 與支援資料分析的客戶服務功能 (Customer Care)，如：Ohaegbu & Devgan [15] 所述。直到最近網際網路帶動了「電子商務」(E-Commerce) 的興起，讓客戶關係管理的觀念已經與網際網路結合，服務目標也延伸到了大量而難以掌握的網路下單 e 世代客戶群。

而客戶關係管理乃是企業爲了鞏固現有客戶、吸引新的客戶群，並增進客戶對企業利潤的貢獻，所採取的一連串足以服務客戶並影響客戶消費行爲的方法。其目的主要有下列幾點：

1. 找出對企業貢獻最多的優良客戶，建立長遠的服務關係。
2. 藉由每一個能接觸到客戶的機會，讓客戶感受到更貼新的服務。
3. 藉由客戶的利潤貢獻，提高整個公司的市場佔有率。

所以，如果能得知現有客戶之消費習性、需求，就可以根據這些資訊擬定最佳行銷策略，配合最新網站技術來提供更好的服務，使得客戶“Come for the product, stay for the service.”，爲公司帶來更多利潤，這才是客戶關係管理的最終目的。

本論文所討論的系統建置目標是希望將所有客戶在證券、期貨、債券、基金四大項業務上的資料進行完全的整合。不過，鑒於國外資料倉儲建置經驗上所提出的報告顯示 [19]：由於整體企業資料倉儲的建置時程往往因異質性資料庫的整合時間拖得太長，平均需要花費三年以上 [6]，導致效果大打折扣。所以爲了不讓專案因時程拖累，因此經過評估之後，第一期的目標是先專注在比較複雜的證券業務上建置「資料市集」(Data Mart) [12]。

本論文的組織方式如下：第貳節簡介資料倉儲系統，以及它與資料庫系統的差

別，並說明一個完整的資料倉儲架構為何；第參節將以前述的架構為主軸，探討證券商 CRM 系統的架構與規劃細節；第肆節說明整個 CRM 系統資料倉儲中所包含的資料庫內容與其分類；第伍節說明我們所建置的資料方塊 (Data Cube) 以便做線上資料分析之用，最後我們提出結論並探討後續的發展空間。

二、資料倉儲系統簡介

1. 資料庫與資料倉儲系統的差別

「資料」(Data) 對企業來說是非常重要的資產，所以傳統在資料庫的整體規劃與設計上，重點工作就是決定資料的儲存與管理方式。其考慮的重點在於：快速存取資料的結構與技巧，並且要能保護資料異動的一致性，以方便管理。

好的資料庫設計可以將整體資料的重覆性 (Redundancy) 降低，以消除資料更新時的異常現象 (Update Anomalies)，並提高整體的運作效率。所以，資料庫系統的目標在於：建立一個高效率的操作型資料庫 (Operational Database)，讓我們可以很容易地建立各式各樣的資料，以及快速地查詢、統計資料。而一般企業成功的關鍵之一也正是妥善運用這類高效率的操作型資料庫。這類資料庫系統基本上常常會有異動的情形產生，因此可稱作「線上資料異動處理系統」(On-Line Transaction Processing Systems, OLTP Systems) [12]。

不過，在資料庫中存放資料、管理資料，還有一個重要目的就是希望用現有及歷史的資料去產生一些能協助我們做決策的資訊。但是，長久以來，在發展決策支援系統的過程中，我們發現由於所面對的問題與所需的資訊相當複雜，以致於線上資料異動處理系統 (OLTP Systems)，爲了大量的異動處理，而將資料儲存方式最佳化 (也就是關聯表的正規化) 的結構，已經無法讓我們可以容易地擷取各式各樣的資訊了。例如：有時候只爲了要調閱某筆客戶與訂購的資料，便需要將三、四個關聯表一起合併才能如願。所以常常可以聽到有資料庫的使用者抱怨說道：

1. 「我每天都有從各個構面 (從時間、物品類別等構面) 去擷取資料的需求，但是現有的資料庫結構讓我不曉得要如何下這些查詢。」

2. 「我想同時比較去年度與今年度，台灣北、中、南、東部的銷售資料在男、女方面的差異與對比情形，但是卻要花很多時間才能將這些統計資料組織起來。」

上面的敘述並非意味著：以往的 OLTP 系統有設計上的錯誤，而應該是說：OLTP 系統存在的目的是用來生產與收集外部資料的。同時它也點出了一項事實，那就是：OLTP 系統的設計並不適合用於資料分析上。因此，經驗的累積讓研究人員們發現：我們必須另行組織所要分析的資料，因為：“適合高效率線上處理的資料結構並不見得適合做大量資料分析與決策”。於是，出現了「資料倉儲」(Data Warehouse) 的概念，此一概念希望將我們想分析的各種資料切面 (Slice) 做一番前置處理後，變成很容易產生各類資訊的型式，然後擺放在系統中，供未來下決策時取用，就好像倉庫中放置各類成品一樣，可以讓我們隨時取用或是加以包裝出售。這類由資料倉儲上去分析資料的系統，現在已經是決策支援系統的一部份了，它們在許多應用上比較少有異動的情形產生，所以相對於前述的「線上資料異動處理系統」(On-Line Transaction Processing Systems, OLTP Systems) 來說，我們可以稱之為「線上資料分析系統」(On-Line Analytical Processing Systems, OLAP Systems) [12]。

一般來說，商業上決策支援 (Decision Support) 的目的，是希望能在策略的規劃上產生各種可行方案 (Alternatives)，以做為制定決策的依據。其做法則是利用現有商業資料 (Business Data) 配合過去的商業模式 (Business Model) 來產生可行方案。所以，決策支援的有效與否？是取決於「資料品質」(Quality of Data) 的好壞。因此我們常常會聽到電腦系統中有所謂的「垃圾進，垃圾出」(Garbage In, Garbage Out) 的說法。因此，在資料倉儲的建置過程中有賴於高品質的資料來源，方能分析出有效的決策。但是，高品質的資料往往需要耗費大量人力與時間來將資料庫中的缺項或雜亂資料做一番修正與整理方能達成，雖然有許多科學化的方法提出 (詳見 [17])，但是根據 [17] 所提及的經驗顯示：這個過程所佔用的時間最大可能超過整個專案時程的 80%。

