

第二章

相關研究

在這一章我們回顧和本研究相關的研究成果與文獻。主要分為兩個部分，第一部分我們將會就目前在分散式分享網路上已有研究成果的 P2PDSN [32] 以及基於 D-S 證據推論理論 (Dempster-Shafer theory) [31] 所建立的信賴模型 [1] 來做概略的介紹，特別是兩者所採用的信賴機制的部份會做較深入的探討；第二部份則會就兩者目前尚有不足之處做些說明，並詳列本文可以加以改進之處。

2.1. P2PDSN 信賴模型

Smirnova 在西元 2002 年曾發表了一篇碩士論文「Multi-agent System for Distributed Data Fusion in Peer-to-Peer Environment」 [32]，並在文中提出了「P2PDSN」，成為目前研究「Brokering system between GRID and P2P」 [19] 上一項十分重要的概念。

Smirnova 認為要有效地在分散式系統中處理分散式分享的問題，最適切的模型便是利用「多代理人」(MAS, Multi-Agent System) [18] 來控管每一項資源的使用以及再使用。多代理人運算環境，指的是在同一運算系統或子系統之中，應用一個以上不同功能的代理人，協同運作處理更多的訊息流向。相對於傳統的單一代理人環境 (SAS, Single-Agent System) [18]，多代理人運算可以掌握到更多更精確的訊息。對於許多複雜系統，多代理人運算的方式變得相當的重

要。而目前在研究多代理人運算環境領域中，應用最為廣泛應屬 DSNs (Distributed Sensor Network) [17] 模型。分散式感應網路 (DSNs) 是一個在最近非常受到重視的一個領域，應用的範圍諸如戰場監視、搖控感測、全球意識等等以時間為關鍵原素，範圍廣大且非常要求訊息的可靠性。Smirnova 將這樣的觀念結合 P2P 分散式分享網路，將每一個電腦節點廣義的定義為一個小型的 DSN 系統，結合無數的 DSN 便成為 P2PDSN 網路的基本模型。圖 2.1 顯示簡單的 P2PDSN 架構圖。

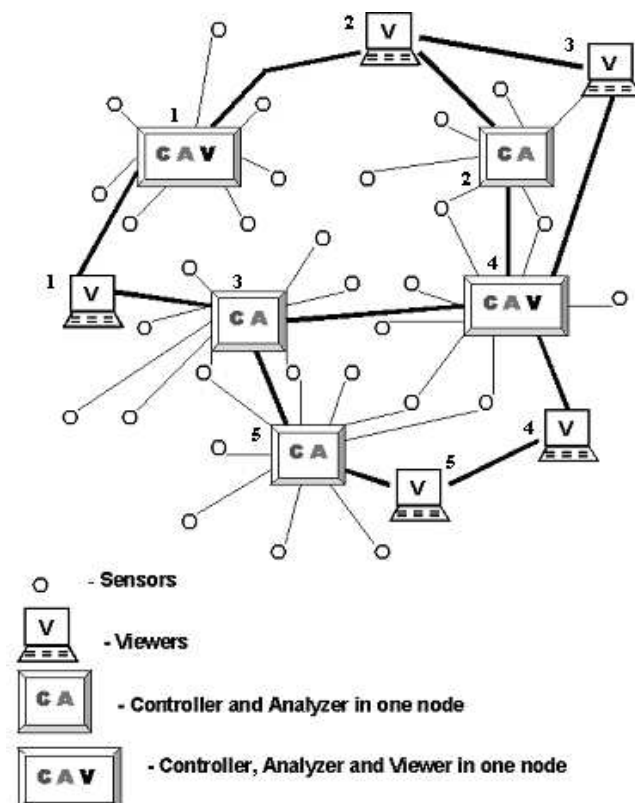


圖 2.1，P2PDSN 示意圖 [18]

如上圖所示，P2PDSN 模型的建立是基於四個合乎邏輯的實體：「傳感器」(O, Sensor)，「控制單元」(C, Control)，「分析器」(A, Analyzer) 和「使用者葉節點」(V, Viewer)。我們將每一個 P2P 分散式分享網路上的使用者（電

