

在無線傳感器網路中之四分位法容錯機制 Improves Fault Tolerant by Using Quartile Analysis in Wireless Sensor and Actor Networks

吳亦超 段裘慶 張瑋修
Y.-C. Wu C.-C. Tuan W.-S. Chang
國立臺北科技大學 電腦與通訊研究所
t5419001@ntut.edu.tw cctuan@ntut.edu.tw t6418084@ntut.edu.tw

摘要

無線傳感器網路為新型態的異質性無線感測網路，其中，如何對錯誤的感測資料作分析與回覆仍是重要的研究議題之一。然而大部分研究都集中在連結容錯方面，在感測資料的容錯無太多的分析與探討。為此，本文提出四分位法容錯機制。此機制主要利用感測資料的離散程度，重新研判可能為正確資料的範圍，以排除掉錯誤的感測資料。透過四分位法容錯機制，除了可排除錯誤的感測資料，以提供給目的地端外，更可與連結的容錯機制作結合，以形成一完整的容錯機制。

關鍵字：無線傳感器網路、感測資料、四分位法、容錯機制。

Abstract

Wireless Sensor and Actor Networks (WSAN) is a new wireless heterogeneous sensor networks. In WSAN, an important issue is to analyze and respond the fault sensed data. The fault tolerant research on link is popular, but on sensed data is less. This paper proposed a fault tolerance based on the quartiles method (FTQM). This policy utilizes the degree of dispersion on sensed data to locate the range of correct sensing data, and to remove the error data. FTQM not only removes the error sensed data, but also

constructs a complete fault tolerant mechanism by integrating combining with the fault tolerant of link.

Keyword: WSAN, sensed data, quartile, fault tolerant.

1. 簡介

無線傳感器網路(Wireless Sensor and Actor Networks, WSAN)[1,2]為一異質性無線感測網路，主要由感測節點(Sensor Node, SN)、反應節點(Actor)、與資料彙集端(Sink)所組成。其中，感測節點具有低電量、低功率消耗、低花費的特性，其主要功能為環境監測，並將感測到的資料傳送至反應節點或資料彙集端。反應節點為電量較大、傳輸距離較長、且具有強大計算能力的無線裝置，其主要功能為依據感測節點傳送過來的資料作適當的判斷和處理。資料彙集端主要功能為蒐集感測節點送來的感測資料。

根據反應節點功能的不同，WSAN 主要分為兩種架構，第一種為半自動式架構(Semi-Automated Architecture)，此架構是感測節點先將感測到的資料傳送至資料彙集端，資料彙集端再將命令送到適當的反應節點作處理，如圖 1 所示。第二種為自動式架構(Automated Architecture)，在此架構下，感測節點直接將感測資料傳送到適當的反應節點，反應節點根據接收到的感測資料直接作適當的