2. 資料倉儲的一般架構

對於資料倉儲的系統架構，Anahory & Murray [1] 曾經提出如圖 1 的完整架構，

這個架構包含了資料倉儲的相關程序、工具與技術，可以說相當完整，因此普遍受到認同。後面我們將根據這樣的架構來設計該證券商 CRM 系統的資料倉儲。

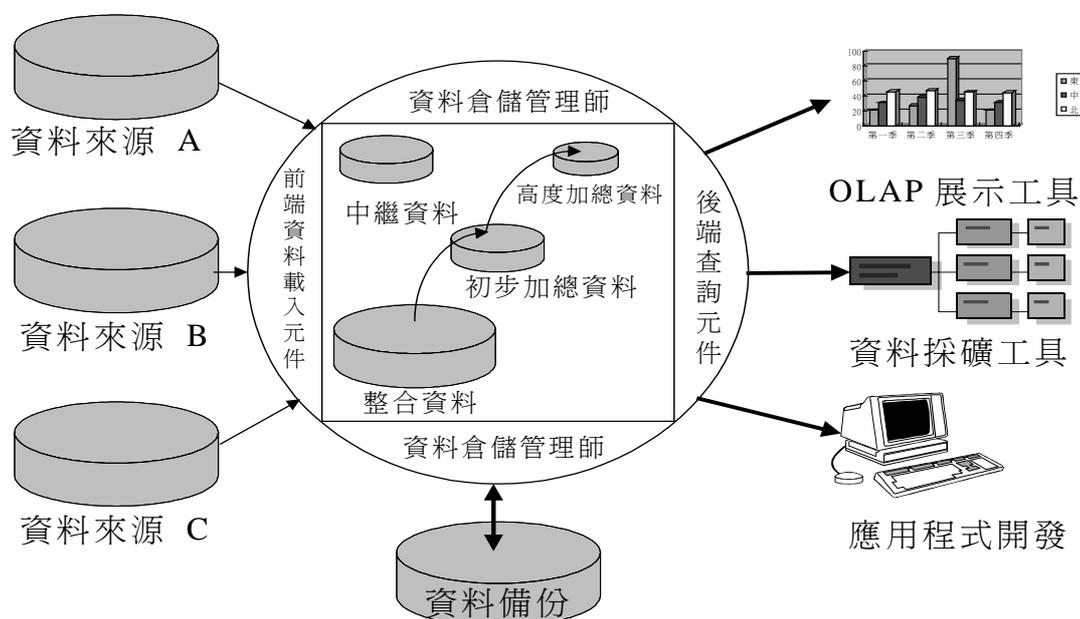


圖 1：資料倉儲的系統架構

整個架構具有多個異質性資料來源，經由「前端資料載入元件」(Front-End Load Component) 轉換與整合後產生整合資料 (Detailed Data)、中繼資料 (Meta-Data)，以及加總資料 (Summarized Data)，並由「資料倉儲管理師」(Data Warehouse Manager) 負責管理與備份。最後由「後端查詢元件」(Back-End Query Component) 負責輸出資料到「線上分析工具」(OLAP Tools) 或「資料採礦工具」(Data Mining Tools) 中進行分析與展示，另外也可以透過應用程式開發更完善的決策支援系統。

在前述的系統架構中總共包含了幾類資料，我們說明如下：

1. 運算資料 (Operational Data)：通常是以「線上資料異動處理系統」(On-Line Transaction Processing Systems, OLTP Systems) 所處理過的資料所組成，放置在大型系統的資料庫中，有些甚至可能是放在老舊的階層式資料庫 (Hierarchical Database) 中，或是專屬的檔案系統 (Propriety File System) 中，如：VSAM 或 ISAM File。當然有的資料也可以直接由 Internet 上的公開網頁上取得。
2. 整合後的詳細資料 (Integrated Detail Data)：將各式各樣所需的運算資料整

合、轉換而成整合後的詳細資料，我們加上‘詳細’一詞的用意乃在於：通常我們在資料倉儲中都會存放各類巨細靡遺的詳細資料，以利我們做維度分析之用。這些整合後的詳細資料也是放在資料庫中，而且會一直持續累積。

3. 預先加總的資料 (Summarized Data)：由於資料倉儲中的資料量很大，所以為了查詢效能的提升，通常會對某些常查詢的統計資料做預先加總的動作。這些資料必須在每次轉入新的運算資料時，跟著更新，維持加總的正確性。
4. 中繼資料 (Meta-Data)：前述的整合資料必須配合能夠用來描述它們的資料，這些「描述資料的資料」(Data about Data) 就稱為「中繼資料」。
5. 備份資料 (Backup Data)：適時備份整合後的詳細資料是有必要的，可以避免資料損壞所造成的損失。

在這樣的架構之下，Hackathorn [7] 也曾針對其中的資料流動方式，定義了以下的五種資訊流 (Information Flow)，我們配合圖 2 中的虛線箭頭部分說明如下：

1. 「輸入資訊流」(Inflow)：代表著由來源資料庫中擷取 (Extraction)、淨化 (Cleansing)，以及載入 (Loading) 到資料倉儲的整個過程。
2. 「向上統計資訊流」(Upflow)：代表著由資料倉儲中的資料進行加總、合併以產生加總資料的過程。
3. 「向下備份資訊流」(Downflow)：代表對資料倉儲中的資料進行備份的過程。
4. 「輸出資訊流」(Outflow)：代表將資料展示給使用者的過程。
5. 「中繼資訊流」(Meta-flow)：代表管理整個中繼資料 (Metadata) 的過程。