腦節點) 視為一個小型的 DSN 系統，每一個 DSN 系統都擁有傳感器、控制單元、分析器、與使用者葉節點。其中，控制元件是能獨立實體執行計算和主機資源的資源掌管與執行任務元件；分析元件則是用來分析運算任務所需用到的計算資源，並掌管是否回應服務請求的權力。使用者葉節點處理計算任務的請求，並將其傳送到所屬的控制元件或者分析元件；控制元件以及分析器則努力滿足被指定的計算任務。每一個電腦節點在服務之初無需認識整個拓樸網路的全貌，只需要在自身無法完成被指定的運算任務時，由所屬的傳感器尋求附近另一個可提供服務的節點，並與其交換訊息請求協助。節點相互之間透過傳感器交換訊息修改他們所認識的「環境」，進而得到整個拓樸網路概觀，以利下次再需要協助時可以更快的得到服務。圖 2.2 到圖 2.4 顯示以 CAV_i 為控制元件的節點請求其他電腦節點服務時的步驟。

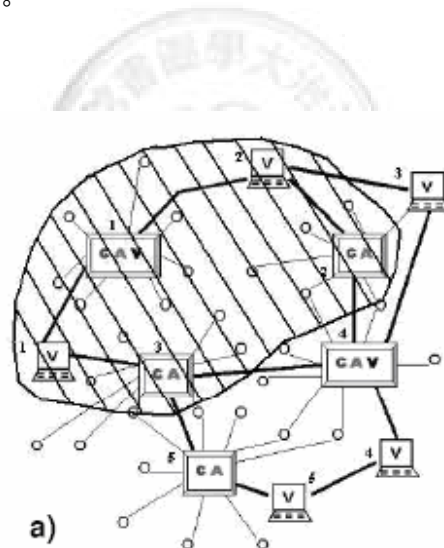


圖 2.2, 控制節點 CAV_i 在服務之初所擁有的拓樸概觀 [18]。

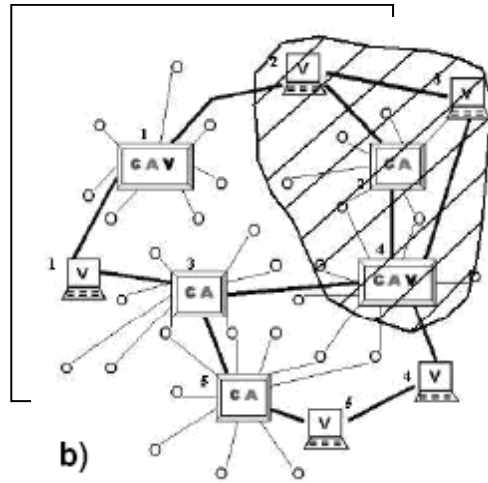


圖 2.3, 被請求服務的控制節點 CAV_2 所擁有的拓樸概觀 [18]。

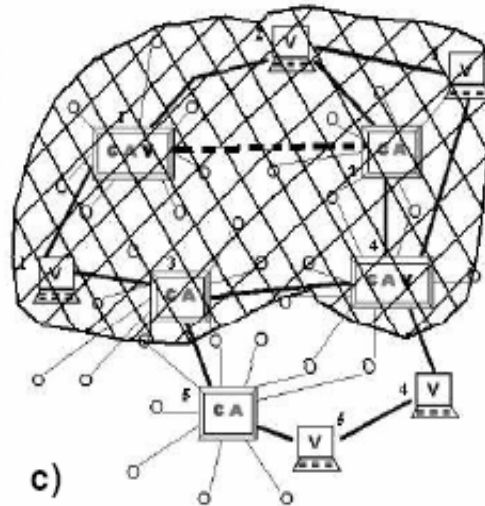


圖 2.4, 當 CAV_2 回應 CAV_1 的請求後, CAV_1 將得到兩者的拓樸概觀集合 [18]。

以圖 2.4 為例, 節點 CAV_1 在服務之初會先得到這樣的領域概觀: $\{V_1, CA_3, V_2, CA_2\}$ 。如果 CAV_1 因為無法完成本身所被託付的任務時, 它便需要其他控制元件的協助而需要其他領域概觀。此時它開始透過本身所屬的傳感器去偵測尋找周遭其他控制元件的位置, 並發送請求得到它們擁有的運算資源。依照圖 2.4

所示， CAV_1 將送一次廣播請求得到它需要的服務， CAV_2 得到請求後將本身的運算資源回應給 CAV_1 ，待 CAV_1 得到服務回應後， CAV_1 將一併擁有查看 CAV_2 拓樸概觀的能力。我們可以從上述的步驟了解到 P2PDSN 具有下面幾點優點：

