



(19) 中華民國智慧財產局

(12) 發明說明書公告本

(11) 證書號數：TW I714496 B

(45) 公告日：中華民國 109 (2020) 年 12 月 21 日

(21) 申請案號：109112410

(22) 申請日：中華民國 109 (2020) 年 04 月 13 日

(51) Int. Cl. : G06N3/08 (2006.01)

G06F17/16 (2006.01)

(71) 申請人：國立清華大學 (中華民國) NATIONAL TSING HUA UNIVERSITY (TW)

新竹市光復路二段 101 號

(72) 發明人：邱偉育 CHIU, WEI-YU (TW)；蔡松佑 TSAI, SUNG-YU (TW)

(74) 代理人：葉璟宗；卓俊傑

(56) 參考文獻：

TW M451751U1

TW 201236301A

TW 201626770A

TW 201817275A

TW 201902152A

審查人員：陳奕昌

申請專利範圍項數：15 項 圖式數：3 共 27 頁

(54) 名稱

無線電力驅動通訊網路的強化學習通訊時間分配方法及基地台

(57) 摘要

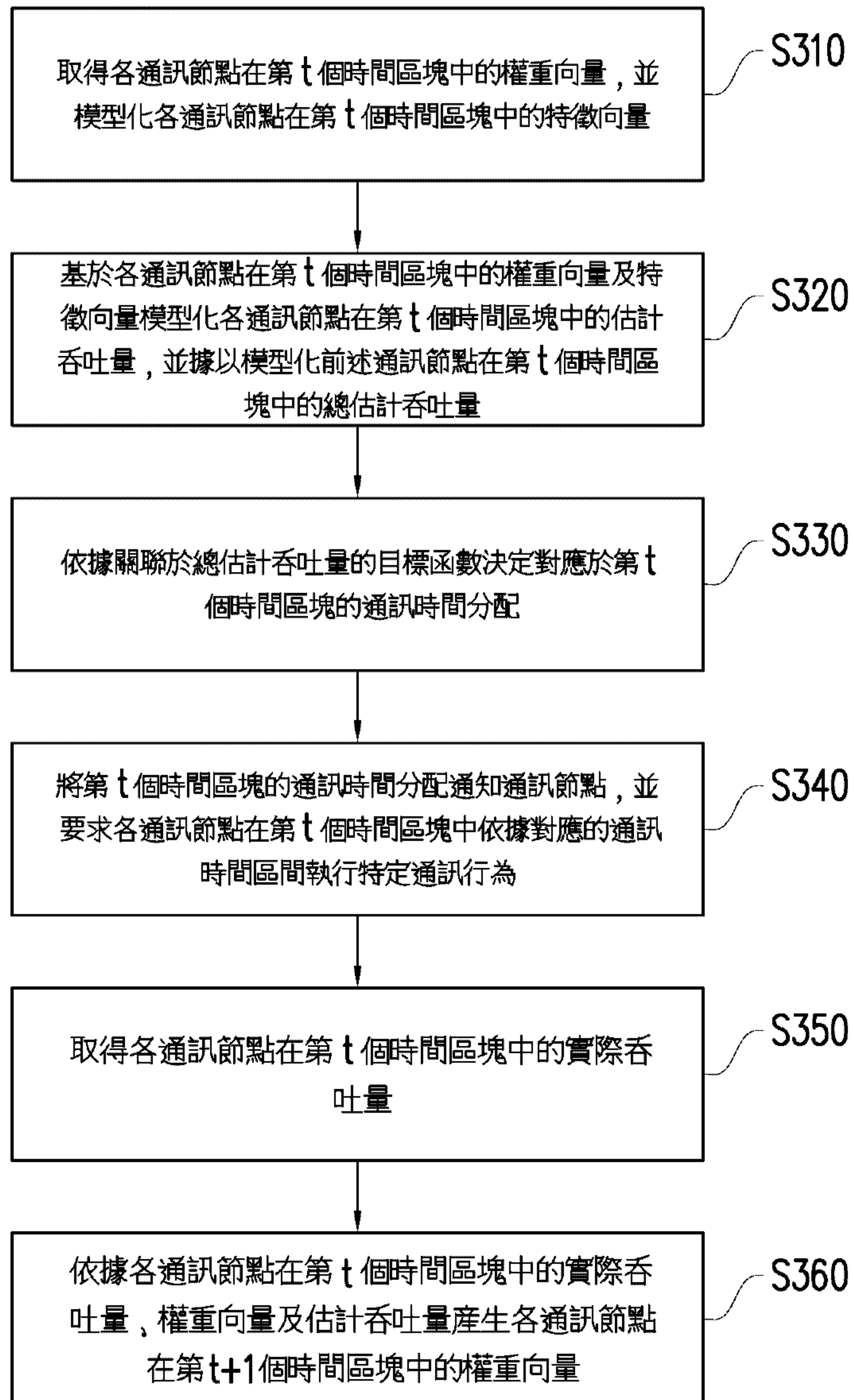
本發明提供一種無線電力驅動通訊網路的強化學習通訊時間分配方法及基地台。所述方法包括：依據關聯於通訊節點的總估計吞吐量的目標函數決定對應於第 t 個時間區塊的通訊時間分配；要求各通訊節點在第 t 個時間區塊中依據對應的通訊時間區間執行特定通訊行為；取得各通訊節點在第 t 個時間區塊中的實際吞吐量；依據各通訊節點在第 t 個時間區塊中的實際吞吐量、權重向量及估計吞吐量產生各通訊節點在第 $t+1$ 個時間區塊中的權重向量。

The invention provides a communication time allocation method using reinforcement learning for a wireless powered communication network and a base station. The method includes: determining a communication time allocation corresponding to the t -th time block according to an objective function associated with the total estimated throughput of the communication nodes; requesting each communication node to perform specific communication behaviors according to the corresponding communication time interval in the t -th time block; obtaining the actual throughput of each communication node in the t -th time block; generating the weight vector of each communication node in the $(t+1)$ -th time block according to the actual throughput, the weight vector, and the estimated throughput of each communication node in the t -th time block.

指定代表圖：

符號簡單說明：

S310~S360: 步驟



【圖3】



I714496

【發明摘要】

【中文發明名稱】 無線電力驅動通訊網路的強化學習通訊時間分配方法及基地台

【英文發明名稱】 COMMUNICATION TIME ALLOCATION METHOD USING REINFORCEMENT LEARNING FOR WIRELESS POWERED COMMUNICATION NETWORK AND BASE STATION

【中文】 本發明提供一種無線電力驅動通訊網路的強化學習通訊時間分配方法及基地台。所述方法包括：依據關聯於通訊節點的總估計吞吐量的目標函數決定對應於第 t 個時間區塊的通訊時間分配；要求各通訊節點在第 t 個時間區塊中依據對應的通訊時間區間執行特定通訊行為；取得各通訊節點在第 t 個時間區塊中的實際吞吐量；依據各通訊節點在第 t 個時間區塊中的實際吞吐量、權重向量及估計吞吐量產生各通訊節點在第 $t+1$ 個時間區塊中的權重向量。

【英文】 The invention provides a communication time allocation method using reinforcement learning for a wireless powered communication network and a base station. The method includes: determining a communication time allocation corresponding to the t -th time block according to an objective function associated with the total estimated throughput of the communication nodes; requesting

each communication node to perform specific communication behaviors according to the corresponding communication time interval in the t-th time block; obtaining the actual throughput of each communication node in the t-th time block; generating the weight vector of each communication node in the (t+1)-th time block according to the actual throughput, the weight vector, and the estimated throughput of each communication node in the t-th time block.

【指定代表圖】圖3。

【代表圖之符號簡單說明】

S310~S360:步驟

【特徵化學式】

無

【發明說明書】

【中文發明名稱】無線電力驅動通訊網路的強化學習通訊時間分配方法及基地台

【英文發明名稱】COMMUNICATION TIME ALLOCATION METHOD USING REINFORCEMENT LEARNING FOR WIRELESS POWERED COMMUNICATION NETWORK AND BASE STATION

【技術領域】

【0001】本發明是有關於一種通訊時間分配方法，且特別是有關於一種無線電力驅動通訊網路（Wireless Powered Communication Network，WPCN）的強化學習通訊時間分配方法及基地台。

【先前技術】

【0002】先前 WPCN 的傳輸吞吐量最佳化多將問題轉化為凸函數問題（convex problems），並利用 Lagrange 乘數（multiplier）或凸函數最佳化演算法（convex optimization algorithms）求解。然而，這些方法須假設所有通道參數（例如通道增益、衰減等）為已知。

【0003】然而，實際上的通道參數會隨時間、基地台位置或通訊節點位置而改變，因此若基地台及各通訊節點的通訊時間無法得到動態調整，將可能使得 WPCN 的總吞吐量大幅下降。