處理，如圖二所示。本論文採用自動式架構作為 WSN 的基本架構。

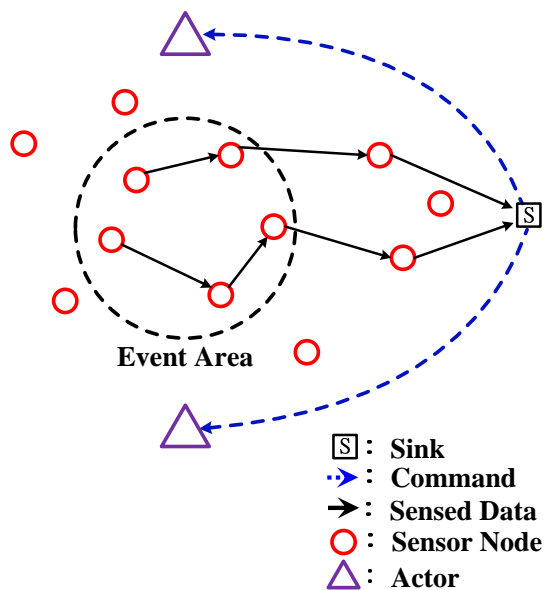


圖 1 Semi-Automated Architecture

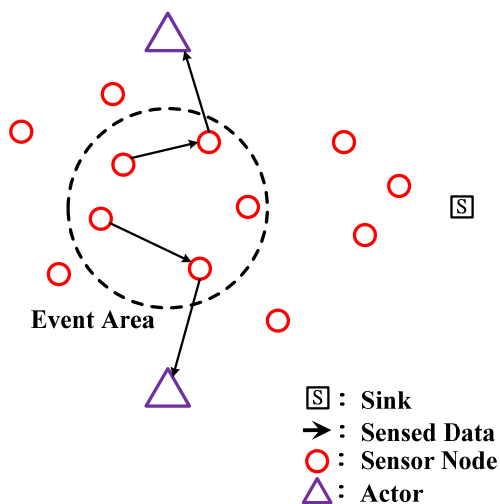


圖 2 Automated Architecture

在 WSN 環境中，容錯機制一直是很重要的議題，其主要目的是在感測節點發生錯誤或傳送資料中發生遺失或錯誤時，能對錯誤的資料作分析與回覆，以維持網路的正常運作[3,4]。因此便有學者利用新型態的 WSN 提出新的容錯機制，以改善 WSN 既有的容錯問題。在 WSN 中，大部分的容錯機制都是針對網路連結方面作探討[5,6]，對於因感測方面而

產生的錯誤感測資料並無太多的分析與研究。然而錯誤的感測資料卻可能讓反應節點造成錯誤的判斷或處理，例如：在灑水系統中，少部分的感測節點因為感測功能的故障而傳送錯誤的濕度資料，造成灑水器無法提供適當的水量給草坪，而造成草坪的死亡。因此本論文利用反應節點強大的計算能力與四分位法的特性，對接收到的感測資料作處理、分析，來找出正確的資料群，並排除故障的感測節點，讓反應節點能正確作出適當的判斷與處理動作。

論文的第二節將介紹相關容錯機制之研究；第三節詳細敘述四分位數與四分位法容錯機制；第四節說明模擬環境建立與評量因子；第五節為結論與未來工作。

2. 相關容錯機制

在無線感測網路或無線傳感器網路的環境中，容錯一直是重要的研究議題之一。在[7]中，作者將容錯分為三個階段，第一個是錯誤預測階段，該階段分為詳盡預測和暗示預測兩個方法。在詳盡預測中當事件被預測時，感測應用程式就會發出警報並直接處理該錯誤。暗示預測是針對感測節點在通訊方面的失效方面討論，其中分為主動模式和被動模式。主動模式是透過感測網路的中央控制器週期性的發送訊息來確定中央控制器周圍節點是否還存在。被動模式是當節點錯誤被預測出來時，中央控制器才通知錯誤節點附近的感測節點。

第二階段為錯誤診斷階段，該階段是藉由網路系統正確的分辨錯誤，或透過其他功能的警報來區別錯誤。第三階段為錯誤恢復階段，主要透過感測網路的重建和重新設定來確保錯誤節點不會影響到預期的網路效能。

在[9]中，作者提出平均法來找尋錯誤節點。為了減少網路傳輸的數量，延長網路的壽

命，作者將整個網路劃分為網格的形式，每個網格內都有聚集節點。當事件發生時，叢集首會先蒐集網格內感測到事件的節點資料，同時以中位數為準，剔除掉與中位數差異過大的節點資料，再將剩下的資料一起傳送到反應節點。

K.Ozaki 在[8]中提出 WSN 的容錯模型。作者假設感測節點與反應節點之間的通訊是強健的，並將錯誤節點定義為 Byzantine Fault。也就是當事件發生時，部分感測節點有可能無法傳送感測資料到反應節點，或是傳送與正常感測節點不同的事件感測資料。因此，K.Ozaki 將感測資料統一由反應節點蒐集並處理，反應節點將蒐集到的資料取平均值，並根據預設的容許誤差值，作錯誤節點的判斷[8]。當該節點的感測資料與平均值的誤差大於容許誤差值時，該感測節點就判定為錯誤節點。如圖 3，當事件發生時，SN₁ 至 SN₅ 感測到的資料分別為 15.1°C、15.3°C、10.5°C、14.7°C、15.0°C，其中容許誤差值為 2°C。反應節點將這些接收到的資料處理後，得平均值為 14.12°C，再比較每個感測資料與平均值的誤差得到 0.98°C、1.18°C、3.62°C、0.58°C、0.88°C。這邊可以發現 SN₃ 所感測到的溫度與平均值的誤差大於容許誤差值，因此設定 SN₃ 為錯誤節點。但是當感測資料與正常的資料誤差值極大時，此方法反而會造成反應節點作出錯誤的判斷與處理。如圖 4，若感測節點蒐集到的資料為 15.1°C、15.3°C、0°C、14.7°C、15.0°C，容許誤差值一樣為 2°C。反應節點透過這些資料處理，得到平均值為 12.02°C，這時再比較每個感測資料與平均值的誤差得到 3.08°C、3.28°C、12.02°C、2.68°C、2.98°C，由於每個感測資料皆大於容許誤差值，因此反應節點會誤判所有節點皆為錯誤節點，但實際上，錯誤節點只有感測到 0°C 的感測節點 SN₃。為了解決此問題，本論文提出了四分位法容錯機制，讓反應節點根據蒐集到的感測資料作更正確