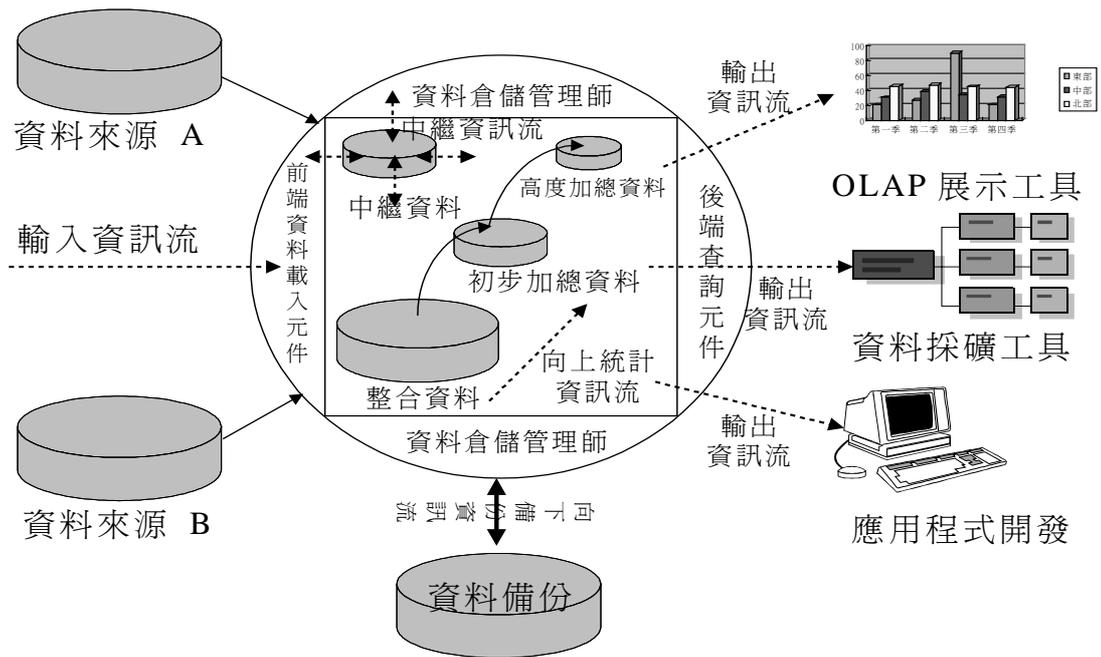


圖 2：資料倉儲系統架構中的五種資訊流

接著，我們說明該證券商目前的線上交易系統架構，以及我們所規劃的 CRM 系統架構。

三、證券商 CRM 系統架構

1. 證券商的線上交易系統架構

該證券商目前的線上交易系統仍然是 Legacy System，所使用的是 Tandem Non-Stop System，其所使用的資料庫是以「階層式資料模式」(Hierarchical Data Model) 為基礎的 VSAM 專屬檔案格式資料庫管理系統。由於該券商全部採用 Microsoft 的整體解決方案，為了讓資料能順利轉到 MS SQL Server 2000 上運作，該公司購買了一支應用程式用來將 Tandem 上的資料轉換成 MS SQL Server 2000 所能讀取的格式，並以每天自動轉換的方式下載到 SQL Server 2000 中。我們會在第肆節說明整個資料庫的綱要內容。

2. CRM 系統架構

前一小節中的資料庫格式適合線上交易處理 (On-Line Transaction Processing, OLTP)，但是不見得適合線上分析處理 (On-Line Analytical Processing, OLAP) 之

用。因此，我們另外設計了 CRM 系統的資料庫綱要 (Database Schema)，以利後續資料方塊 (Data Cubes) 的建置。

從資料可否異動的角度來看，我們將 CRM 系統中的資料庫分成以下幾大類：

1. 不可更新、刪除與修改的部分：這些資料是由 Tandem 端的線上交易資料所轉入，必須保持原始交易的原貌，不得任意更改。這些資料包含了：原始的交易資料、統計資料等。
2. 可以人工更新、刪除與修改的部分：這些資料是由前述的基本資料，以及我們為了分析方便而在 CRM 系統另行加入的新增基本資料，我們也提供了 Web 的介面讓該證券商的內部使用者可以對它進行修改以便維護。這些資料包括了：郵遞區號列表、各種下單方式的原始資料、各種交易金額級距與分類、證券商的基本資料、營業員的基本資料 (如：進入公司日期、所屬分公司)、自然人客戶的基本資料 (如：地址、電話、性別、教育程度)、法人客戶的基本資料 (如：統一編號、電話、資本額) 等。
3. 由 CRM 系統進行更新、刪除與修改的部分：這些資料是依照客戶的交易資料，隨著交易時間的移動而對客戶的偏好類股、偏好個股，以及風險偏好進行統計與分析而得。由與這些資料乃動態由交易資料中所計算，所以是由系統計算結果後填入資料，不可以由使用者來進行內容的更動。

整個 CRM 系統的資料庫架構如圖 3 所示，除了上述的資料分類以外，另外也建置了許多的視界 (Views) 與使用者自訂函數 (User-Defined Functions)。整個 CRM 資料庫綱要中共含有 242 個自定資料型態、179 個關聯表、38 個視界、18 個使用者自訂函數，以及 35 個預儲程序等。