【發明內容】

【0004】有鑑於此，本發明提供一種無線電力驅動通訊網路的強

化學習通訊時間分配方法及基地台，其可用於解決上述技術問題。

【0005】 本發明提供一種無線電力驅動通訊網路的強化學習通訊時間分配方法，適於一無線電力驅動通訊網路中的一基地台，其中基地台管理無線電力驅動通訊網路中的多個通訊節點，所述方法包括：取得各通訊節點在第 t 個時間區塊中的一權重向量，並模型化各通訊節點在所述第 t 個時間區塊中的一特徵向量，其中各通訊節點在所述第 t 個時間區塊的特徵向量關聯於各通訊節點在所述第 t 個時間區塊中的一通訊時間區間；基於各通訊節點在所述第 t 個時間區塊中的權重向量及特徵向量模型化各通訊節點在所述第 t 個時間區塊中的一估計吞吐量，並據以模型化前述通訊節點在所述第 t 個時間區塊中的一總估計吞吐量；依據關聯於總估計吞吐量的一目標函數決定對應於所述第 t 個時間區塊的一通訊時間分配，其中通訊時間分配包括基地台及前述通訊節點個別在所述第 t 個時間區塊中的通訊時間區間，且通訊時間分配滿足至少一限制條件；將所述第 t 個時間區塊的通訊時間分配通知前述通訊節點，並要求各通訊節點在所述第 t 個時間區塊中依據對應的通訊時間區間執行一特定通訊行為；取得各通訊節點在所述第 t 個時間區塊中的一實際吞吐量；依據各通訊節點在所述第 t 個時間區塊中的實際吞吐量、權重向量及估計吞吐量產生各通訊節點在第 $t+1$ 個時間區塊中的權重向量。

【0006】 本發明提供一種基地台，其屬於一無線電力驅動通訊網

路，並管理無線電力驅動通訊網路中的多個通訊節點，基地台經配置以：取得各通訊節點在第 t 個時間區塊中的一權重向量，並模型化各通訊節點在所述第 t 個時間區塊中的一特徵向量，其中各通訊節點在所述第 t 個時間區塊的特徵向量關聯於各通訊節點在所述第 t 個時間區塊中的一通訊時間區間；基於各通訊節點在所述第 t 個時間區塊中的權重向量及特徵向量模型化各通訊節點在所述第 t 個時間區塊中的一估計吞吐量，並據以模型化前述通訊節點在所述第 t 個時間區塊中的一總估計吞吐量；依據關聯於總估計吞吐量的一目標函數決定對應於所述第 t 個時間區塊的一通訊時間分配，其中通訊時間分配包括基地台及前述通訊節點個別在所述第 t 個時間區塊中的通訊時間區間，且通訊時間分配滿足至少一限制條件；將所述第 t 個時間區塊的通訊時間分配通知前述通訊節點，並要求各通訊節點在所述第 t 個時間區塊中依據對應的通訊時間區間執行一特定通訊行為；取得各通訊節點在所述第 t 個時間區塊中的一實際吞吐量；依據各通訊節點在所述第 t 個時間區塊中的實際吞吐量、權重向量及估計吞吐量產生各通訊節點在第 $t+1$ 個時間區塊中的權重向量。

【圖式簡單說明】

【0007】

圖 1 是依據本發明之一實施例繪示的 WPCN 系統示意圖。

圖 2A 是依據本發明第一實施例繪示的 WPCN 系統示意圖。

圖 2B 是依據本發明第二實施例繪示的 WPCN 系統示意圖。

圖 2C 是依據本發明第三實施例繪示的 WPCN 系統示意圖。

圖 2D 是依據本發明第四實施例繪示的 WPCN 系統示意圖。

圖 3 是依據本發明之一實施例繪示的 WPCN 的強化學習通訊時間分配方法流程圖。

【實施方式】

【0008】 請參照圖 1，其是依據本發明之一實施例繪示的 WPCN 系統示意圖。如圖 1 所示，WPCN 系統 100 包括基地台 110 及多個通訊節點 121~12N（N 為通訊節點的總數），其中基地台 110 可用於管理通訊節點 121~12N。在一些實施例中，基地台 110 可用於(同時)傳遞能量 P_H 至通訊節點 121~12N，以對通訊節點 121~12N 進行充電。另一方面，通訊節點 121~12N 可在所分配的通訊時間區間分別發送資料 $P_1 \sim P_N$ 至基地台 110，但可不限於此。

【0009】 在本發明的實施例中，WPCN 系統 100 係假設為基於先獲取能量再傳送資料協定（harvest-then-transmit protocol）運作。亦即，在一時間區塊（time block）中，基地台 110 將先對各通訊裝置 121~12N 進行充電（即，能量獲取（energy harvest）），之後再由各通訊裝置 121~12N 在對應的通訊時間區間內傳送資料至基地台 110。為便於說明，基地台 110 在第 t 個時間區塊中對各通訊裝置 121~12N 進行充電的通訊時間區間表示為 $\tau_H(t)$ （其大於等於 0），而通訊節點 121~12N 在第 t 個時間區間內所佔用的總傳輸時

間表示為 $\tau_T(t)$ ，其中 $\tau_H(t)$ 與 $\tau_T(t)$ 的總和假設為一個時間區塊的長度。在一實施例中，為便於說明，假設一個時間區塊的長度為 1 (即 $\tau_H(t) + \tau_T(t) = 1$)，但可不限於此。

【0010】 另外，通訊節點 121~12N 中的第 n 個通訊節點 (以下以通訊節點 12n 代稱) 在第 t 個時間區間中從基地台 110 獲取能量 (即，被基地台 110 充電) 的時間以 $\tau_{0,n}(t)$ 表示，而通訊節點 12n 在第 t 個時間區間中的通訊時間區間以 $\tau_n(t)$ 表示。

【0011】 在不同的實施例中， $\tau_H(t)$ 、 $\tau_T(t)$ 、 $\tau_{0,n}(t)$ 及 $\tau_n(t)$ 的態樣會依 WPCN 系統 100 的組態而有所不同，以下將另輔以圖 2A 至圖 2D 作進一步說明。

【0012】 請參照圖 2A，其是依據本發明第一實施例繪示的 WPCN 系統示意圖。在第一實施例中，假設基地台 110 僅具有 1 根天線，因此基地台 110 同一時間僅能傳遞能量或從通訊裝置 121~12N 的其中之一接收資料。另外，假設第 t 個時間區塊被均勻分配予基地台 110 及通訊裝置 121~12N，因此通訊裝置 121~12N 的 $\tau_{0,1}(t) \sim \tau_{0,N}(t)$ 及基地台 110 的 $\tau_H(t)$ 可皆相等，且 $\tau_T(t)$ 可以是 $\tau_1(t) \sim \tau_N(t)$ 的總和，如圖 2A 所示。

【0013】 請參照圖 2B，其是依據本發明第二實施例繪示的 WPCN 系統示意圖。在第二實施例中，假設基地台 110 具有 2 根天線，亦即基地台 110 可同時以此二天線分別進行向通訊節點 121~12N 傳送能量及從通訊節點 121~12N 之一接收資料。

【0014】 另，假設各通訊節點 121~12N 具有睡眠模式，亦即通訊

節點 12n 在對應的 $\tau_{0,n}(t)$ 之後即不會再從基地台 110 獲取能量。因此，對於通訊節點 121 而言，其對應的 $\tau_{0,1}(t)$ 即為 $\tau_H(t)$ ，而對於其他的通訊節點 12n 而言，其對應的 $\tau_{0,n}(t)$ 可表示為 $\tau_H(t) + \sum_{j=1}^{n-1} \tau_j(t)$ 。此外， $\tau_T(t)$ 仍可以是 $\tau_1(t) \sim \tau_N(t)$ 的總和，如圖 2B 所示。

【0015】 請參照圖 2C，其是依據本發明第三實施例繪示的 WPCN 系統示意圖。在第三實施例中，與第二實施例之間的差別僅在於第三實施例中的通訊節點 121~12N 皆不具有睡眠模式，亦即通訊節點 12n 在對應的 $\tau_{0,n}(t)$ 之後仍會從基地台 110 繼續獲取能量。因此，對於通訊節點 12n 而言，其對應的 $\tau_{0,n}(t)$ 可表示為 $\tau_H(t) + \sum_{\substack{j=1 \\ j \neq n}}^N \tau_j(t)$ 。此外， $\tau_T(t)$ 仍可以是 $\tau_1(t) \sim \tau_N(t)$ 的總和，如圖 2C 所示。

【0016】 請參照圖 2D，其是依據本發明第四實施例繪示的 WPCN 系統示意圖。在第四實施例中，假設基地台 110 具有 N 根天線，因此基地台 110 可同時接收通訊節點 121~12N 所發送的資料。在此情況下，通訊裝置 121~12N 的 $\tau_{0,1}(t) \sim \tau_{0,N}(t)$ 及基地台 110 的 $\tau_H(t)$ 可皆相等，且 $\tau_1(t) \sim \tau_N(t)$ 與 $\tau_T(t)$ 亦可皆相等，如圖 2D 所示。