- 1) 沒有單一節點失敗的瓶頸。藉由傳感器偵測的方式，當電腦節點發現本身無法完成計算任務時，才向外請求其他控制單元（電腦節點）的協助。當傳感器偵測時，是以廣播的方式向節點周遭詢問，所以可尋得服務的路徑並不是唯一，只要是在節點四周的控制單元，都可以成為尋求服務路徑上的一點。如此一來，便不必擔心單點瓶頸形成的危險。
- 2) 無需事先得到整個網路拓樸，便可以進行搜尋。傳感器透過網路底層的通訊，互相交換訊息，在節點接獲運算任務的初期，若無須向外請求其他節點協助，則傳感器僅需知道四周其他的控制單元的位置，無須知道在他們之外是否還有其他控制單元。如此一來可以節省單一電腦節點資源，以供其他需要計算資源的節點使用。
- 3) 分層處理資料的融合（Data Fusion），減少頻寬需求，加快訊息的處理。P2PDSN 架構有一個非常重要的好處，就是當請求服務的路徑上的節點大於等於三個時，由於每一個節點的控制單元都有獨立運算的能力，所以當節點 A 發出協助運算的請求，通過節點 B 到達節點 C 時，此時若 C 已經完成前一部分的運算，則當運算結果返回到 B 時，無須再返回到 A 即可做兩者的資料融合（Data Fusion），再將新算式返回給 C，讓節點 C 可以很快的得到新的算式，繼續運作。當所有的算式都完成後，再由節點 B 作總和，一次將結果返回給 A，如此一來便可以節省 A 到 B 之間的頻寬，加速運算結果的出爐，並且節省節點 A 的計算資源。這個現象，在 Smirnova 的文章中，稱之為「分散式資料融合」（Distributed Data Fusion, DDF）或稱為「中間層資料融合」（Broker Data Fusion, BDF）[18]。請求服務的路徑越長，則此一現象得到的好處變越明顯，但前提是在經過幾次的訊息傳遞後，節點 B 應從節點 A 獲得所有計算任務的算式，形同複製計算任務。

P2PDSN 架構嚴謹，各元件之間分工詳細，並利用中間層資料融合的架構來提升計算任務的運算速度，改進了傳統 P2P 分散式分享網路的許多缺點，目前也已經應用在很多地方，例如 NaradaBrokering System [20] 就是一個有名的例子。

2.2. 基於 D-S 證據推論理論所建立的信賴模型

2.2.1. D-S 證據推論理論

演繹統計法 (The Deductive-Statistics method, the D-S method)，是統計學中針對研究的主題，提供一種對於經驗現象的系統化理解，且將這些現象放置在符合法則的模式中來理解現象之間的關係。例如硬幣的正反面出現機率在每一次的丟硬幣過程中，都是 50%，因此可以推論每一次丟硬幣的正面反面出現機率都是 50%。這種推論的方法在 1976 年由 G.Shafer 定義為證據理論 (Evidence theory) [31]。以過去的經驗做為往後抉擇的依據，並以之做為準則的系統，這便是所謂的 D-S 證據推論理論 (Dempster-Shafer evidence theory)。

舉例來說，我們假設 Q 是一個問題，其全部可能的答案用集合 Θ 來表示， $\Theta = \{T, \neg T\}$ ，其中 T 表示一個節點 A 對節點 B 信任，而 $\neg T$ 表示節點 A 對節點 B 不信任，且 $t \in \{T, \neg T\}$ 是問題的正確答案，我們希望藉由網路上其他的節點的「推薦」找到正確的答案 t 。所謂的「推薦」是指一種信賴關係，當本身對事件 X 所能掌握的資訊不足時，完全仰賴他人對 X 的資訊來做決定，並相信他人所給予的資訊，這樣的關係便是所謂的推薦，而以這樣的關係為基礎所建立的網路系統便稱為「推薦網路」。集合 Θ 便稱為問題 Q 的識別範圍。由於證據的不完備、不精確、不完全可靠等等原因（例如在 A 節點

上沒有 B 節點的交易行為紀錄，也就是當 A 無法掌握 B 之前的歷史證據時），我們便無法確定 t 的正確性。但由於推薦網路所獲得的推薦證據在一定程度上可以確定答案 t 所在的範圍是可能的，也因此 Shafer 定義了基本可信度函數的概念 [31]。