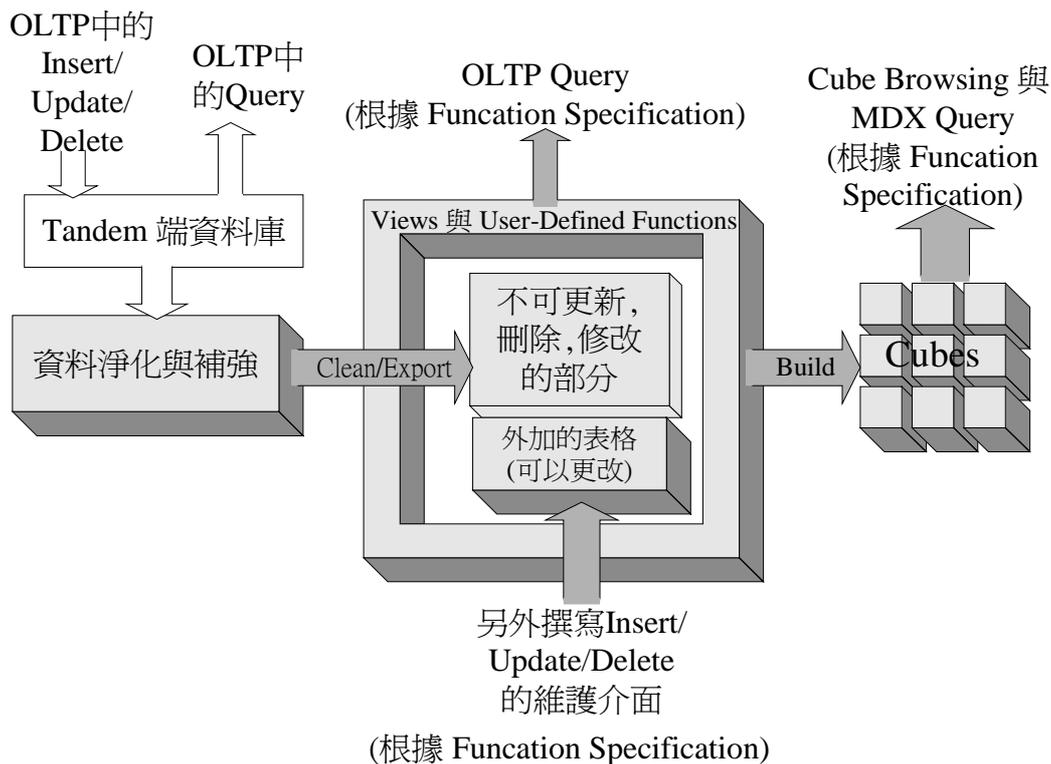


圖 3：CRM 系統的簡化架構圖

另外，由於在轉換的過程中，系統如果發現欲轉入的資料值組 (Tuple) 之主鍵 (Primary Key) 已經存在時，往往會發出錯誤訊息並停止轉換動作。為了防止此一情況發生，我們也各自替所有的基本資料表格都撰寫了一個觸發程序 (Trigger)，其目的是希望在轉換過程中，如果發現欲轉入的資料值組 (Tuple) 之主鍵 (Primary Key) 已經存在時，那麼就將原本要新增 (Insert) 的動作改成以更新 (Update) 動作來替代。而由於此動作必須在新增之前就要檢查，雖然 SQL Server 2000 並沒有支援 Before Trigger (在異動真正執行前就會先執行的觸發程序) 的功能，但是它提供了有 INSTEAD OF Trigger 的功能，可以用來實現此一作用。所以我們利用 SQL Server 2000 新提供的 INSTEAD OF Trigger 來加以實現。舉例來說，假設有一個 $Accounts(key, attribute_1, attribute_2, \dots, attribute_n)$ 關聯表代表證券帳戶，其中 Accounts 是關聯表名稱、key 則是其主鍵，而 $attribute_1, attribute_2, \dots, attribute_n$ 則是其他非主鍵的屬性 (Non-Key Attributes)，則其上的觸發程序如下列程式所示：

```
create trigger instead_of_INSERT_Accounts on Accounts instead of INSERT
as
if not exists(select B.key from Accounts B, inserted I where B.key = I.key)
insert into Accounts select * from inserted
```

```

else
  update Accounts set attribute1 = I.Attribute1, Attribute2 = I.Attribute2,
    ..., attributen = I.attributen
  from Accounts B, inserted I
  where B.key = I.key

```

由於這些觸發程序的邏輯都一樣，再加上關聯表相當多，因此，我們是另外撰寫一支程式來自動讀取資料庫中的「系統目錄」(System Catalog)，以取得所有表格的名稱、主鍵、還有其他屬性資料後，自動產生這些觸發程序。所以，任何關聯表後續如果有任何欄位上的更動，該程式會自動產生新的觸發程序，大大降低了我們在維護上的工作。

接下來，我們以圖 3 的資料倉儲系統架構為基礎，一一說明整個證券商 CRM 系統中五個資訊流的設計與規劃方式。

(一) CRM 系統的輸入資訊流

由來源資料庫中擷取 (Extraction)、淨化 (Cleansing)，以及載入 (Loading) 到資料倉儲的整個過程可以分成兩個過程來說明：

1. 第一次轉換過程 (One-Time Transformation)：第一次將所有 Tandem 端所存放的歷史資料 (但是交易的歷史資料只轉三年) 轉換到 MS SQL Server 2000 中，由於第一次轉換動作只做一次，所以轉換時間的需求上比較有彈性，並沒有規定一定要在幾天內轉完。
2. 往後每天的轉換過程 (Daily Transformation)：往後每一個交易日所產生的新客戶、新帳戶、新股票、新營業據點、新營業員、新交易資料等通通要轉到 MS SQL Server 2000 中，並重新調整所有的加總資料，所有的計算工作與資料方塊重建，都必須在晚上七點到隔天早上七點以前轉換與計算完畢。因為它只是做漸進式的轉換 (Incremental Transformation) 而已，所以這個工作的計算量比起第一次的要少很多。

為了降低後續維護上的工作，我們一開始便設定這兩個程式模組必須要具備完全一樣的邏輯，也就是說兩者的系統架構與程式內容必須完全一樣，兩者的差別僅在於所計算的對象資料不同而已。如此一來，即使因為調整效能而做的任何變更，也都只需改變一份即可，大大降低了我們在維護上的工作。

我們透過 MS SQL Server 2000 的 Data Transformation Service (DTS) 總共撰寫了 22 個 DTS Packages，分別進行：

1. Tandem 到 MS SQL Server 2000 的資料庫轉換：由於 Tandem 端的資料超過 16 Gigabytes，第一次轉換到 CRM 資料庫後加上以下所需的初步加總資料與高度加總資料後會變成 24 Gigabytes。因此，這個部分大概是佔最多時間的部分，如果只使用 Data Transformation Service 來實現時，由於有許多表格需要執行前述的觸發程序，因此，在八顆 CPU「對稱式多重處理架構」(Symmetric Multi-Processing) 的 Compaq Proliant 伺服器運算下，配合 EMC (<http://www.emc.com>) 的 Disk Storage Device，第一次的轉換時間仍然要超過五天整 (120 小時以上)。稍後我們將最大的關聯表「客戶歷史帳卡」改以 C++ 程式撰寫多重執行緒 (Multi-Threading) 程式，竭盡全力將八顆 CPUs 用到極限，並配合透過呼叫 MS SQL Server 2000 所提供的 Dos 指令 bcp (bulkcopy 的簡稱，為一提供大量資料匯入與匯出的工具程式)，進行更快速的轉換，結果竟然可以縮短到兩天內 (47 個小時多)。
2. 初步加總計算：包括計算客戶的偏好類股與偏好個股、貢獻度、目前持股庫存等；還有營業員的績效、分公司的績效；分公司的股票餘額與信用收付資料等。第一次的計算時間大約要花上 8 個小時。
3. 高度加總計算：將台灣證券交易所轉來的資料與上述的加總資料進一步彙整後產生更高度的加總統計資料。例如：客戶透過網路下單的金額與比率、各個分公司每日交易金額統計與所有券商每日交易金額統計資料等。第一次的計算時間大約要花上 5 個小時。
4. 建立大型表格的常用索引：由於整個資料倉儲建置完成後，尚有許多的需求規格是屬於「線上交易處理」(On-Line Transaction Processing, OLTP) 方面的查詢，因此為了效能的提升我們挑選了幾個常用而且內容超過三百萬筆的關聯表進行常用索引建立，並置於 Data Transformation Service 的 Package 中，以加快查詢速度，經過我們實驗的結果，建立索引後的查詢效能約比尚未建索引前平均要快上 50 幾倍以上，效果相當顯著。
5. 處理所有資料方塊 (Data Cubes) 所用到的維度 (Dimensions)：由於大部分的維度都很小，只有幾個如：「客戶結構維度」(因為帳戶的數目超過 730,000 個，同時每天平均以近兩百個帳戶的數量增加)，以及「各個區域

劃分與券商關係」的大型維度我們挑出來採用「漸進式更新」(Incremental Update) 的方式做重新計算之外，其餘我們都採用完整處理 (Full Process) 的方式來完成。

6. 處理所有的資料方塊：雖然 MS SQL Server 的 Analysis Services 提供了 Relational OLAP (簡稱 ROLAP)、Multi-Dimensional OLAP (簡稱 MOLAP)、以及 Hybrid OLAP (HOLAP) (見 [18]) 三種模式可供選擇，但是我們爲了執行效率上的考量，決定全部採用 MOLAP 方式處理。

(二) CRM 系統的向上統計資訊流

整個 CRM 系統的初步加總資料與高度加總資料是放在前小節中所述的 Data Transformation Service 中來計算。另外，有一些則是放在資料方塊中來做多維度的加總與計算。