【0017】 在習知技術中，若欲取得基地台 110 及通訊節點 121~12N 在第 t 個時間區塊內的最佳通訊時間分配（基地台 110 及通訊節點 121~12N 在第 t 個時間區塊內的通訊時間區間，其可表徵為 $\tau(t) = [\tau_H(t) \tau_1(t) \dots \tau_N(t)]$ ），一般須對以上第一至第四實施例所示的 WPCN 系統架構個別採用不同的演算法方能順利取得（且還需已知相關的通道參數），並無法以單一套演算法即適用於第一至第四實施例所示的 WPCN 系統架構。

【0018】然而，本發明提出的方法除了可在通道參數未知的情況下找出基地台 110 及通訊節點 121~12N 在第 t 個時間區塊內的最佳通訊時間分配（即 $\tau(t)$ ）之外，還可廣泛地適用於上述第一至第四實施例的 WPCN 系統架構。以下將對本發明的方法作進一步說明。

【0019】請參照圖 3，其是依據本發明之一實施例繪示的 WPCN 的通訊時間分配方法流程圖。本實施例的方法可由圖 1 的基地台 110 執行，以下即搭配圖 1 所示的元件說明圖 3 各步驟的細節。

【0020】首先，在步驟 S210 中，基地台 110 可取得各通訊節點 121~12N 在第 t 個時間區塊中的權重向量，並模型化各通訊節點 121~12n 在所述第 t 個時間區塊中的特徵向量。

【0021】在本發明的實施例中，對於通訊節點 12n 而言，其在第 t 個時間區塊中的權重向量及特徵向量分別可表示為 $\mathbf{w}_n(t)$ 及 $\mathbf{x}_n(t)$ ，其中 $\mathbf{x}_n(t)$ 可關聯於通訊節點 12n 在第 t 個時間區塊中的通訊時間區間。

【0022】在一實施例中， $\mathbf{w}_n(t) = [w_{n,1}(t) w_{n,2}(t) \dots w_{n,D}]$ ，且 $\mathbf{x}_n(t) = [x_1(\tau_H(t), \tau_n(t)) \dots x_D(\tau_H(t), \tau_n(t))]$ ，其中 D 為 $\mathbf{w}_n(t)$ 及 $\mathbf{x}_n(t)$ 的維度。在本發明的實施例中，步驟 S210 的細節可參照相關的現有技術文獻（例如「*R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, Massachusetts, London, England: MIT Press, 2018*」），於此不另贅述。

【0023】概略而言，在本發明的實施例中， $\mathbf{w}_n(t)$ 可藉由更新 $\mathbf{w}_n(t -$

1)而得，而相關的更新機制將在之後詳述。另外，對於第 1 個時間區塊而言，其對應的 $\mathbf{w}_n(1)$ 可由設計者依考量而產生（例如隨機產生等），但可不限於此。

【0024】 接著，在步驟 S220 中，基地台 110 可基於各通訊節點 121~12N 在第 t 個時間區塊中的權重向量及特徵向量模型化各通訊節點 121~12N 在第 t 個時間區塊中的估計吞吐量，並據以模型化通訊節點 121~12N 在第 t 個時間區塊中的總估計吞吐量。

【0025】 在一實施例中，各通訊節點 121~12N 在第 t 個時間區塊中的估計吞吐量可表示為 $\hat{R}_n(t)$ ，且其可模型化為 $\hat{R}_n(t) = \mathbf{w}_n(t) \mathbf{x}_n^T(t) = \hat{R}_n(\tau_H(t), \tau_n(t), \mathbf{w}_n(t)) = \sum_{d=1}^D w_{n,d}(t) x_d(\tau_H(t), \tau_n(t))$ 。另外，通訊節點 121~12N 在第 t 個時間區塊中的總估計吞吐量可表示為 $\hat{R}(t)$ ，而其可相應地模型化為 $\hat{R}(t) = \sum_{n=1}^N \hat{R}_n(t)$ ，但可不限於此。

【0026】 在其他實施例中，上述總估計吞吐量（即， $\hat{R}(t)$ ）還可依設計者的考量而以不同的方式進行模型化。例如，習知的 WPCN 相關文獻並無考量各通訊節點 121~12N 的電池壽命，然而當電池壽命終了，整個 WPCN 將無法繼續使用。因此，假設設計者欲令所決定的 $\tau(t)$ 能夠進一步考量並延長各通訊節點 121~12N 的電池壽命，則 $\hat{R}(t)$ 例如可相應地調整為 $\hat{R}(t) = \sum_{n=1}^N \hat{R}_n(t) - \beta \sum_{n=1}^N SoC_n(t)$ ，其中， β 為一權重係數，而 $SoC_n(t)$ 為通訊節點 12n 在第 t 個時間區塊中所獲得的電量。在本發明的實施例中， β 及 $SoC_n(t)$ 的相關說明可參考相關的現有技術文獻（例如「*P. Shen, M. Ouyang, L. Lu, J. Li, and X. Feng, "The co-estimation of state of charge, state of health,*

and state of function for lithium-ion batteries in electric vehicles,” *IEEE Trans. Veh. Technol.*, vol. 67, no. 1, pp. 92–103, Jan. 2018」),

故其細節於此不另贅述。

【0027】 之後，在步驟 S330 中，基地台 110 可依據關聯於總估計吞吐量（即， $\hat{R}(t)$ ）的目標函數決定對應於第 t 個時間區塊的通訊時間分配（即， $\tau(t)$ ）。

【0028】 在一實施例中，上述目標函數例如包括 $\max_{\tau(t)} \hat{R}(t)$ （即，最大化 $\hat{R}(t)$ ），且上述限制條件例如包括：「 $\tau_H(t) + \tau_T(t) = 1$ 」、「 $\tau_H(t) \geq 0, \tau_T(t) \geq 0$ 」及「 $\tau_n(t) \geq 0$ 」，但可不限於此。

【0029】 在其他實施例中，上述目標函數及限制條件可依設計者的考量而進行調整。舉例而言，習知 WPCN 文獻較少探討各通訊節點 121~12N 的傳輸公平性的問題，然而在 WPCN 中，若無考量到傳輸公平性，則在多個通訊節點的情況下，距離基地台 110 較遠的通訊節點僅能獲得少量傳輸時間，進而可能導致距離較遠的通訊節點的吞吐量遠少於其他距離較近的通訊節點。

【0030】 因此，在一些實施例中，上述限制條件還可進一步包括「 $\hat{R}_n(t) \geq \bar{R}_n$ 」，其中 \bar{R}_n 為通訊節點 12n 的吞吐量下限值。如此一來，基地台 110 所取得的 $\tau(t)$ 即可進一步考量並保證通訊節點 121~12N 之間的傳輸公平性，但可不限於此。

【0031】 在本發明的實施例中，透過步驟 S330 所取得的 $\tau(t)$ 可理解為滿足上述限制條件且可最大化 $\hat{R}(t)$ 的最佳通訊時間分配。為便於說明，以下將以 $\tau^*(t)$ 代稱步驟 S330 中所取得的 $\tau(t)$ ，但可不限

於此。

【0032】此外，在一些實施例中，為避免所取得的 $\tau(t)$ 出現過度擬合或落入局部最優解等情形，在步驟S330之前，基地台110還可依據一 ε -貪婪策略（ ε -greedy policy）判斷是否依據關聯於總估計吞吐量（即， $\hat{R}(t)$ ）的目標函數決定對應於第 t 個時間區塊的通訊時間分配（即， $\tau(t)$ ）。若是，則基地台110可依據關聯於總估計吞吐量的目標函數決定對應於第 t 個時間區塊的通訊時間分配；若否，則基地台110可隨機產生基地台110及通訊節點121~12N個別在第 t 個時間區塊中的通訊時間區間，以決定對應於第 t 個時間區塊的通訊時間分配，其中上述通訊時間分配滿足上述限制條件。

【0033】簡言之，基地台110可依據上述 ε -貪婪策略決定是否執行步驟S330。具體而言，若採用上述 ε -貪婪策略，則基地台110有 ε （其例如是一極小值）的機率會選擇不執行步驟S330而以隨機方式決定 $\tau(t)$ （但所決定的 $\tau(t)$ 仍須滿足所設定的限制條件），而有 $1-\varepsilon$ 的機率會選擇執行步驟S330來決定 $\tau(t)$ （即，先前提及的 $\tau^*(t)$ ）。藉此，可避免所取得的 $\tau(t)$ 出現過度擬合或落入局部最優解等情形，但可不限於此。

【0034】此外，在一實施例中，當 t 等於1時，基地台110亦可基於上述隨機方式來決定 $\tau(1)$ ，而當 t 大於1時，基地台110可依上述教示決定 $\tau(t)$ ，但本發明可不限於此。

【0035】之後，在步驟S340中，基地台110可將第 t 個時間區塊的通訊時間分配（即， $\tau(t) = [\tau_H(t) \tau_1(t) \dots \tau_N(t)]$ ）通知通訊節點