的判斷與處理。

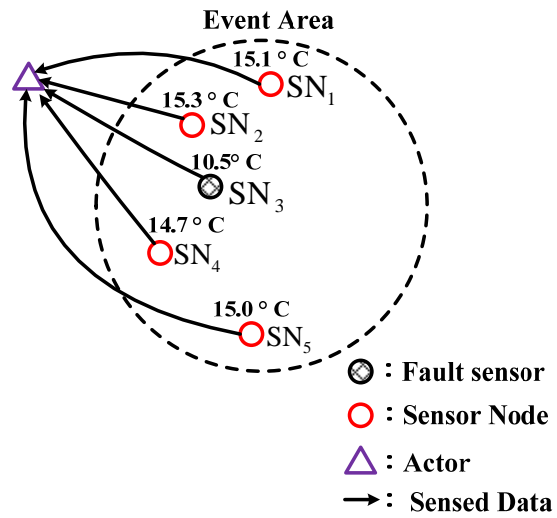


圖 3 平均法(1)

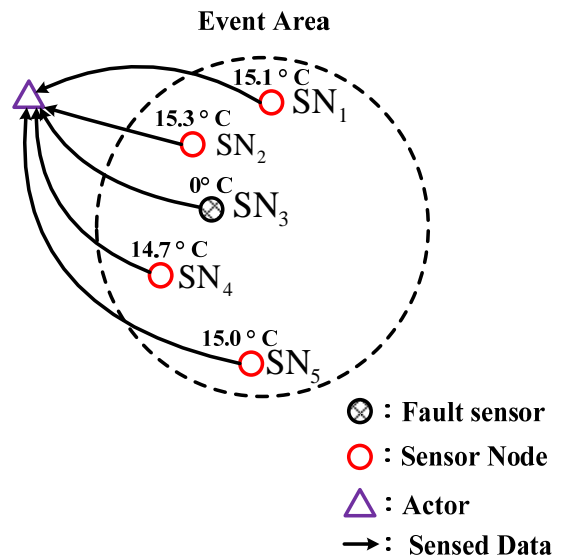


圖 4 平均法(2)

3. 四分位法容錯機制 (Fault Tolerance by Using Quartiles Method, FTQM)

在無線傳感器網路中，反應節點往往需要藉由蒐集到的感測資料來決定相對應的判斷或處理。然而，反應節點並無法有效的篩選出正確資料的範圍，導致反應節點可能會因為錯

誤資料的影響，作出不正確的判斷或處理方式。

因此，本論文提出四分位法容錯機制來解決此問題。此機制主要利用四分位法針對反應節點蒐集到的感測資料作分析預測，透過四分位差中資料的離散程度來預測正確資料的分佈位置，以篩選出正確節點所傳送的資料。

3.1 四分位數與盒形圖

四分位數是將所有資料由小到大排序後，將全部資料中第 $\frac{1}{4}$ 位置的數，稱為第1四分位數(Q_1)；在全部資料中第 $\frac{1}{2}$ 位置的數，稱為中位數(M)；在全部資料中排第 $\frac{3}{4}$ 位置的數，稱為第3四分位數(Q_3)，如圖5所示。舉例來說，若反應節點蒐集到 n 個已經排序好的資料，則第1四分位數和第3四分位數的算法如下，假