(三) CRM 系統的向下備份資訊流

整個 CRM 系統的資料保持三年的交易資料，但是全部的基本資料則完全保留在資料倉儲中，所以我們提供了一個備份功能在每年的年底便可以將三年前的資料匯出，並予以備份。當然所有初步加總資料與高度加總資料則要配合重新計算。

(四) CRM 系統的輸出資訊流

在 OLAP 分析結果輸出方面，我們採用的是 Arcplan 公司 (<http://www.acrplan.com>) 出品的 Insight 開發系統來展示各種輸出畫面與統計圖表。該產品可以產出相當強大的展示功能，同時可以與 WWW 技術整合，讓所有使用者在任何時間、任何地點都可以透過網路瀏覽器查看各種線上分析結果。

(五) CRM 系統的中繼資訊流

有關 CRM 系統的中繼資訊流方面是由該證券商的資訊處負責「資料倉儲管理師」(Data Warehouse Manager) 以及「資料品質管理師」(Data Quality Manager) 的工作。由於在開發專案的過程中，該處人員均全程參與，故對於系統的設計理念

與架構能充分掌握，再加上後續透過教育訓練的方式讓資訊處的工程人員更能進一步完全了解整個維護的過程。

四、證券商 CRM 系統的說明

1. 資料庫綱要的分類

目前在 Tandem 中與證券業務相關的資料檔案共分成以下幾類，分成這幾類的主要原因之一，是爲了要在轉入 CRM 系統時，能區分出哪些資料需要、以及哪些資料不需要做修正與補強的動作：

- (1) 基本資料 (Preliminary Data)：用來存放客戶、營業員、券商、銀行、帳戶、股票，等基本資料，不會每天跟著交易大量變動的資料，每天只有少量的新增與變動。這些資料大多是文字型態的資料，由於各個營業據點各自維護自己內部所用到的資料，所以目前在 Tandem 端的錯誤率很高，或者是有欄位名稱但是根本沒有輸入資料 (Null Value)，這也是我們必須要先花大量時間加以正確化 (Data Cleaning) 與資料補強 (Enhancement) 的部分。
- (2) 原始交易資料 (Transaction Data)：分開存放客戶委託交易的原始委託資料 (不管有沒有撮合成功都存)，另外還有一個檔案存放了所有撮合成功的每一筆交易原始資料，這些資料大多爲數字資料，而且所有的人都會很關心其數字的正確性，所以資料的錯誤率極低，不需要也不可以修正它，因爲所有的交易內容如果需要更動時，必須經過相當的查證才能加以更改，否則會引發與客戶之間的爭議。
- (3) 證券商自行統計加總的資料 (Local Summarized Data)：存放原始交易資料的加總，其中有以股票爲主的加總資料，例如：各營業據點的股票餘額檔；也有以客戶爲主的加總資料，例如：客戶手中的持股庫存檔、客戶的信用交易餘額彙總檔、客戶每月的貢獻度與交易量等。這些資料也大多爲數字，其正確性可以經由對原始交易資料重算，或是透過加總檢查方式 (Checksum) 來加以驗證，所以即使資料有錯誤，後續也可以透過資料倉儲的重新計算來獲得正確的資料，不需要在轉換過程加以正確化。當然以原有 Tandem 端所算出來的與資料倉儲所計算出來的做雙重確認

(Double Checking)，以便修正原有 Tandem 系統中的錯誤程序更佳。我們後續是以 SQL Server 2000 的 User-Defined Functions 來計算這些加總的資料，這個動作對應到前面所述的「向上統計資料流」(Upflow)，分別會產生「初步加總資料」(Lightly Summarized Data) 與「高度加總資料」(Highly Summarized Data)。

- (4) 台灣證券交易所統計後的資料 (Global Summarized Data)：存放整體交易市場各種資料的加總，例如：集中市場各股票的融資、融券數量、各券商營業據點的交易量資訊、各股票當日收盤價檔等。這些資料大多為數字，其正確性應由台灣證券交易所 (Taiwan Stock Exchange Corporation, <http://www.tse.com.tw>) 負責，所以即使資料有錯誤，也要以台灣證券交易所重新公佈的為準，無法重新計算來獲得。除非台灣證券交易所有發佈更正資料，否則我們認為不需要加以正確化。這些資料應該依照新設計的資料庫綱要，每天將最新的資料匯入 SQL 資料庫中。其中有些資料，如：其他券商每個月的電子下單統計資訊是使用 MS Office 系列的 MS Excel 與 MS Word 格式來傳送，需要做特殊的轉換處理。
- (5) 使用者的資料 (User Profile)：這些資料是存放所有營業據點中的營業員資料，由於該證券商所採用的「知識工作者系統」中的使用權限是放在 Windows Advanced Server 2000 的 Active Directory Services 中，所以後續這些資料會透過 LDAP¹ (Lightweight Directory Access Protocol) 的方式匯入 Active Directory 中，以便進一步與「知識工作者系統」中的「安全控制子系統」做整合，這個安全控制子系統提供任何使用者單一登入的個人化服務，只需登入一次便可以使用所有授與權限的任何資料與查詢。

這些表格後續將透過 MS SQL Server 2000 中所提供的 Data Transformation Service (DTS) 進行資料的轉換，也就是圖 2 中所述的「輸入資訊流」(Inflow)。

很妙的是：我們發現整個 Data Transformation Service 的資料轉換圖與資料庫

¹ LDAP 是一種開放式的網際網路標準，由 IETF (Internet Engineering Task Force) 所制定，該協會所制定的 TCP/IP、DNS、SMTP、NNTP、SNMP、HTTP 均已成功於網際網路上實現。如要深入了解 LDAP，可以參考 [10] [14]，或連到下列網址 <http://www.innosoft.com/ldapworld/ldapv3.html>。

綱要設計圖是完全一樣的。只不過在資料庫綱要設計圖中，各個關聯表之間的連接線箭頭乃是用來呈現「外來鍵參考」(Foreign Key Reference) 的關係，因此在 DTS 轉換圖中的箭頭卻是完全相反的順序，如此才能確保轉換過程中不會違反外來鍵參考的關係。

2. CRM 系統功能

整個 CRM 系統的整體目標在於：從不同的角度，對客戶做不同的分類、分析，以便能對客戶提供適當的服務、適當的產品，採行有效的行銷方式，從而創造公司之最大利益。底下我們列出從各個角度來看每個部分之目標：