121~12N，並要求各通訊節點 121~12n 在第 t 個時間區塊中依據對應的通訊時間區間執行特定通訊行為（例如從基地台 110 獲取能量及/或發送資料至基地台）。

【0036】 接著，在步驟 S350 中，基地台 110 可取得各通訊節點 121~12N 在第 t 個時間區塊中的實際吞吐量。亦即，基地台 110 可實際測量各通訊節點 121~12N 在所分配的通訊時間區間（即 $\tau_1(t) \sim \tau_N(t)$ ）內傳送的資料量。為便於說明，通訊節點 12n 在第 t 個時間區塊中的實際吞吐量可表示為 $R_n(t)$ 。

【0037】 之後，在步驟 S360 中，基地台 110 可依據各通訊節點 121~12N 在第 t 個時間區塊中的實際吞吐量、權重向量及估計吞吐量產生各通訊節點 121~12N 在第 $t+1$ 個時間區塊中的權重向量。

【0038】 在一實施例中，基地台 110 可依據各通訊節點 121~12N 在第 t 個時間區塊中的實際吞吐量、權重向量及估計吞吐量執行一隨機梯度下降（stochastic gradient decent，SGD）法，以產生各通訊節點 121~12n 在第 $t+1$ 個時間區塊中的權重向量。

【0039】 舉例而言，通訊節點 12n 在第 $t+1$ 個時間區塊中的權重向量例如可表徵為 $\mathbf{w}_n(t+1)$ 。在一實施例中， $\mathbf{w}_n(t+1)$ 可表徵為：

$$\begin{aligned} \mathbf{w}_n(t+1) &= \mathbf{w}_n(t) + \frac{1}{2} \alpha \nabla (R_n(t) - \hat{R}_n(t))^2 \\ &= \mathbf{w}_n(t) + \frac{1}{2} \alpha \nabla (R_n(t) - \hat{R}_n(\tau_H(t), \tau_N(t), \mathbf{w}_n(t)))^2 \\ &= \mathbf{w}_n(t) + \alpha [R_n(t) - \hat{R}_n(\tau_H(t), \tau_N(t), \mathbf{w}_n(t))] \mathbf{x}_n(t) \end{aligned}$$

，其中 α 為一步階尺寸值， $\nabla(\cdot)$ 為一梯度運算子。

【0040】 在一實施例中，在取得 $\mathbf{w}_n(t+1)$ 之後，基地台 110 還可再

次依據先前實施例的教示決定第 $t+1$ 個時間區塊中的通訊時間分配（即， $\boldsymbol{\tau}(t+1)$ ），並據以要求各通訊節點 $121\sim 12N$ 在第 $t+1$ 個時間區塊中依據對應的通訊時間區間執行特定通訊行為。之後，基地台 110 同樣可取得各通訊節點 $121\sim 12N$ 在第 $t+1$ 個時間區塊中的實際吞吐量，並相應地產生各通訊節點在第 $t+2$ 個時間區塊中的權重向量，而相關細節可參照先前實施例中的說明，於此不另贅述。

【0041】 經驗證，隨著 t 的增加， $E[(R_n(t) - \hat{R}_n(\tau_H(t), \tau_N(t), w_n(t)))^2]$ （即 $\hat{R}(t)$ 與 $\hat{R}_n(t)$ 的均方誤差（mean square error, MSE））的值將逐漸下降。亦即，隨著時間的推移，經由本發明方法所決定的 $\boldsymbol{\tau}(t)$ 可讓通訊節點 $12n$ 的實際吞吐量（即， $R_n(t)$ ）逐漸接近通訊節點 $12n$ 的估計吞吐量（即， $\hat{R}_n(t) = \mathbf{w}_n(t)\mathbf{x}_n^T(t)$ ）。

【0042】 綜上所述，除了不需將問題轉化為凸函數問題之外，本發明提出的方法及基地台還可在通道參數為未知的情況下找出各時間區塊中的最佳時間分配，且可廣泛地適用於各式 WPCN 系統架構。並且，透過適當地在限制條件中引入各通訊節點的吞吐量下限值，經本發明方法所決定的 $\boldsymbol{\tau}(t)$ 可保證各通訊節點之間的傳輸公平性，從而避免資料傳輸量過度集中於離基地台較近的通訊節點上。此外，透過在通訊節點的總估計吞吐量中引入與電量相關的參數（即， $\beta \sum_{n=1}^N SoC_n(t)$ ），經本發明方法所決定的 $\boldsymbol{\tau}(t)$ 可進一步考量各通訊節點的電池壽命。

【0043】 雖然本發明已以實施例揭露如上，然其並非用以限定本

發明，任何所屬技術領域中具有通常知識者，在不脫離本發明的精神和範圍內，當可作些許的更動與潤飾，故本發明的保護範圍當視後附的申請專利範圍所界定者為準。

【符號說明】

【0044】

100:WPCN 系統

110:基地台

121~12N:通訊裝置

S310~S360:步驟

$P_1 \sim P_N$:資料

P_H :能量

【發明申請專利範圍】

【請求項1】 一種無線電力驅動通訊網路的強化學習通訊時間分配方法，適於一無線電力驅動通訊網路中的一基地台，其中該基地台管理該無線電力驅動通訊網路中的多個通訊節點，所述方法包括：

取得各該通訊節點在第 t 個時間區塊中的一權重向量，並模型化各該通訊節點在所述第 t 個時間區塊中的一特徵向量，其中各該通訊節點在所述第 t 個時間區塊的該特徵向量關聯於各該通訊節點在所述第 t 個時間區塊中的一通訊時間區間；

基於各該通訊節點在所述第 t 個時間區塊中的該權重向量及該特徵向量模型化各該通訊節點在所述第 t 個時間區塊中的一估計吞吐量，並據以模型化該些通訊節點在所述第 t 個時間區塊中的一總估計吞吐量；

依據關聯於該總估計吞吐量的一目標函數決定對應於所述第 t 個時間區塊的一通訊時間分配，其中該通訊時間分配包括該基地台及該些通訊節點個別在所述第 t 個時間區塊中的該通訊時間區間，且該通訊時間分配滿足至少一限制條件；

將所述第 t 個時間區塊的該通訊時間分配通知該些通訊節點，並要求各該通訊節點在所述第 t 個時間區塊中依據對應的該通訊時間區間執行一特定通訊行為；

取得各該通訊節點在所述第 t 個時間區塊中的一實際吞吐量；

依據各該通訊節點在所述第 t 個時間區塊中的該實際吞吐量、

該權重向量及該估計吞吐量產生各該通訊節點在第 $t+1$ 個時間區塊中的該權重向量。

【請求項2】 如請求項1所述的方法，其中該通訊時間分配表徵為 $\boldsymbol{\tau}(t) = [\tau_H(t) \tau_1(t) \dots \tau_N(t)]$ ， $\tau_H(t)$ 為該基地台在所述第 t 個時間區塊中的該通訊時間區間， $\tau_n(t)$ 為該些通訊節點中的第 n 個通訊節點在所述第 t 個時間區塊中的該通訊時間區間，該總估計吞吐量表徵為 $\hat{R}(t)$ ，該目標函數包括 $\max_{\boldsymbol{\tau}(t)} \hat{R}(t)$ ，且該至少一限制條件包括：

$$\tau_H(t) + \tau_T(t) = 1,$$

$$\tau_H(t) \geq 0, \tau_T(t) \geq 0,$$

$$\tau_n(t) \geq 0$$

，其中 $\tau_T(t)$ 為該些通訊節點在所述第 t 個時間區間內所佔用的一總傳輸時間。

【請求項3】 如請求項2所述的方法，其中該至少一限制條件更包括 $\hat{R}_n(t) \geq \bar{R}_n$ ，其中 $\hat{R}_n(t)$ 為所述第 n 個通訊節點在所述第 t 個時間區塊中的該估計吞吐量， \bar{R}_n 為所述第 n 個通訊節點的一吞吐量下限值。

【請求項4】 如請求項2所述的方法，其中該些通訊節點中的第 n 個通訊節點在所述第 t 個時間區塊中的該權重向量、該特徵向量及該估計吞吐量分別表徵為 $\mathbf{w}_n(t)$ 、 $\mathbf{x}_n(t)$ 及 $\hat{R}_n(t)$ ，且 $\hat{R}_n(t) = \mathbf{w}_n(t) \mathbf{x}_n^T(t)$ 。

【請求項5】 如請求項4所述的方法，其中該些通訊節點在所述第t個時間區塊中的該總估計吞吐量表徵為 $\hat{R}(t)$ ，且 $\hat{R}(t) = \sum_{n=1}^N \hat{R}_n(t)$ ，其中N為該些通訊節點的總數。

