1 Network Model

The network model assumes that the network consists of a number of clusters. In each cluster, one node is selected as the CH that takes charge of central functions, and the others (MNs) equally participate in team-functions. It is well known that WSNs can be categorised as hierarchical and flat according to routing protocols [**?**]. A hierarchical WSN directly coincides with this model while, when the detector is operated by a flat WSN, we have to employ the additional techniques summarised in [**?**] to intentionally cluster the nodes. Moreover, we assume that the topology of the network is static, with all the nodes time-synchronised.

For a flat WSN, we have to enforce the linear relationships on every cluster and achieve a trade-off between cost and effectiveness. In particular, it has to ensure that

- all the nodes within a cluster are spatially close to each other,
- information can be transmitted between any pair of nodes without multi-hopping and
- the scale of each cluster is moderate.

We assume that the BS knows the geographical coordinates of the nodes and that the topology of the network remains unchanged. Therefore, it is able to cluster the nodes in advance using a fixed-width clustering algorithm [**?**]. Suppose that ω denotes the fixed-width and ρ the transmission radius of a node. Then, as the farthest distance between two nodes within a cluster will be 2ω, any two nodes can directly communicate when $\rho \geq 2\omega$. The BS also randomly selects a CH for each cluster which may be reselected periodically. Moreover, as all the nodes are assumed to be time-synchronised, each MN can collect a data segment of equal size during a given period of time.

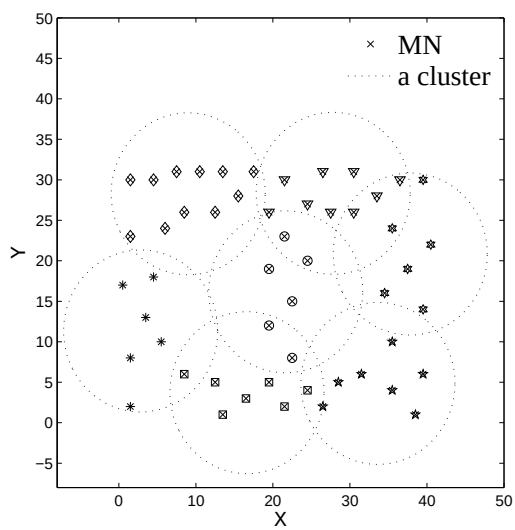

Fig. 1. Topology and clusters of the network deployed at the IBRL

As an example, we cluster the flat WSN deployed at the IBRL which is used for the later numerical experiments. This network is comprised of 54 Mica2 sensor nodes and Figure 1 illustrates its topology together with the resulting clusters from which damaged nodes are removed (Nos. 5 and 15) and $\omega = 10$.

2 Algorithms and Full Scheme

In this section, the main algorithms mentioned in the main file are presented. In particular, each MN produces its compressed difference sequence locally with Algorithm 1 *difference sequence compressor* (DSC) and the CH retrieves the raw rank sequence through Algorithm 2 *difference sequence decompressor* (DSD). The recovery of $\bar{Q}$ is implemented by Algorithm 3 *approximate sample covariance matrix* (ASCM) and then the detector is described by Algorithm 4 *prediction variance detector* (PVD). These algorithms comprise the full segment-based anomaly detection scheme (SADS) given by Algorithm 5.

The full scheme works as follows. After a MN locally collects a data segment, it obtains the rank sequence and sample standard deviation, and performs Algorithm 1 to produce the compressed difference sequence. When the CH collects all the compressed difference sequences from the MNs, they are decompressed to their raw rank sequences by Algorithm 2 and, if tied ranks are found in a raw rank sequence, it will be transformed to a reassigned rank sequence. Then, the CH can approximate the sample covariance matrix using Algorithm 3, according to which the detection can be implemented with Algorithm 4.

Algorithm 1 DSC

$\hat{\mathrm{S}}^* = \mathrm{DSC}\,(\mathrm{S})$
$\hat{\mathrm{S}}\,(1) = \mathrm{S}\,(1)$
for $\mathrm{i} = 2$ to n **do**
  $\Delta = \mathrm{S}\,(\mathrm{i}) - \mathrm{S}\,(\mathrm{i}-1)$
  if $\Delta \leq 0$ **then**
    $\Delta = 127 - \Delta$
  end if
  $\hat{\mathrm{S}}\,(\mathrm{i}) = \Delta$
end for
if if any $\Delta \in \hat{\mathrm{S}}$ repeat more than 2 times **then**
  $\hat{\mathrm{S}}^* : \left[\underbrace{\Delta \cdots \Delta}_{\text{repeat } \tau \text{ times}} \right] \rightarrow [\Delta\ 0\ \mathrm{d}]$
end if
return $\hat{\mathrm{S}}^*$