- (1) 客戶：從不同的角度，對客戶做不同的分類、分析，以便能對客戶提供適當的服務、適當的產品，採行有效的行銷方式。
- (2) 行銷：從客戶資料庫中，要能依據不同條件篩選出所需的客戶名單，提供電子行銷或其它行銷活動所用。對每個行銷活動要能統計其績效，將接觸到的投資人資料建檔，並留存行銷活動資料供別人日後參考。
- (3) 公關：透過簡便的查詢，便能將公司與外界的關係、人脈，清楚地呈現出來。
- (4) 行政管理：建立包含人員專才、證照、績效、考核等項目的人才庫，讓公司能搜尋到所需人才。並將公司的公用資源建檔，追蹤管理其使用情形。以及建立教育訓練系統，透過投影片、文件、影片、遠距教學等方式，讓員工、客戶能透過此系統學習到所需知識。
- (5) 財務：即時呈現資產、負債的配置，各單位的資金需求及資金運用的效益。
- (6) 風險管理：對公司所承擔的風險，予以量化呈現，在面臨重大事件發生時，要能估算出最大風險為何。對公司的資產與負債的真實價值，則要即時呈現。
- (7) 營運：分析出該證券商在同業間的各项相對強弱勢，及本公司在各项業務上的績效及品質。
- (8) 投資：收集國內外政經、產業情報及研究報告，提供公司、客戶在投資方面的參考資訊。從客的交易歷史中，找出贏家的操盤模式，發展成專家系

統，讓營業員對其客戶提供建議。

(9) 事件通報：建立事件內容與事件主間的關連，當事件發生時，便由系統主動以適當方式通知事件主。

(10) 訊息服務：提供訊息傳送服務，將事件內容根據指定的方式、途徑傳送到事件主。

整體系統功能上還要建立讓所有應用系統都能共用的標準，如：資料搜尋、系統登入、稽核管理、資料存取權，以及程式使用權，還有個人客製化的使用者介面等。

五、維度分析模式與資料探勘

1. 維度分析模式

在資料倉儲的建立方面，我們針對軟體需求規格書加以分析後，歸納出 18 個維度，以及 19 個資料方塊，其交互參考對照表如表 3 所示。其中的維度名稱採用以底線分隔的命名方式來表示，如果該維度名稱如果有底線區隔分項的話，就表示分項名稱之間形成階層關係，可以進一步「深入探索」(Drill-Down) 的意思。例如：“分公司_營業員_證券帳號”表示此維度是以“分公司”為維度的第一層、分公司下面再區分其所屬的“營業員”，最後營業員下面則再分成所負責的證券帳號。不過有一個「現股_融資_融券_當沖維度」例外，此維度是用來分割不同交易型態之用。

針對各個維度，我們說明如下：

- (1) 分公司：表示所有分公司營業據點。
- (2) 分公司_營業員_證券帳號：表示所有分公司之中所有營業員所負責的證券帳號。
- (3) 分公司_營業員：表示所有分公司之中所有營業員。
- (4) 市場別：目前僅區分「集中市場」與「櫃檯市場」。
- (5) 交易方式：大致上區分為「電話下單」、「語音下單」、「網路下單」，等。將來可以動態加入「有線電視下單」、「WAP 下單」、「PDA 下單」

等。

- (6) 交易年_月_日：表示所有交易的日期，依照年、月、日構成階層關係。
- (7) 交易金額級距：將委託交易金額的數字轉換成爲五個級距。這個級距可以動態由使用者自訂。
- (8) 年齡層級距：將客戶的生日轉換成爲五個年齡級距。也可以動態由使用者自訂。
- (9) 券商_營業據點：表示國內目前所有券商與其營業據點的階層關係。
- (10) 券商區域分佈維度：將券商的營業據點依照各縣市區分其所在區域，但是台北市的營業據點另行細分。
- (11) 性別：區分爲「男」、「女」、「無」(法人客戶)、未知等四類。
- (12) 客戶種類：區分爲「現場」、「電話」、「VIP」、「未知」四類。
- (13) 貢獻度級距：將證券帳號的貢獻度區分成數個級距。這個級距可以動態由使用者自訂。
- (14) 現股_融資_融券_當沖維度：將交易的模式區分爲「現股」、「融資」、「融券」與「當沖」四類。
- (15) 市場別_類股_證券：表示集中與櫃檯市場之中所有類股分類下的所有證券。
- (16) 縣市_鄉鎮：依照客戶地址郵遞區號區分其所在縣市、鄉鎮，並形成階層關係。
- (17) 投資人身分別：區分爲「本國自然人」、「僑外自然人」、「本國政府機構」、「僑外金融機構」、「本國銀行」、「僑外投資信託基金」等二十幾種身份別。
- (18) 類股名稱_證券簡稱：表示所有類股分類下的所有證券。

依據這些維度，我們依據需求規格的要求建立了以下的資料方塊 (Data Cubes)，並進行各種形式的維度分析，分析結果透過 Arcplan Insight 以 Web 畫面方式呈現。這些資料方塊說明如下：

- (1) 證券客戶結構分析：依據年齡層級距、性別、客戶種類，以及投資人身分別來分析目前在各營業據點有開戶的客戶之組成結構。
- (2) 市場消長趨勢分析：分析集中與櫃檯市場在不同時間於不同縣市鄉鎮地區的消長情況。
- (3) 分公司業務狀況分析：分析所有坐落於各縣市的分公司與所屬營業員在不同市場中的業務狀況。
- (4) 市場總成交金額分析：「集中市場」與「櫃檯市場」在不同時間的成交金額分析。
- (5) 活動戶數分析：分析每個月在各營業員旗下之客戶的活動戶數有多少。
- (6) 累計開戶數分析：分析每個月在各營業員旗下累計之客戶數有多少。
- (7) 新開戶數分析：分析每個月在各營業員旗下新開設之客戶數有多少。
- (8) 營業據點市場達成率分析：依據每年所訂定的營運目標來分析各個營業據點在各個市場的達成率情形。
- (9) 區域券商家數分析：分析各縣市區域的券商營業據點數目，包括所有在台灣設立的外國券商與本土券商。
- (10) 券商營業據點業績分析：依據台灣證券交易所傳來的資料，分析各縣市區域的券商營業據點之業績，包括所有在台灣設立的外國券商與本土券商。
- (11) 券商營業據點電子交易業績比較分析：依據台灣證券交易所傳來的資料，分析各縣市區域的券商營業據點在電子下單方面之業績，包括所有在台灣設立的外國券商與本土券商。
- (12) 歷史業績分析：分析各縣市區域的營業據點之業績與比較分析。
- (13) 類股成交金額比重分析：將證券依據其所屬類股進行成交金額與比重的分析。
- (14) 融資融券維持率分佈：分析各營業據點的融資與融券狀況。
- (15) 市場融資融券分析：分析集中與櫃檯市場中的融資與融券狀況。
- (16) 客戶身份別業績分析：依照客戶的身份別來分析其對該證券商的業績貢