【請求項6】 如請求項5所述的方法，其中：

$$\mathbf{w}_n(t) = [w_{n,1}(t) w_{n,2}(t) \dots w_{n,D}] ;$$

$$\mathbf{x}_n(t) = [x_1(\tau_H(t), \tau_n(t)) \dots x_D(\tau_H(t), \tau_n(t))] ;$$

$$\hat{R}_n(t) = \hat{R}_n(\tau_H(t), \tau_n(t), \mathbf{w}_n(t)) = \sum_{d=1}^D w_{n,d}(t) x_d(\tau_H(t), \tau_n(t)) , \text{ 其中 } D \text{ 為 } \mathbf{w}_n(t) \text{ 及 } \mathbf{x}_n(t) \text{ 的維度。}$$

【請求項7】 如請求項4所述的方法，其中該些通訊節點在所述第t個時間區塊中的該總估計吞吐量表徵為 $\hat{R}(t)$ ，且 $\hat{R}(t) = \sum_{n=1}^N \hat{R}_n(t) - \beta \sum_{n=1}^N SoC_n(t)$ ，其中N為該些通訊節點的總數， β 為一權重係數，而 $SoC_n(t)$ 為該些通訊節點中的第n個通訊節點在所述第t個時間區塊中所獲得的一電量。

【請求項8】 如請求項1所述的方法，其中該基地台在所述第t個時間區塊中的對應的該通訊時間區間中傳遞能量至該些通訊節點的至少其中之一。

【請求項9】 如請求項1所述的方法，其中該特定通訊行為包括發送資料至該基地台。

【請求項10】 如請求項1所述的方法，其中依據各該通訊節點在所述第t個時間區塊中的該實際吞吐量、該權重向量及該估計吞吐量產生各該通訊節點在第t+1個時間區塊中的該權重向量的步驟包括：

依據各該通訊節點在所述第 t 個時間區塊中的該實際吞吐量、該權重向量及該估計吞吐量執行一隨機梯度下降法，以產生各該通訊節點在第 $t+1$ 個時間區塊中的該權重向量。

【請求項11】 如請求項10所述的方法，其中該些通訊節點中的第 n 個通訊節點在所述第 t 個時間區塊中的該實際吞吐量、該權重向量及該估計吞吐量分別表徵為 $R_n(t)$ 、 $w_n(t)$ 及 $\hat{R}_n(t)$ ，所述第 n 個通訊節點在所述第 $t+1$ 個時間區塊中的該權重向量表徵為 $w_n(t+1)$ ，其中 $w_n(t+1) = w_n(t) + \frac{1}{2}\alpha\nabla(R_n(t) - \hat{R}_n(t))^2$ ， α 為一步階尺寸值， $\nabla(\cdot)$ 為一梯度運算子。

【請求項12】 如請求項1所述的方法，其中在依據關聯於該總估計吞吐量的該目標函數決定對應於所述第 t 個時間區塊的該通訊時間分配的步驟之前，所述方法更包括：

依據一 ε -貪婪策略判斷是否依據關聯於該總估計吞吐量的該目標函數決定對應於所述第 t 個時間區塊的該通訊時間分配；

若是，依據關聯於該總估計吞吐量的該目標函數決定對應於所述第 t 個時間區塊的該通訊時間分配；

若否，隨機產生該基地台及該些通訊節點個別在所述第 t 個時間區塊中的該通訊時間區間，以決定對應於所述第 t 個時間區塊的該通訊時間分配，其中該通訊時間分配滿足該至少一限制條件。

【請求項13】 如請求項1所述的方法，其中 t 大於1。

【請求項14】 如請求項1所述的方法，其中若 t 等於1，所述方法更包括隨機產生該基地台及該些通訊節點個別在所述第 t 個時間區

塊中的該通訊時間區間，其中該通訊時間分配滿足該至少一限制條件。

【請求項15】 一種基地台，其屬於一無線電力驅動通訊網路，並管理該無線電力驅動通訊網路中的多個通訊節點，該基地台經配置以：

取得各該通訊節點在第 t 個時間區塊中的一權重向量，並模型化各該通訊節點在所述第 t 個時間區塊中的一特徵向量，其中各該通訊節點在所述第 t 個時間區塊的該特徵向量關聯於各該通訊節點在所述第 t 個時間區塊中的一通訊時間區間；

基於各該通訊節點在所述第 t 個時間區塊中的該權重向量及該特徵向量模型化各該通訊節點在所述第 t 個時間區塊中的一估計吞吐量，並據以模型化該些通訊節點在所述第 t 個時間區塊中的一總估計吞吐量；

依據關聯於該總估計吞吐量的一目標函數決定對應於所述第 t 個時間區塊的一通訊時間分配，其中該通訊時間分配包括該基地台及該些通訊節點個別在所述第 t 個時間區塊中的該通訊時間區間，且該通訊時間分配滿足至少一限制條件；

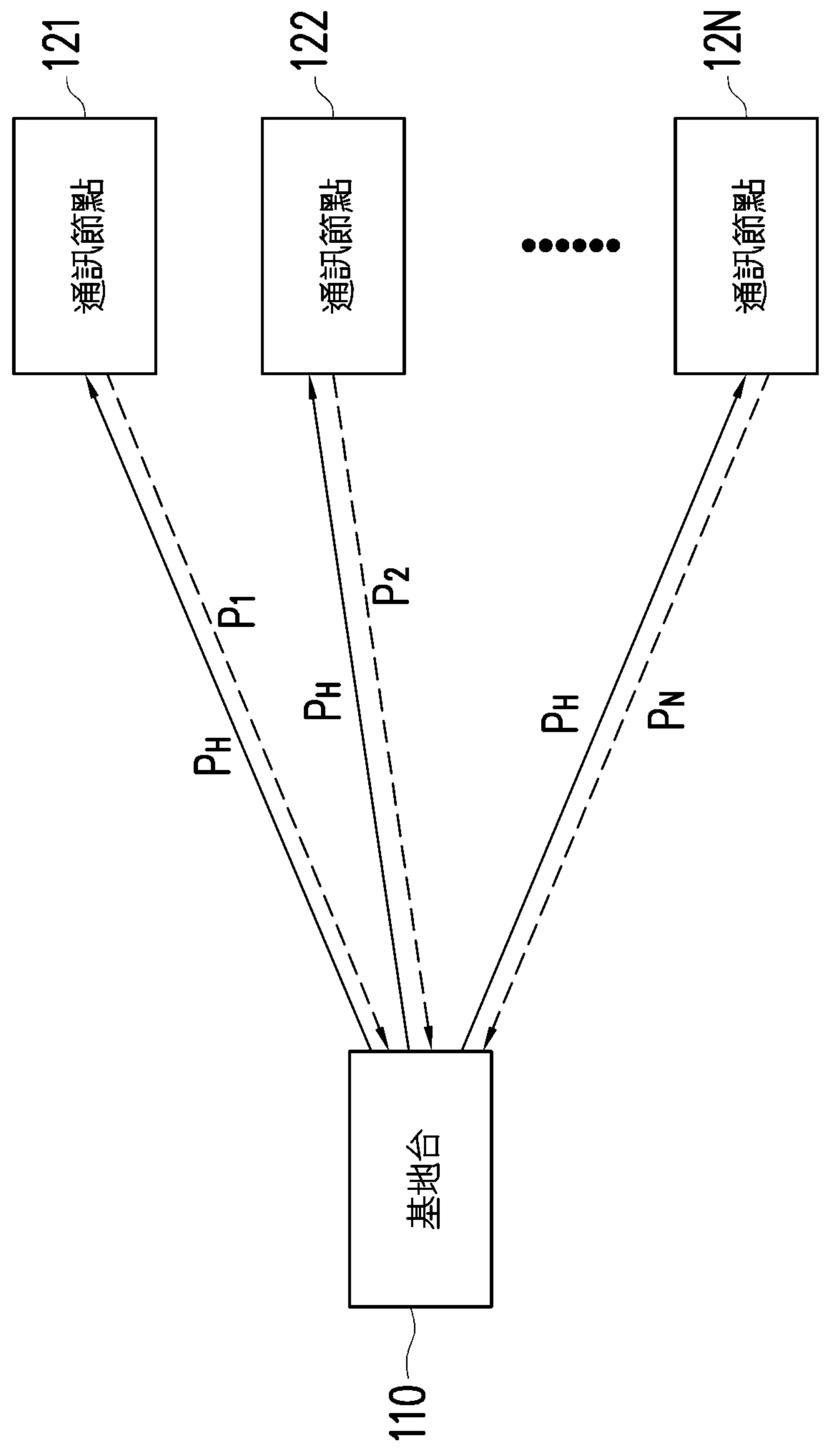
將所述第 t 個時間區塊的該通訊時間分配通知該些通訊節點，並要求各該通訊節點在所述第 t 個時間區塊中依據對應的該通訊時間區間執行一特定通訊行為；