2.2.2. 基於 D-S 證據推論理論的信賴模型

中國科學技術大學計算機科學技術系的朱俊茂等人在 2003 年便提出了「Grid 與 P2P 混合計算環境下基於推薦證據推理的信任模型」[1]。朱先生認為，在分散式混合計算的環境（Grid&P2P）中，因為 Grid 節點（文中稱為 GP）可以提供 QoS 的服務保證，而 P2P 節點（文中稱為 PP）的計算資源屬於自發性的分享資源，並不提供 QoS 保證，所以要在混合計算的環境中建立信任推薦機制，尤其是在整個系統之初並沒有任何歷史經驗可以參考的情況下，通過 D-S 證據推薦理論作為基礎來實現，可以很簡單且有效的解決 Grid&P2P 中的信任問題。

基於 D-S 證據推論理論的信賴模型，最主要的目標是在解決 Grid 跟 P2P 兩種不同分散式分享網路混合運作時所產生的信賴度認定方式無法統一的問題。Grid&P2P 混合環境中，一般仍是讓 GP 節點（Grid Point）充當運算資源以及任務的註冊與調配；PP 節點（P2P Point）擔任葉節點的角色，處理被指派的計算任務。由於 Grid 與 P2P 的基本信任制度不同，所以當 GP 需要指定一個計算任務時，或是 PP 請求一個計算任務或要求資源，就很容易發生詐欺或是沒有法則可以作為任務指派的依據。透過 D-S 證據推論信賴模型，我們可以看到針對 Grid&P2P 混合運算環境中的信賴度評量有以下幾點尚可加以改進：

- 1) 定義 PP 節點的信賴度初值設定。我們了解在原先的 P2P 分散式分享網路架構下，任何一個分享節點是透過登入與登出網路時更新鄰近集線器的分享

資源名單。分享資源的好壞事實上完全仰賴 *PP* 節點的「道德修養」，也就是參予提供計算資源的 *PP* 節點會不會提供詐欺服務，端視 *PP* 節點是否為懷有惡意的節點。所以當 *PP* 節點在 Grid&P2P 混合環境下運作時，我們必須給定一個信賴度的初值。初值的設定，必須不能過小，以防初次分享計算資源的 *PP* 節點因為信譽值過低而永遠得不到計算任務，以至於長時間處於飢餓狀態 (Starvation)；也不能過高，否則 *PP* 節點很容易以詐欺的手段取得計算任務，導致計算任務的品質低落甚至失敗。D-S 證據推論信賴模型針對 Shafer 原先的可信度提出這樣的改進，但在「基於 D-S 證據推論系統所建立的信賴模型」文中並無詳細描述。不過一般而言，初值 α 應該不脫 $\Theta = \{T, \bar{T}\}$ 的信賴集合範圍。

- 2) 引入欺騙行為懲罰機制。在信賴模型中，信賴度評價的好壞是選擇服務的重要依據。對於服務良好的 *PP* 節點我們必須給予獎勵，增加其信譽值，使其下一次可以很容易的取得計算任務；同樣地針對不誠實的 *PP* 節點也應該予以懲罰，促使節點注重服務的品質。在原先 Shafer 的定義中，每一個 *PP* 節點只有「信賴」與「不信賴」兩種情形，這樣的定義顯然有不公平之處。D-S 證據推論信賴模型將這樣的定義加以改良，加入了懲罰係數 ρ ，讓 *PP* 節點的信譽值在 $\Theta = \{T, \bar{T}\}$ 之間增減，增加遴選品質優良的服務的公平性。
- 3) 建立賦予 *GP* 節點的權值機制。Grid&P2P 混合運算環境，是由「分配計算任務的 *GP* 節點」與「分享計算資源提供服務的 *PP* 節點」所組成。如同第二點，*PP* 節點與 *GP* 節點間有懲罰機制的問題，同樣的 *GP* 節點相互之間亦然。接受計算任務的 *GP* 節點在切割任務並評估所屬的 *PP* 節點無法完成所有的計算任務時，便會向外請求其他的 *GP* 節點支援，此時發生 *GP* 節點的權值問題，也就是 *GP* 的懲罰機制。權值越高的 *GP* 節點，下次獲得分割任務指派的機會就越高；反之若時常出現詐欺行為的 *GP* 節點，勢必漸漸地沒有人願意將計算任務分割賦予。D-S 證據推論信賴模型對不同的推薦 *GP* 節點，賦予了不同的權值。假設自身節點權值定義為 ω ， $0 < \omega < 1$ ，則其他 *GP* 節點的權值為 $(1 - \omega) / (n - 1)$ 。這個公式代表每一個 *GP* 節點在信賴度排序上的比重，其中 n 代表了 *GP* 節點的總數。當其