獻情況。

(17) 客戶交易方式分析：分析客戶在交易方式方面的習性改變。

(18) 交易種類業績比較：分析各種交易種類對整體業績的貢獻情形。

(19) 投資類股比重分析：分析不同分公司在不同類股與證券上的投資比重配置情形。

表 3：證券商 CRM 系統的維度與資料方塊交互參考對照表

維度 (Dimension) 資料方塊 (Cube)	分公司	分公司 營業員 證券帳號	分公司 營業員	市場別	交易方式	交易年 月 日	交易金額級距	年齡層級距	券商 營業據點	券商區域 分佈維度	性別	客戶 種類	貢獻度級距	現股 融資 融券 當沖維度	市場別 類股 證券	縣市 鄉鎮	投資人 身分別	類股名稱 證券簡稱
證券客戶結構分析								√			√	√	√				√	
市場消長趨勢分析				√		√										√		
分公司業務狀況分析			√	√		√						√				√		
市場總成交金額分析				√		√							√					
活動戶數分析			√			√						√						√
累計開戶數分析			√			√						√						√
新開戶數分析			√			√						√						
營業據點市場達成率分析	√					√												
區域券商家數分析									√	√								
券商營業據點業績分析						√			√	√								
券商營業據點電子交易 業績比較分析						√			√	√								
歷史業績分析	√					√								√				
類股成交金額比重分析						√								√	√			
融資融券維持率分佈	√	√			√	√								√				
市場融資融券分析						√								√	√			
客戶身份別業績分析											√	√						√
客戶交易方式分析	√				√						√	√						√
交易種類業績分析		√			√	√												√
投資類股比重分析	√				√	√												√

另外，19 個資料方塊則依據需求規格的要求進行各種形式的分析，並將分析結果以 Arcplan Insight 以 Web 畫面方式呈現。

2. BCG 模式與資料分析

除了建立資料倉儲以利多維度的資料分析之外，另外我們還透過了管理學上長久以來被採用的 BCG (Boston Consulting Group) Growth/Share matrix 模式 [8] [9] 分析下列的結果：

- (1) 將各營業據點的市佔率與成長率分別求出後，以橫軸表示市佔率，縱軸表示成長率，然後將各營業據點依照這兩個指標置於 BCG 模式中，並加以分類為「狗」(Dogs)、「母牛」(Cash Cows)、「明星」(Stars)、「問題」等 (Question Marks)。如圖 4 所示即為透過 Arcplan Insight 所呈現的 2001 年第三季各營業據點針對某個下單方式的 BCG 模式圖。
- (2) 將所有券商的市佔率與成長率分別求出後，置於 BCG 模式中，並加以分類為「狗」、「母牛」、「明星」、「問題」等。以便找出自己目前的狀態。

當然，此一應用也可以擴展到將傳統下單方式、網路下單，以及其他下單方式分成「狗」、「母牛」、「明星」、「問題」等類別，以看出網路下單是否會成爲明日之星。

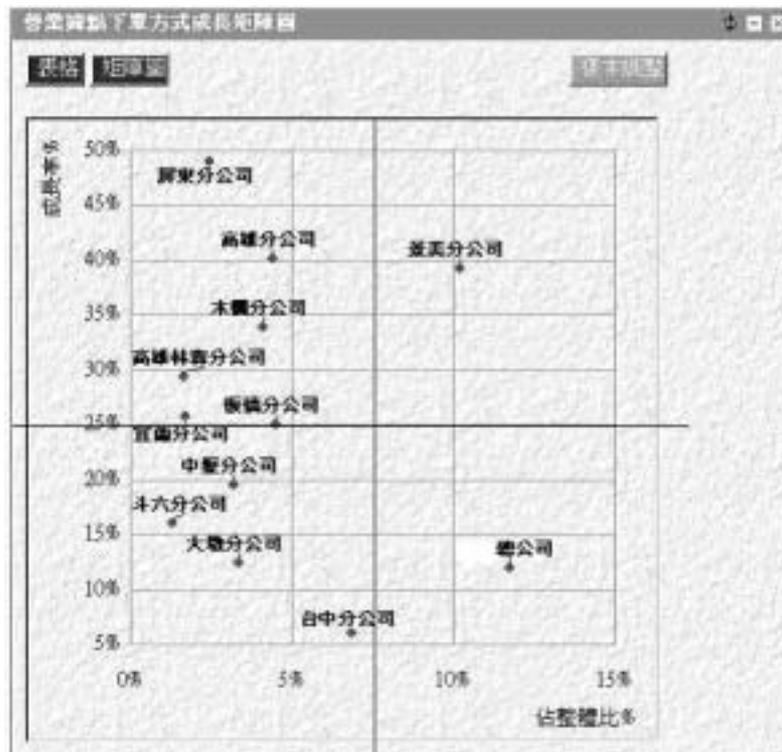


圖 4：2001 年第三季各營業據點針對某個下單方式的 BCG 模式圖

3. 運用 SQL Sever 2000 做資料探勘

由於本系統所採用的資料庫管理系統 MS SQL Server 2000 已經具備資料探勘的 (Data Mining) 功能，因此我們希望能利用該系統所提供的 Microsoft Decision Trees 與 Microsoft Clustering 兩個演算法來建立適當的模型並找出可用的樣型 (Patterns)。其中的 Microsoft Decision Tree 可以依據過去的客戶之「風險偏好」與「風險承受度」資料建立一個決策樹 (Decision Tree)，以使用來預測新客戶的「風險偏好」與「風險承受度」。而 Microsoft Clustering 演算法則可以協助我們將現有客戶依據各種不同的屬性分成幾大類，以方便行銷部門在做行銷方案規劃時的參考。由於資料探勘乃一持續性的分析與挖掘過程，故目前尚未整合到整個 CRM 系統中，而是採用「線外運作」(Off-Line Operation) 的方式進行。

六、經驗總結與未來研究方向

在整個系統的建置過程中，該公司資訊處扮演了相當重要的聯絡窗口角色，尤其在資料轉換、系統規格、使用者介面等方面幾乎至少每隔一個月便會召開一次會議做充分的討論，同時也會將證券業的專業知識 (Domain Knowledge) 傾囊相授，而 N 科技股份有限公司 在人員技術訓練上的紮實程度，以及透過建置 Enterprise Portal Site 形成綿密的溝通表達管道，則提供了技術與討論作業上的後勤支援。這兩者的完美搭配是造就本系統成功的主要因素，此番系統開發經驗也讓 N 科技股份有限公司 在榮獲台灣微軟 Microsoft FY01 Best Windows 2000 Deployment Award 之後，得到了更上層樓的肯定與好評。