取得各該通訊節點在所述第 t 個時間區塊中的一實際吞吐量；

依據各該通訊節點在所述第 t 個時間區塊中的該實際吞吐量、

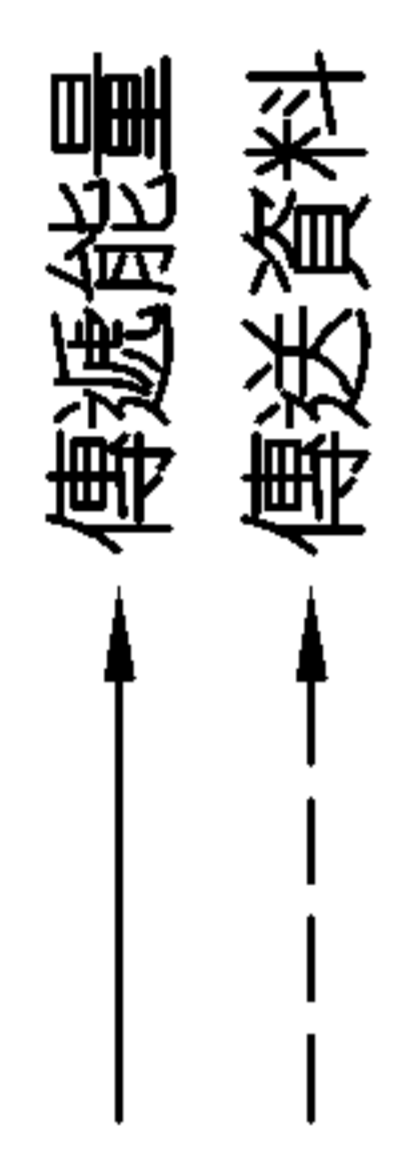
該權重向量及該估計吞吐量產生各該通訊節點在第 $t+1$ 個時間區塊中的該權重向量。

【發明圖式】

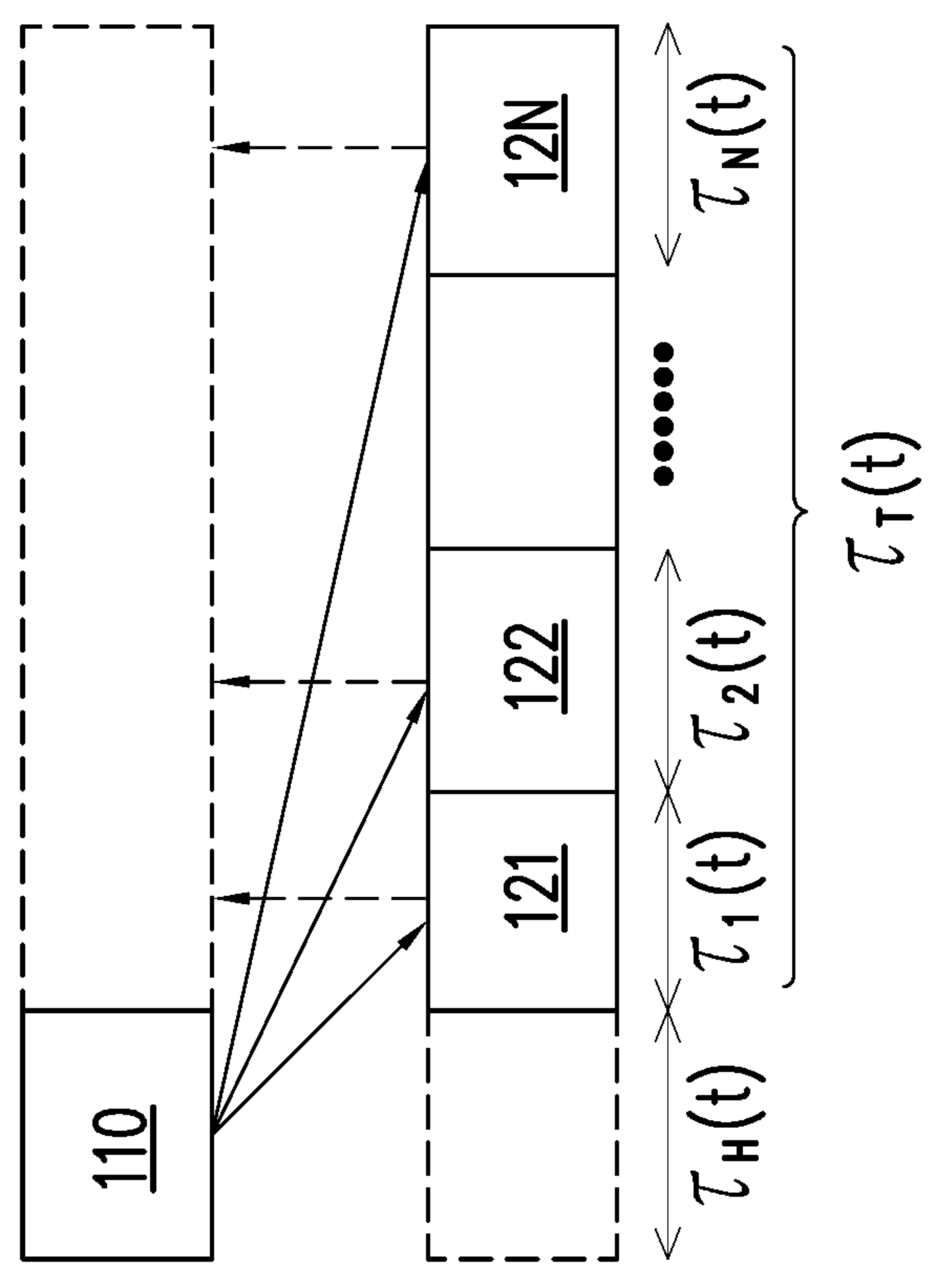


【圖1】

100



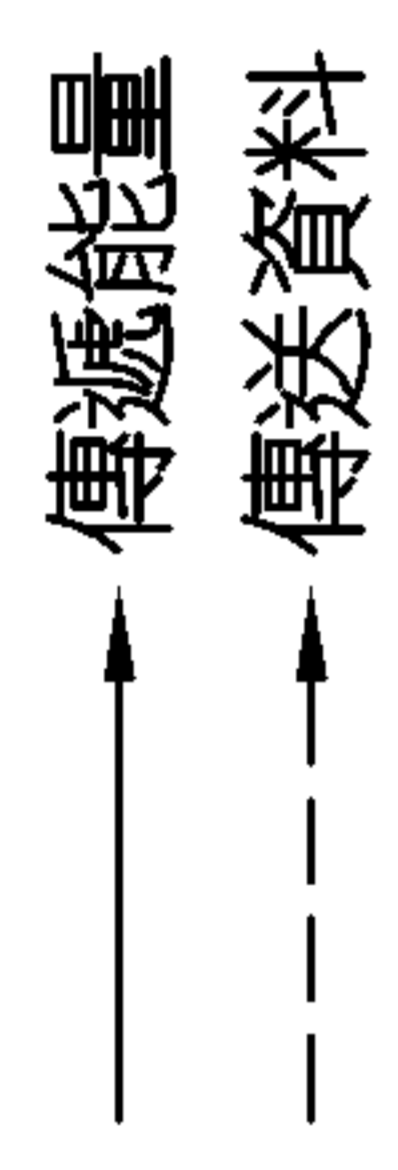
(基地台只有 1 根天線)