Algorithm 2 DSD

$\mathrm{S} = \mathrm{DSD}\left(\hat{\mathrm{S}}^*\right)$
if any $[\Delta\ 0\ \tau]$ found in $\hat{\mathrm{S}}^*$ **then**
  $\hat{\mathrm{S}} : [\Delta\ 0\ \tau] \rightarrow \left[\underbrace{\Delta \cdots \Delta}_{\text{repeat } \tau \text{ times}} \right]$
end if
$\mathrm{S}\,(1) = \hat{\mathrm{S}}\,(1)$
for $\mathrm{i} = 2$ to n **do**
  if $\hat{\mathrm{S}}\,(\mathrm{i}) \geq 127$ **then**
    $\Delta = 127 - \hat{\mathrm{S}}\,(\mathrm{i})$
  end if
  $\mathrm{S}\,(\mathrm{i}) = \mathrm{S}\,(\mathrm{i}-1) + \Delta$
end for
return S

TABLE 2
Anomaly injectors

Type	Function	k	γ	$\lvert s \rvert$
I	$X_k(s) = \gamma \cdot mean(X_k(s))$	$[1, m]$	$[0.98, 1.02]$	$0.75n$
II	$X_k(s) = X_k(s) + \gamma \cdot std(X_k(s))$	$[1, m]$	$[-10, 10]$	$0.1n$
III	$X_k(s) = X_k(s) + noise$	$[1, m]$	$[1.2, 1.4]$	$0.75n$
IV	$X_k(s) = X_k(s) + noise$	$[1, m]$	$[2.25, 3.25]$	$0.33n$

$noise \sim N(0, \gamma \cdot Var(X_k(s)))$

Algorithm 3 ASCM

$\bar{\mathrm{Q}} = \mathrm{ASCM}\left([\mathrm{S}_1\mathrm{S}_2 \cdots \mathrm{S}_\mathrm{m}], [\bar{\sigma}_{\mathrm{X}_1}\bar{\sigma}_{\mathrm{X}_2} \cdots \bar{\sigma}_{\mathrm{X}_\mathrm{m}}]\right)$
if tied ranks are found in S **then**
  $\mathrm{S} \to \bar{\mathrm{S}}$
end if
for $\mathrm{i} = 1$ to m **do**
  for $\mathrm{j} = 1$ to m **do**
    $\bar{\mathrm{r}}_{\mathrm{X_i X_j}} = \mathrm{r}_{\bar{\mathrm{S}}_\mathrm{i}\bar{\mathrm{S}}_\mathrm{j}}$ or $\bar{\mathrm{r}}_{\mathrm{X_i X_j}} = 1 - \frac{6\left(\mathrm{S_i} - \mathrm{S_j}\right)\left(\mathrm{S_i} - \mathrm{S_j}\right)^\mathrm{T}}{\mathrm{l}(\mathrm{l}^2 - 1)}$
    $\bar{\mathrm{c}}_{\mathrm{X_i X_j}} = \bar{\mathrm{r}}_{\mathrm{X_i X_j}}\bar{\sigma}_{\mathrm{X_i}}\bar{\sigma}_{\mathrm{X_j}}$
  end for
end for
return $\bar{\mathrm{Q}}$

Algorithm 4 PVD

$\left[\mathrm{label}, \bar{\mu}_\mathrm{Y}^{\mathrm{t}+1}\right] = \mathrm{PVD}\left(\bar{\mathrm{Q}}, \bar{\mu}_\mathrm{Y}^\mathrm{t}\right)$
for $\mathrm{i} = 1$ to m **do**
  $\mathrm{A} = \bar{\mathrm{Q}}$ deleting i^{th} row and i^{th} column
  $\mathrm{B} = \mathrm{i}^{\mathrm{th}}$ column of $\bar{\mathrm{Q}}$ deleting $\bar{\mathrm{c}}_{\mathrm{X_i X_i}}$
  $\mathrm{C} = \bar{\mathrm{c}}_{\mathrm{X_i X_i}}$
  if A is (almost) singular **then**
    $\mathrm{A}^{-1} = \mathrm{A}^\dagger$
  end if
  $\beta_0 = \mathrm{B}^\mathrm{T}\mathrm{A}^{-1}\mathrm{B}$
  $\beta_1 = -\left(\mathbf{1}^\mathrm{T} \cdot \mathrm{A}^{-1} \cdot \mathbf{1}\right)^{-1}$
  $\beta_2 = \mathbf{1} \cdot \mathrm{A}^{-1}\mathrm{B}$
  $\mathrm{y}_\mathrm{i}^{\mathrm{t}+1} = \mathrm{C} - \beta_0 - \beta_1(\beta_2 - 1)^2$
end for
$\bar{\mu}_\mathrm{Y}^{\mathrm{t}+1} = \mathrm{mean}\left(\mathrm{Y}^{\mathrm{t}+1}\right)$
$\hat{\mu}_\mathrm{Y}^{\mathrm{t}+1} = \theta\bar{\mu}_\mathrm{Y}^\mathrm{t} + (1 - \theta)\,\bar{\mu}_\mathrm{Y}^{\mathrm{t}+1}$
for $\mathrm{i} = 1$ to m **do**
  if $(\mathrm{m} - 1)\,\frac{\mathrm{y}_\mathrm{i}^{\mathrm{t}+1}}{\hat{\mu}_\mathrm{Y}^{\mathrm{t}+1}} \in \left[\mathrm{F}^{-1}_{\chi^2_{\mathrm{m}-1}}(\alpha), \mathrm{F}^{-1}_{\chi^2_{\mathrm{m}-1}}(1 - \alpha)\right]$ **then**
    $\mathrm{label}\,(\mathrm{i}) = 0$ #normal
  else
    $\mathrm{label}\,(\mathrm{i}) = 1$ #abnormal
  end if
end for
$\bar{\mathrm{Y}}_{\mathrm{t}+1} = \left\{\mathrm{y}_\mathrm{i}^{\mathrm{t}+1} \middle| \mathrm{i} = 1, 2, \cdots \mathrm{n} \text{ and } \mathrm{label(i)} = 0\right\}$
$\bar{\mu}_\mathrm{Y}^{\mathrm{t}+1} = \mathrm{mean}\left(\mathrm{Y}^{\mathrm{t}+1}\right)$