設 Y_x 表第 x 筆資料的數，並令 $m = \frac{n+1}{4}$ ：

當 m 為整數時，

$$Q_1 = Y_m \quad (1)$$

$$Q_3 = Y_{3m} \quad (2)$$

當 m 非為整數時，

$$Q_1 = (\lceil m \rceil - m)Y_{\lfloor m \rfloor} + (m - \lfloor m \rfloor)Y_{\lceil m \rceil} \quad (3)$$

$$Q_3 = (\lceil 3m \rceil - 3m)Y_{\lfloor 3m \rfloor} + (3m - \lfloor 3m \rfloor)Y_{\lceil 3m \rceil} \quad (4)$$

中位數的算法如下：

當 $2m$ 為整數時，

$$M = Y_{2m} \quad (5)$$

當 $2m$ 非為整數時，

$$M = \frac{1}{2} \times (Y_{\lfloor 2m \rfloor} + Y_{\lceil 2m \rceil}) \quad (6)$$

例如：有一個16筆的資料，可以知道 $n=16$ ，

則 $\frac{n+1}{4} = 4.25$ ，可知 $\lfloor m \rfloor = 4$ ， $\lceil m \rceil = 5$ ，即可得

到 $Q_1 = 0.75 \times Y_4 + 0.25 \times Y_5$ 。

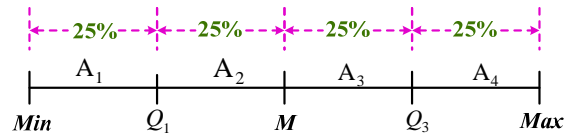


圖 5 四分位數示意圖

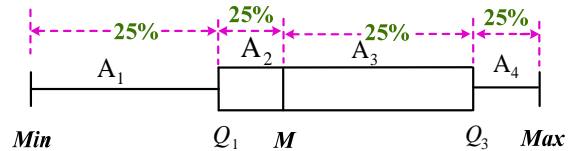


圖 6 盒形圖示意圖

假設最小值(Min)是由小到大排序後資料中的第一個數值，最大值(Max)是全部資料中的最後一個數值。而 Min 、 Q_1 、 M 、 Q_3 、 Max 又可以稱為五數總結(Five - Number Summary)，其中 Min 與 Q_1 之間的資料標示為(Area 1, A_1)、 Q_1 與 M 之間的資料標示為(Area 2, A_2)、 M 與 Q_3 之間的資料標示為(Area 3, A_3)， Q_3 與 Max 之間的資料標示為(Area 4, A_4)。

盒形圖是五數總結中的五個參數的關係圖。如圖6所示，由於每個區域的資料都佔了全部資料的 $\frac{1}{4}$ ，若該區域的差距越大，代表資料分佈的越離散、差異越大；相反的，若該區域的差距越小，代表其中的數據越集中、差異越小。

3.2 四分位法容錯機制

由於正常節點傳送的資料經過排序後，會集中在某一個連續範圍內，而錯誤感測資料的離散度則較高，因此可利用此特性配合四分位法來預估正確資料的分佈位置。

首先利用 Q_1 、 M 、 Q_3 找出資料的分佈情形；接著依據 Q_1 和 M 、 M 和 Q_3 數這兩個資料差距與反應節點的容許誤差值 r 分為四個方案

來探討。由於差距越小的區域包含的正確資料就越多，因此這邊將這兩個差距當成權重，並重新調整使兩個區域的權重相等，其步驟如下：

1. 反應節點蒐集感測節點傳送過來的感測資料，並由小到大進行排序。
2. 找出五數總結中的 Min 、 Q_1 、 M 、 Q_3 、 Max 。
3. 計算 Max 與 Min 的差距是否超過容許的誤差範圍，若無，則表示沒有錯誤資料產生，所有資料皆為正確的感測資料，此時 $C_{min} = Min$ 、 $C_{max} = Max$ 。
4. 若最大值與最小值的差距超過容許誤差值，則根據 $(M - Q_1)$ 和 $(Q_3 - M)$ 和容許誤差值比較，分為四個方案來探討。