$$\tau_{0,1}(t) = \tau_{0,2}(t) = \dots = \tau_{0,N}(t) = \tau_H(t)$$

$$\tau_T(t) = \tau_1(t) + \dots + \tau_N(t)$$

【圖2A】

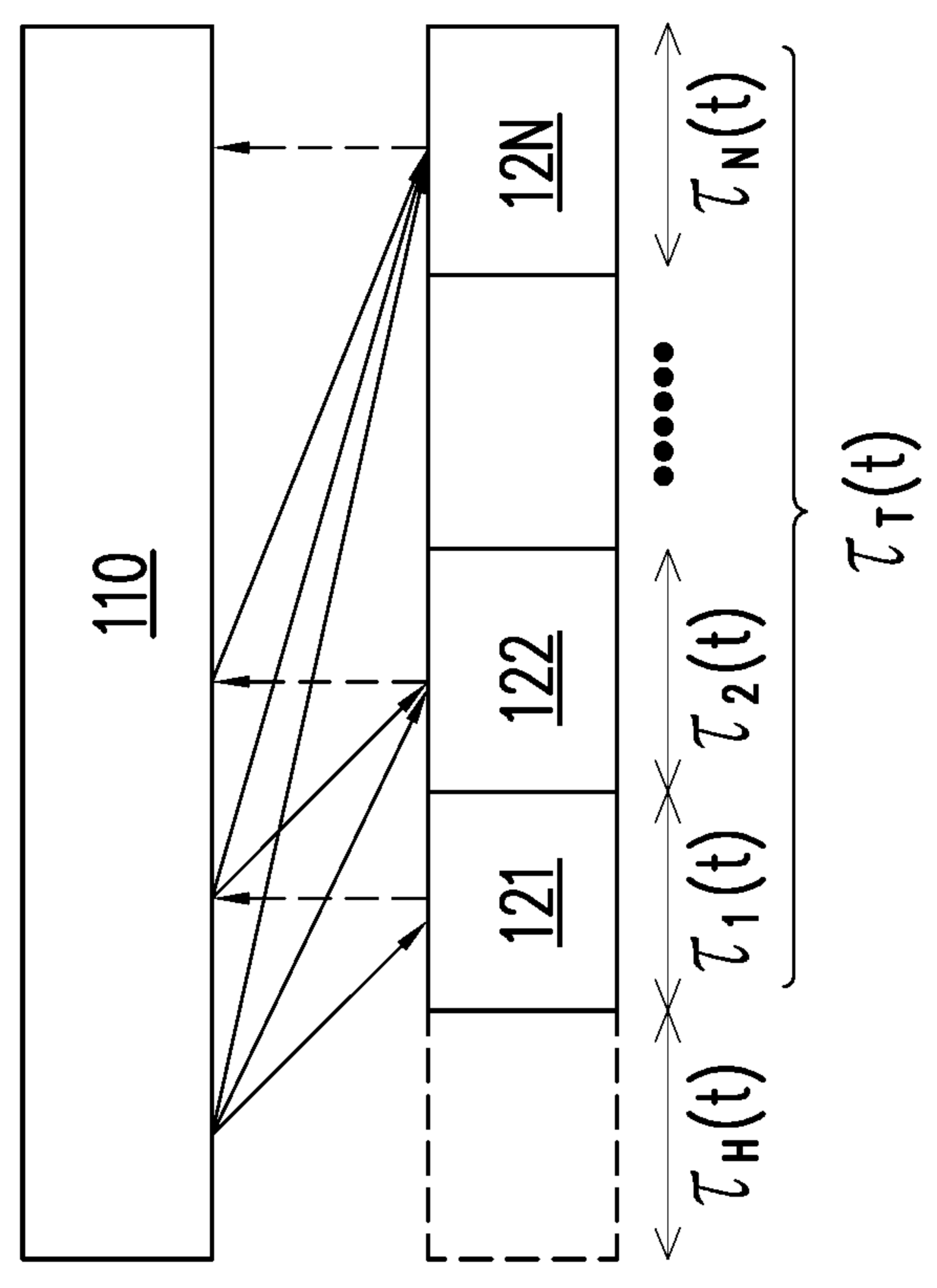


$$\tau_{0,1}(t) = \tau_H(t)$$

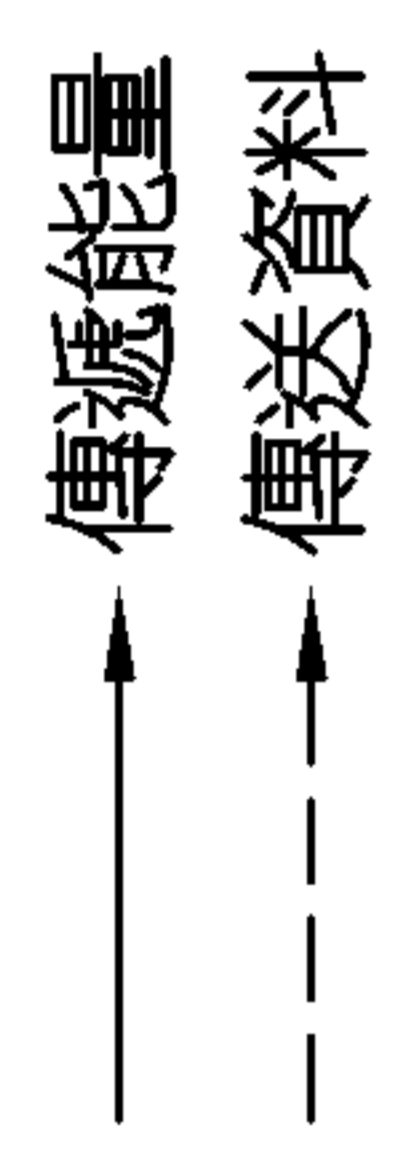
$$\tau_{0,n}(t) = \tau_H(t) + \sum_{j=1}^{n-1} \tau_j(t), n=2, \dots, N$$

$$\tau_T(t) = \tau_1(t) + \dots + \tau_N(t)$$

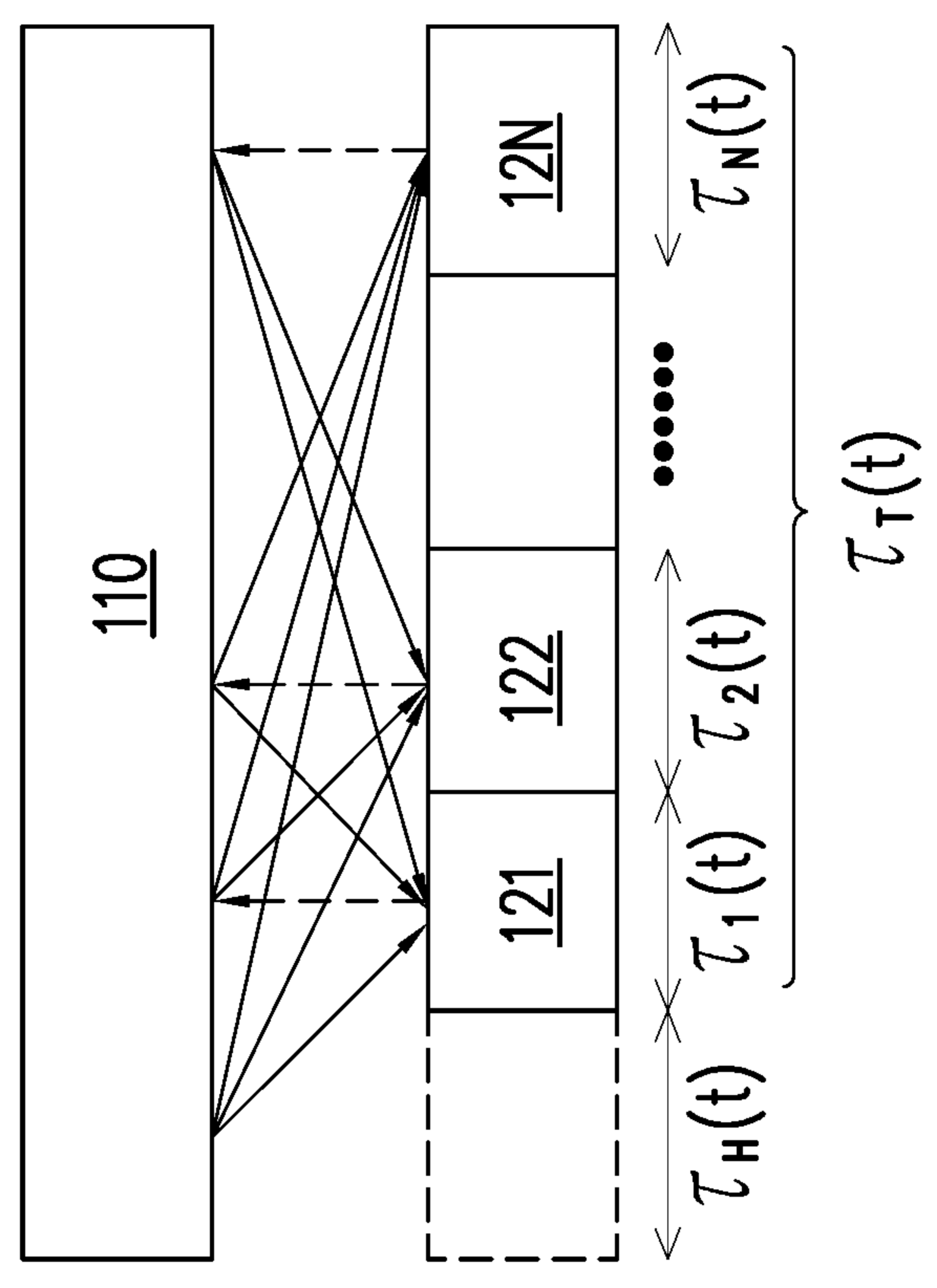
(基地台有 2 根天線，且通訊節點具有睡眠模式)



【圖2B】



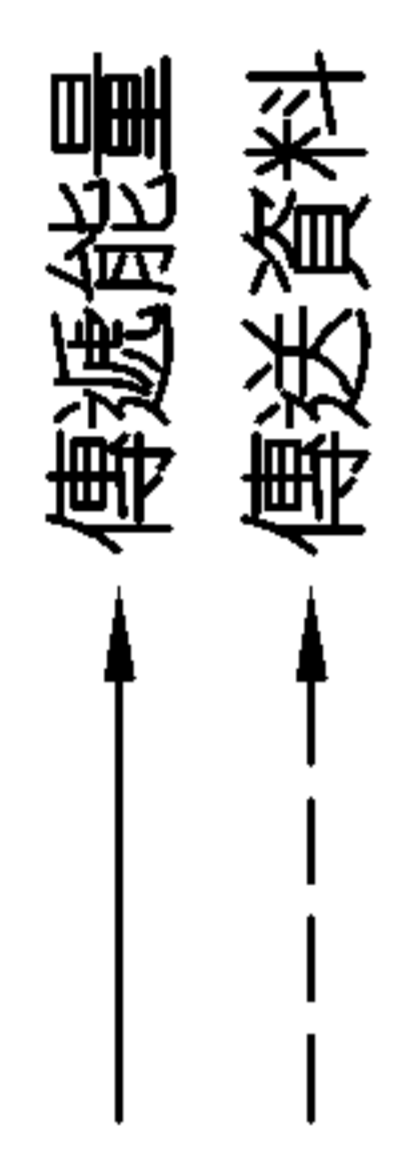
(基地台有 2 根天線，且通訊節點不具有睡眠模式)