由於資料庫與資料倉儲的規劃佔了相當大的一部份比重，據此我們也發現並吸收到了許多寶貴的經驗。尤其我們運用了大量傳統關聯式資料庫的「索引」(Indexing) 建置技術與「反正規化」(De-Normalization) 技巧，也發現它們在效能調整方面的確佔有相當大的影響力。另外，為了避免錯誤的資料維護過程造成系統的不穩定，或產生不正確的分析資料，我們必須將資料庫中的關聯表做妥善的分類，如 4.1 節所述。同時，對於被反正規化的關聯表格所參考的表格也必須撰寫妥善的「觸發程序」(Triggers)，將所有因為反正規化而產生的欄位複本做連鎖的更新動作，方能維持整個資料庫的一致性。

目前整個系統已經建置完成，正進行驗收程序，由於後續還有第二期將進行期貨與債券的 CRM 資料倉儲建置，並且希望能與證券部分做一完整的整合，所以我們希望能先將第一期的功能部分做驗收。後續尚待完成以及未來可以在此資料庫

做研究的課題有下列幾項：

- (1) 整體系統效能的調整與改善：持續改善整體系統的運作效能，並藉以訂定整個 CRM 系統的測試標記 (Benchmarking)，做為後續第二期的參考。後續不排除以 PC Clusters 做平行處理以加速效能的架構測試。
- (2) 用戶滿意度調查：進行高階使用者與營業員本身，以及他們所收集的客戶滿意度情況進行調查，以做為系統功能與使用介面的改善依據。
- (3) 採用更多的 Data Mining 工具進行資料的分析：由於目前 MS SQL Server 2000 僅有 Microsoft Decision Trees 與 Microsoft Clustering 兩種演算法可供使用，略嫌不足。故將來希望能配合 Microsoft 的協力廠商 (Third Parties) 所開發的其他 Data Mining 工具做進一步的資料分析。
- (4) 以 XML 整合異質性資料庫之環境建置：為因應未來各種異質性資料的整合需求，以及與 Internet 應用的接軌，因此我們也會朝向以 XML 做為異質性資料庫整合的媒介，並且基於目前已經具有的成果 (見 [21]) 朝向將 XML 與異質性資料庫應用合而為一的目標邁進。有關 XML 的相關簡介，讀者們可以參考 [2] [4] [5]。有關 XML 與資料庫的整合說明，有興趣的讀者們也可以參考 [3]。
- (5) 完整的「證券」、「期貨」、「債券」、「基金」整合式資料倉儲：本期專案主要是建置「證券」方面的資料倉儲，後續會漸次加入「期貨」、「債券」、「基金」的部分，當然將這些資料倉儲完全整合正是未來唯一的方向與目標。

客戶關係管理是一項持續不斷改進的工作，唯有持續的進步才能獲得客戶的認同與好評，因此我們希望將此一系統建置經驗提供給相關產業界參考。尤其在台灣已經成為 WTO 會員國的今日，我們更希望藉此讓國內的產業界能持續向上提昇自我的競爭力，再創台灣經濟奇蹟。

[謝啟]

本研究承蒙 N 科技股份有限公司 計畫補助、證券公司 資訊處 鍾協理、施經理、陳襄理 暨所有同仁在資料提供與系統建置過程上的協助，以及 國立高雄第一科技大學 創新育成中心 的輔導，特此感謝。

參考文獻

- [1] Anahory, S. and Murray D., *Data Warehousing in the Real World: A Practical Guide for Building Decision Support Systems*, Harlow, England: Addison-Wesley Longman, 1997.
- [2] Bergholz, A., “Extending Your Markup: An XML Tutorial”, *IEEE Internet Computing*, Vol. 4, No. 4, July-August 2000, pp.74-79.
- [3] Bertino, Elisa and Barbara Catania, Integrating XML and Databases, *IEEE Computer*, Vol. 5, No. 4, 2001, pp. 84-88.
- [4] Bosak, J., T. Bray, D. Connolly, E. Maler, G. Nicol, C.M. Sperberg-McQueen, L. Wood, and J. Clark, “W3C XML Specification DTD”, 1998, <http://www.w3.org/XML/1998/06/xmlspec-report-19980910.htm>
- [5] Bray, T., J. Paoli, C.M. Sperberg-McQueen, “Extensible Markup Language (XML) 1.0”, 1998, <http://www.w3.org/TR/REC-xml>.
- [6] T. Connolly, C. Begg, and A. Strachan, *Database Systems: A Practical Approach to Design, Implementation, and Management*, Addison-Wesley, Second Ed., 1999.
- [7] R. Hackathorn, Data Warehousing Energizes Your Enterprise, *Datamation*, Feb. 1995, Vol. 1, pp. 38-42.
- [8] P. Haspeslagh, Portfolio Planning: Uses and Limits, *Harvard Business Review*, Vol. 60, No. 1, 1982, pp.58-67.
- [9] R. Hamermesh, *Making Strategy Work*, New York: John Wiley & Sons, 1986.
- [10] Howes, T. and M. Smith, *LDAP: Programming Directory-Enabled Applications with Lightweight Directory Access Protocol*, Macmillan Technical Publishing, USA, 1997.
- [11] Inmon, W.H., *Building the Data Warehouse*, New York, NY: John Wiley and Sons, 1993.
- [12] Kimball, R, *The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses*, John Wiley & Sons, Inc., 1996.
- [13] Kimball, R. and R. Merz, *The Data Webhouse Toolkit—Building the Web-Enabled Data Warehouse*, John Wiley and Sons, Inc., 2000.
- [14] P. Loshin, ed., *Big Book of Lightweight Directory Access Protocol (LDAP) RFCs*, Morgan Kaufman, San Francisco, Calif., 2000.
- [15] Ohaegbu, K., S.S. Devgan, Customer Relationship Management in E-Commerce: The Call Center Solution, *Proceedings of the IEEE*, 2000, pp. 391-394.
- [16] Peppers, D., and Rogers, M. (1997). *The One to One Future: Building Relationships One Customer at a Time*, Bantam Doubleday Dell Publishing.
- [17] D. Pyle, *Data Preparation for Data Mining*, Morgan Kaufmann, 1999.
- [18] Spofford, G., *MDX Solutions: With Microsoft SQL Server Analysis Services*, John Wiley & Sons, Inc., 2001.
- [19] Srivastava, J. and P.Y. Chen, Warehouse Creation—A Potential Roadblock to Data Warehousing, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11, No. 1, 1999, pp. 118-126.
- [20] Thomsen, E., G. Spofford, and D. Chase, *Microsoft OLAP Solutions*, John Wiley & Sons, Inc., 1999.

- [21] L.H. Lin, Tseng, F.S.C and W.M. Chuang, “An XML-Based Framework for Inter-Organization Business Workflow Management and Information Interchange”, *Proceedings of the 13th International Conference on Information Management*, May 2002, pp. 401-408. 中文名稱：以 XML 為基礎之跨組織工作流程管理與資訊交換架構之研究，第十三屆國際資訊管理學術研討會論文集。