圖 7 為四分位預估法流程圖。

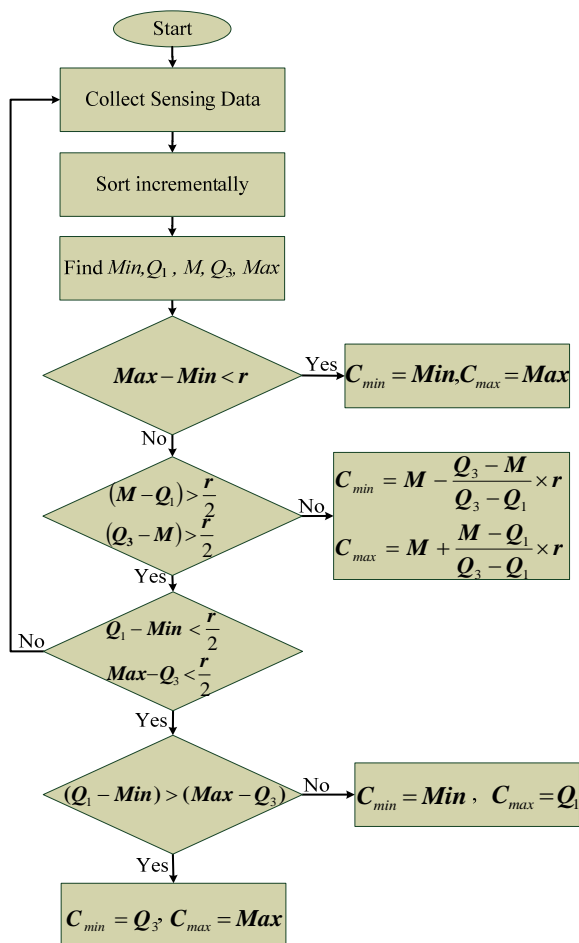


圖 7 四分位預估法流程圖

下面分別為四個方案作介紹：

方案 1: $M - Q_1 \leq \frac{r}{2}$ 且 $Q_3 - M \leq \frac{r}{2}$

此情況表示在 A_2 區域和 A_3 區域的資料都很集中，可以判斷在這兩區域內的資料都是正確資料。然而正確資料可能不只這 50%，為了找到更多的正確資料，這邊利用 $(M - Q_1)$ 和 $(Q_3 - M)$ 兩個差距權重和反應節點容許的誤差範圍，重新找尋正確資料的範圍。如圖 8 所示，假設 $(M - Q_1)$ 和 $(Q_3 - M)$ 的比值為 $a:b$ ，則修正範圍時，往左分配 b 個比例，往右分配 a 個比例，因此可以算出修正後的 C_{max} 與 C_{min} ：

$$C_{min} = M - \frac{Q_3 - M}{Q_3 - Q_1} \times r \quad (7)$$

$$C_{max} = M + \frac{M - Q_1}{Q_3 - Q_1} \times r \quad (8)$$

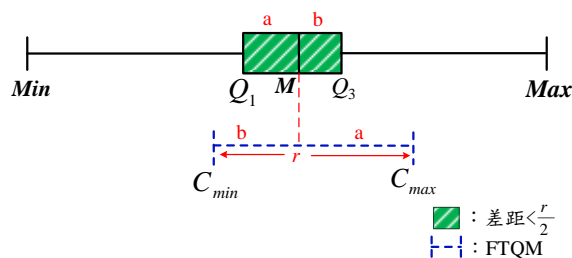


圖 8 四分位法(方案一)

方案 2: $M - Q_1 > \frac{r}{2}$ 且 $Q_3 - M \leq \frac{r}{2}$

此情況表示在 A_3 區域內的資料較集中，所以該區域內的資料較正確，而 A_2 內的資料較離散，所以正確數值偏高。因此透過誤差權重的方法，往數值大的方向取較大的範圍，而往數值小的方向取較小的範圍，如圖 9 所示，其中

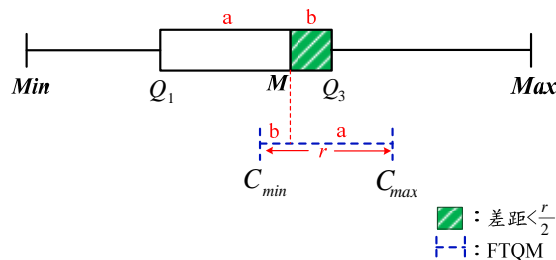


圖 9 四分位法(方案二)