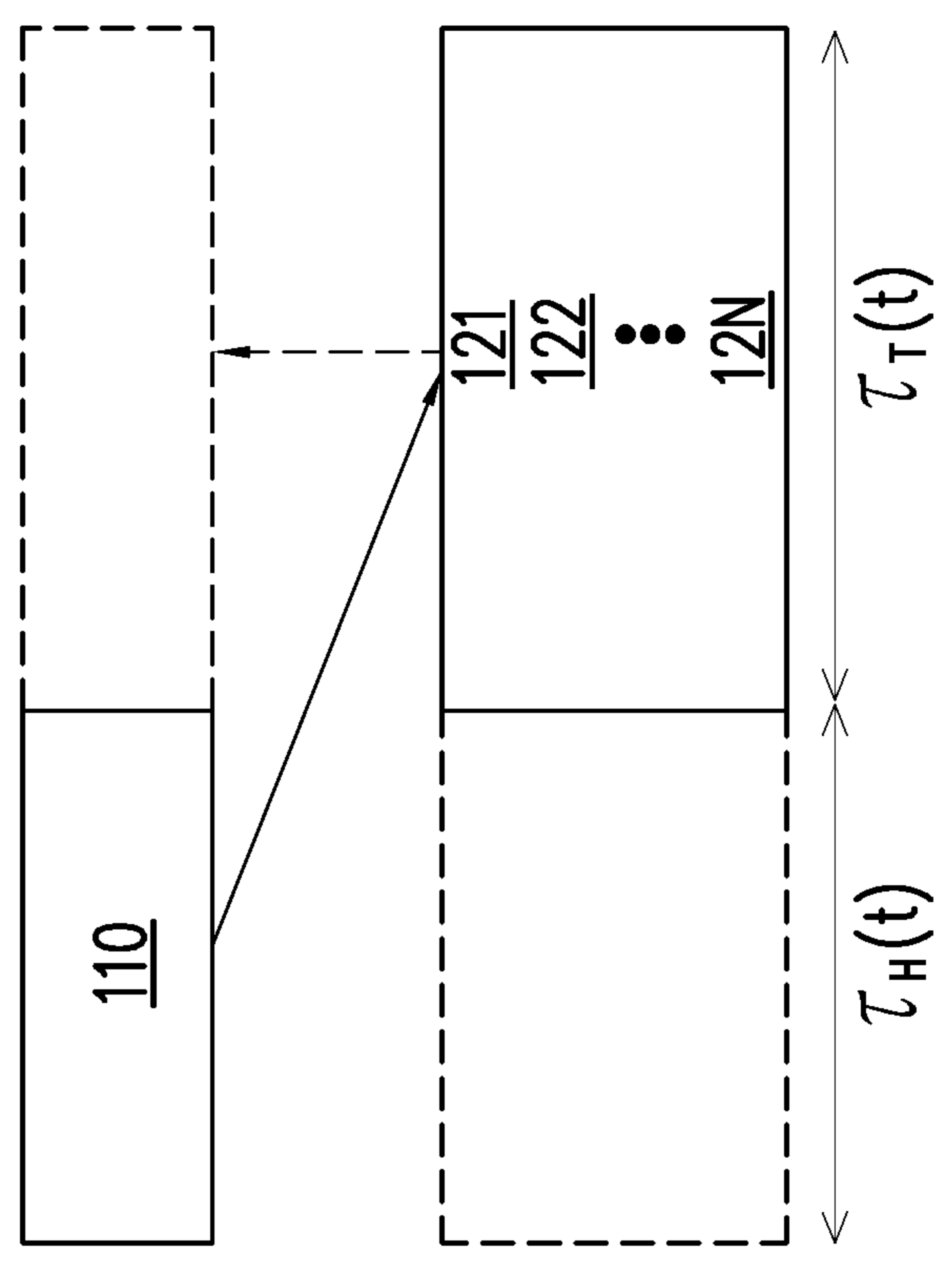
$$\tau_{0,n}(t) = \tau_H(t) + \sum_{j \neq n}^N \tau_j(t)$$

$$\tau_T(t) = \tau_1(t) + \dots + \tau_N(t)$$

【圖2C】



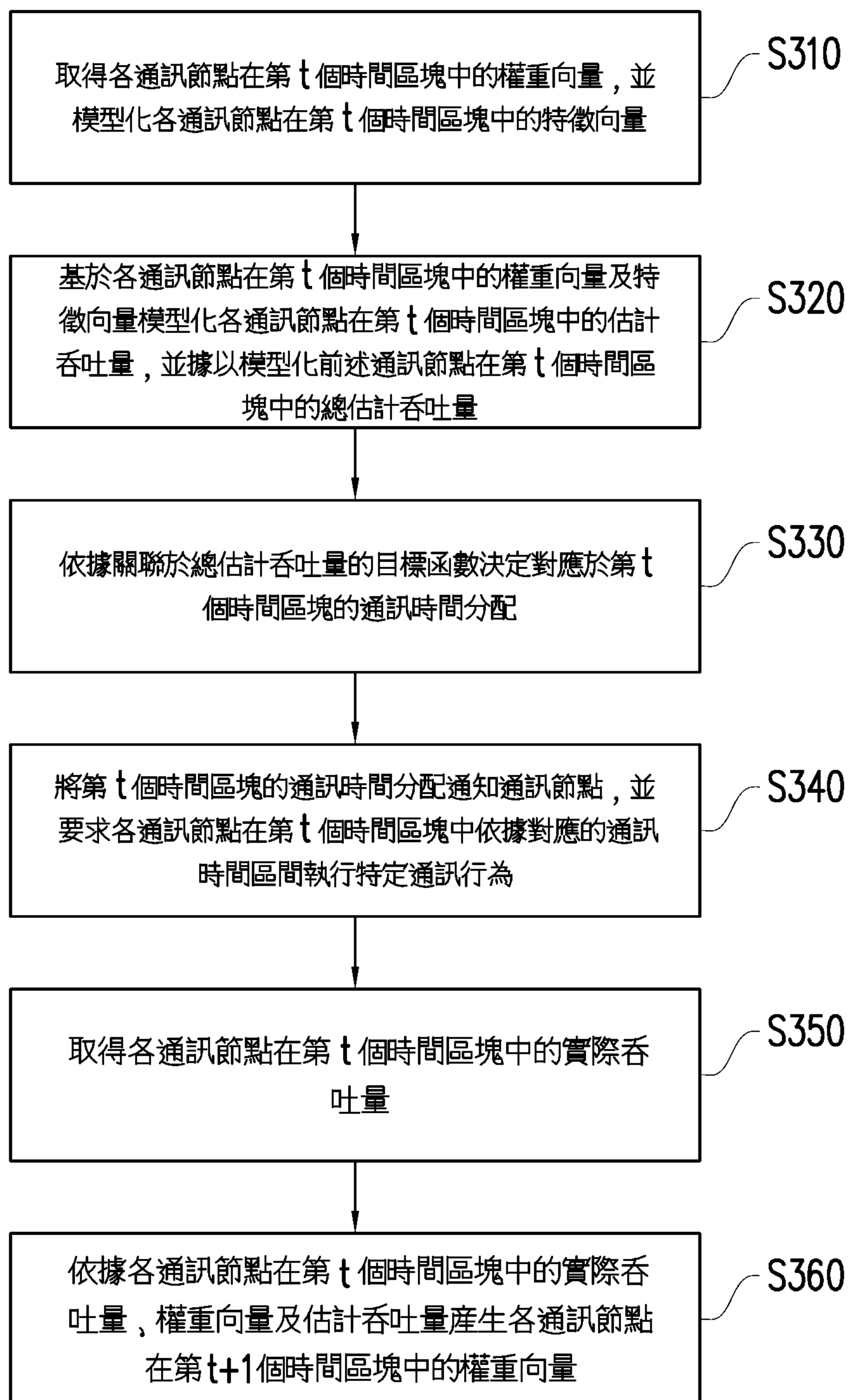
(基地台有N根天線)



$$\tau_{0,1}(t) = \tau_{0,2}(t) = \dots = \tau_{0,N}(t) = \tau_H(t)$$

$$\tau_T(t) = \dots = \tau_N(t) = \tau_T(t)$$

【圖2D】



【圖3】