中一個 *GP* 節點獲得計算任務且需要其他 *GP* 節點的協助時，便會以此做為選擇協助的依據。同時為了防止 *GP* 節點的詐欺行為，文中定義 *GP* 懲罰係數 ε ，在每次計算任務結束後，對有不誠實行為的 *GP* 節點依懲罰係數 ε 降低其權值。文中還提到懲罰的方式有兩種：如果自身節點認為目前的混合計算環境中不存在其他 *GP* 節點協同的欺騙行為的話，那麼該把懲罰係數均分給其他誠實節點；反之自身節點把該懲罰係數留給自己 [1]。

- 4) 最後則是 *PP* 節點的選擇機制。在不失一般性的情況下，直覺上應該把計算任務給信賴度最高的 *PP* 節點，但是如果長時間這樣分配下來的話，將會出現負載不平衡的問題：某一個 *PP* 節點成為服務熱點，不斷的忙碌，但是其他也符合條件的節點卻分配不到計算任務。D-S 證據推論信賴模型採用的選擇方式是，假設 N 是一個可以設定的參數，則在符合信譽條件的 q 個 *PP* 節點： $\{PP_1, PP_2, \dots, PP_q\}$ 中，取 $\min(N, q/2)$ 個節點，並隨機從中選取一個節點指派任務，進而有效地避免計算熱點的現象產生。

D-S 證據推論信賴模型定義了較 P2PDSN 更為明確的信賴制度，並利用懲罰機制來區分服務節點的優劣，有效的避免節點出現詐欺行為以及服務熱點的產生，改進了許多混合環境網路的缺點。

2.3. 本論文認為尚可改進之處

在觀察 P2PDSN 與 D-S 證據推論信賴模型之後，我們認為以下幾點仍是

可以加以改進的地方：

- 1) 信賴機制的建立：P2PDSN 的控制單元在需要其他計算資源的協助時，必須透過所屬的傳感器向周遭詢問，是否有其他控制單元可以提供服務；D-S 證據推論信賴模型則是透過賦予 *GP* 節點權值的方式，挑選權值較高的 *GP* 節點來執行計算任務。嚴格說來，兩篇論文僅僅建立了信賴模型的「詢問制度」以及「信譽值的賦予」，實質上並不能算是信賴模型。如何去評量一

個節點的信賴度，以及透過信賴度的評量去得到良好的計算資源，兩篇論文沒有進一步地加以考量。我們認為，在現實社會中有所謂的「口耳相傳」，透過其他友好節點的「層層推薦」，這樣才能讓即使是不認識的優良節點也可以有機會得到它的服務。

- 2) 懲罰值的賦予：我們認為懲罰值的賦予應當如信譽值的賦予相同，具有「推移性」以及「遞增性」。當自身節點需要其他節點提供服務時，必然會詢問其他友好節點。假設節點 A 在詢問節點 B 後，在透過 B 的好友 C 得到節點 D 的服務，卻因為節點 D 的詐欺行為導致計算任務的失敗，則此時節點 B 、 C 、 D 都應當遭受懲罰值的賦予，且懲罰值的大小應當是 $D > C > B$ 。理由是因為 D 出現了詐欺行為，而 C 推薦了 D ，所以必須負起連帶責任。但 B 很有可能不是 D 的朋友，所以 B 的懲罰值理論上應該小於 C 。也就是說，在經過了「一連串」的推薦之後，不論最後計算任務成功或是失敗，「責任」都應該被「依照比例分攤」在信賴路徑上的每一個節點，而非平均分攤。
- 3) 避免熱點的產生：P2PDSN 與 D-S 證據推論信賴模型都在避免熱點的產生或是因為計算任務導致某個節點特別忙碌，而其他具有優良服務且條件相同的節點卻處於飢餓狀態。這兩篇論文都從信賴度的分配方面去著手，或是隨機挑選，但本論文認為，這樣仍然有很大的機率產生熱點。本論文認為，當一個節點信譽良好自然得到很多的計算任務；同樣的當它無法再負荷時，它為了保有自身良好的信譽，必將無法完成的任務「轉手」給其他友好且擁有相同條件的節點，這也就是本論文提出的「轉手路徑」(Pass-over path) 以及「同業公會」(Guild) 的概念。透過轉手與同業公會的推薦，可以協助將過量的計算任務很快地轉到其他閒置的節點手上，進一步降低熱點產生的可能性。

我們在下一章將就本論文所依據的社會科學概念做較深入的探討，詳細描述上述的三項改進措施所依據的理論並印證我們的研究目標。