C_{max} 和 C_{min} 的求法採用公式(1)、公式(2)。

方案 3: $M - Q_1 \leq \frac{r}{2}$ 且 $Q_3 - M > \frac{r}{2}$

此情況表示在 A_2 區域內的資料較集中，所以該區域內的資料較正確，而 A_3 內的資料較離散，所以正確數值偏低。這時可透過誤差權重的方法往數值小的方向取較大的範圍，而往數值大的方向取較小的範圍，如圖 10 所示。其中 C_{max} 和 C_{min} 的求法採用公式(1)、公式(2)。

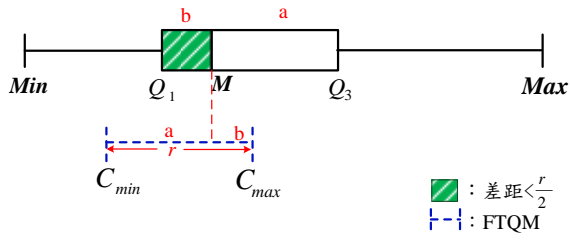


圖 10 四分位法(方案三)

方案 4: $M - Q_1 > \frac{r}{2}$ 且 $Q_3 - M > \frac{r}{2}$

此情況表示 A_2 和 A_3 兩區域內的資料都很離散。因此無法從此兩區域內的資料判斷正確或錯誤的節點。這時觀察 A_1 區域或 A_4 區域的差距是否小於 $\frac{r}{2}$ ，若成立，則將差距較小的那區域當作正確節點的資料。假設 A_2 到 A_4 區域的差距都大於 $\frac{r}{2}$ ， A_1 區域的差距小於 $\frac{r}{2}$ ，這時可以判斷出只有 A_1 區域的資料較集中，且 $C_{min} = Min$ ， $C_{max} = Q_1$ 。若所有區域的差值都大於 $\frac{r}{2}$ ，則表示資料分部的太離散，無法判斷正確資料的分佈區域，因此放棄該筆資料並重新部屬所有的感測節點。

由此可知，四分位預估法可預測正確資料的分佈位置，以讓反應節點作正確的處理，其演算法如下所示：

Input: the sensing data set V of N sensor nodes, permissible error rate r .

Output: C_{min} 、 C_{max} .

Procedure:

1. sorting V ;
2. Find Min, Q_1, M, Q_3, Max of all SNs;
3. if($Max - Min < r$)
4. { $C_{min} = Min, C_{max} = Max$;}
5. else
6. {if($M - Q_1 > \frac{r}{2}$ and $Q_3 - M > \frac{r}{2}$)
7. if($(Q_1 - Min) < \frac{r}{2}$ or $(Max - Q_3) < \frac{r}{2}$)
8. if($(Q_1 - Min) < (Max - Q_3)$)
9. { $C_{min} = Min, C_{max} = Q_1$;}
10. else
11. { $C_{min} = Q_3, C_{max} = Max$;}
12. End if;
13. else
14. discard the data and re-deployment all nodes;
15. End if;
16. else
17. { $C_{min} = M - \frac{Q_3 - M}{Q_3 - Q_1} \times r$;
 $C_{max} = M + \frac{M - Q_1}{Q_3 - Q_1} \times r$;}
18. End if;}
19. End if;

4. 模擬環境建立與評量因子

本研究以 .Net 為平台並使用 C++ 撰寫模擬程式。其模擬部署範圍大小為 200×200 平方公尺，感測節點以均勻隨機方式部署，反應節點以人工方式部署，使每個感測節點皆可將感測資料傳送到至少一個反應節點。錯誤節點的定義為感測數值大於容許誤差值。這邊以節點正確率、準確值誤差和錯誤節點率為本論文之評估因子，並與平均法[8]作比較。其中節點正確率(Accurate Rate, AR)是 FTQM 方法篩選出來的節點正確率與實際正確節點率的比較，目