Algorithm 5 SADS

Time t
Initialise $\bar{\mu}_\mathrm{Y}^\mathrm{t}$
Go to time $\mathrm{t} = \mathrm{t} + 1$
i^{th} **MN**
collect $\mathrm{X_i}$
obtain $\bar{\sigma}_{\mathrm{X_i}}$
obtain $\mathrm{S_i}$ by sorting $\mathrm{X_i}$ in ascending order
$\hat{\mathrm{S}}_\mathrm{i}^* = \mathrm{DSC}\,(\mathrm{S_i})$
transmit $\hat{\mathrm{S}}_\mathrm{i}^*$ and $\bar{\sigma}_{\mathrm{X_i}}$ to the CH
CH
collect $\hat{\mathrm{S}}_\mathrm{i}^*$ and $\bar{\sigma}_{\mathrm{X_i}}$, $\mathrm{i} = 1, 2, \cdots, \mathrm{m}$
$\mathrm{S_i} = \mathrm{DSD}\left(\hat{\mathrm{S}}_\mathrm{i}^*\right)$
$\bar{\mathrm{Q}} = \mathrm{ASCM}\left([\mathrm{S}_1\mathrm{S}_2 \cdots \mathrm{S}_\mathrm{n}], [\bar{\sigma}_{\mathrm{X}_1}\bar{\sigma}_{\mathrm{X}_2} \cdots \bar{\sigma}_{\mathrm{X}_\mathrm{n}}]\right)$
$\left[\mathrm{L}^\mathrm{t}, \bar{\mu}_\mathrm{Y}^{\mathrm{t}+1}\right] = \mathrm{PVD}\left(\bar{\mathrm{Q}}, \bar{\mu}_\mathrm{Y}^\mathrm{t}\right)$
repeat
  3-14
until the end

Algorithm 6 PMD

Time t
i^{th} **MN**
collect an observation $\mathrm{x}_\mathrm{i}^\mathrm{t}$
transmit $\mathrm{x}_\mathrm{i}^\mathrm{t}$ to the CH
CH
collect $\mathrm{Z}^\mathrm{t} = \{\mathrm{x}_1^\mathrm{t}, \mathrm{x}_2^\mathrm{t}, \cdots, \mathrm{x}_\mathrm{m}^\mathrm{t}\}$
for $\mathrm{i} = 1$ to m **do**
  $\mathrm{d_i} = \mathrm{x}_\mathrm{i}^\mathrm{t} - \mathrm{median}\left(\mathrm{Z}^\mathrm{t} - \{\mathrm{x}_\mathrm{i}^\mathrm{t}\}\right)$
end for
$\{\bar{\mathrm{d}}_1, \bar{\mathrm{d}}_2, \cdots, \bar{\mathrm{d}}_\mathrm{m}\} = \mathrm{abs}\left(\mathrm{standarise}\left(\{\mathrm{d}_1, \mathrm{d}_2, \cdots, \mathrm{d}_\mathrm{m}\}\right)\right)$
for $\mathrm{i} = 1$ to m **do**
  if $\bar{\mathrm{d}}_\mathrm{i} > \eta$ **then**
    $\mathrm{L}_\mathrm{i}^\mathrm{t} = 1$
  else
    $\mathrm{L}_\mathrm{i}^\mathrm{t} = 0$
  end if
end for
Go to time $\mathrm{t} = \mathrm{t} + 1$
repeat
  2-18
until the end

TABLE 1
Notation list

Notation	Explanation
subscript, $1, 2, \cdots, m$	index of node
superscript, $t \in \mathbb{Z}$	period of time
m	total number of sensor nodes
$\bar{}$	an estimation
E/μ	expectation
Var/σ^2	variance
X	a random variable (sensed measurement)
Z	a set of random variables
$\hat{Z}$	$Z - \{X\}$
w	weight
ε	estimation error
W	a vector of weights
Q	sample covariance matrix
Λ	Lagrange function
θ	Lagrange multiplier
A, B, C	intermediate quantities
$\beta_0, \beta_1, \beta_2$	intermediate quantities
N	Gaussian distribution
χ^2	chi-squared distribution
F	cumulative density function
$\alpha, \alpha \in (0, 0.5)$	confidence level
Y	$Var\,(X)$
y	predication variance/observation of Y