地是驗證 FTQM 篩選出資料的正確性，如公式 (9)：

$$AR = \frac{N_{pc}}{N_p} \times 100\% \quad (9)$$

N_p ：利用 FTQM 求出的正確節點數。

N_{pc} ： N_p 中實際正確的節點數。

準確值誤差 (Accurate Value Deviation, AVD) 是 FTQM 篩選出來的資料平均值與實際準確數值作比較，目地為驗證 FTQM 方法判斷事件的準確性。如公式(10)：

$$AVD = \left| V_c - \frac{\sum_{i=1}^{N_p} V_p^i}{N_p} \right| \quad (10)$$

V_c ：事件實際數值。

V_p^i ：FTQM 下正確節點中第 i 個節點的感測數值。

N_p ：利用 FTQM 求出的正確節點個數。

錯誤節點率 (Error Node Rate, ENR) 為 FTQM 篩選出來的錯誤節點個數和實際錯誤節點個數的比率，目地為觀察篩選錯誤節點的容錯性。如公式(11)：

$$ENR = \frac{N_p^e}{N_c^e} \times 100\% \quad (11)$$

N_p^e ：FTQM 下錯誤節點個數 ($N - N_p$)。

N_c^e ：實際錯誤節點個數 ($N - N_c$)。

本研究將以上述三個評估因子和平均法作比較，來驗證 FTQM 的方法能夠得到較正確的資料。

5. 結論與未來工作

本論文針對反應節點蒐集的感測資料提出四分位法容錯機制，利用四分位數中的盒形圖將整筆資料分為四個區域，再利用區域間間

距的大小來判斷正確資料的範圍。若正確資料在中間兩區域內，就在進一步的找出所有可能為正確的資料。若中間資料的離散度過大，則針對另外兩個區域作判斷。若四個區域資料的離散度都過大，則表示該環境中故障的感測節點過多，需要重新部署新的感測節點。透過本機制可以淘汰因容許度異常而造成感測到錯誤資料的感測節點，並刪除錯誤的感測資料，使反應節點能正確的判斷出事件發生的情況作進一步的分析及解決。

後續將討論 WSN 環境的佈置方式與模擬結果的分析，並藉由調整感測節點的錯誤率，來探討在不影響反應節點判斷事件的前提下，容許故障的感測節點的比率為多少。由於本論文目前主要針對單一區域內單一事件區域所提出的容錯機制，未來將針對單一區域內有多個事件發生的容錯機制加以探討，並利用資料探勘制定新的容錯機制，以作更完善的錯誤分析和處理。

6. 致謝

本研究獲「臺北科技大學 97 年教學卓越計畫 — 伯樂計劃」贊助，特此致謝。

參考文獻

- [1] I. F. Akyildiz, and I. H. Kasimoglu, "Wireless Sensor and Actor Networks: Research Challenges," *Ad Hoc Networks Journal (Elsevier)*, 2004, pp. 351-367.
- [2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: a Survey," *Computer Networks*, 2002, pp. 393-422.
- [3] M. - H. Lee, and Y. - H. Choi, "Fault Detection of Wireless Sensor Networks,"

- IEEE Journal of Computer Communications*, 2008, pp. 3469-3475.
- [4] M. Yu, H. Mokhtar, and M. Merabti, "Fault Management in Wireless Sensor Networks," *IEEE Wireless Communications*, Dec. 2007, pp. 13-19.
- [5] K. Morita, K. Watanabe, N. Hayashibara, and M. Takizawa, "An Efficient Data Transmission Protocol in a Wireless Sensor-Actuator Network," *On Conf. of Complex, Intelligent and Software Intensive Systems(CISIS)*, 2007, pp. 11-18.
- [6] T. Melodia, D. Pompili, V. C. Gungor, and I. F. Akyildiz, "A Distributed Coordination Framework for Wireless Sensor and Actor Networks," *In Proc. of ACM Intl. Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, Urbana - Champaign, IL, May 2005.
- [7] M. Yu, H. Mokhtar, and M. Merabti, "Fault Management in Wireless Sensor Networks," *IEEE Wireless Communications*, Dec. 2007, pp. 13-19.
- [8] K. Ozaki, K. Watanabe, S. Itaya, N. Hayashibara, T. Enokido, and M. Takizawa, "A Fault-Tolerant Model of Wireless Sensor-Actor Network," *In Proc. of Object and Component-Oriented Real-Time Distributed Computing, 2006 (ISORC 2006)*, April 2006.
- [9] E. Ngai, YF Zhou, MR Lyu, and JC Liu, "Reliable Reporting of Delay-Sensitive Events in Wireless Sensor-Actuator Networks," *In Proc. of the 3rd IEEE on Mobile Ad-Hoc and Sensor Systems (MASS)*, 2006, pp. 101-108.