A Quantitative Measure of Reliability for Wireless Sensor Networks

Waltenegus Dargie, *Senior Member, IEEE*

***Abstract*—The reliability of a wireless sensor network is often associated with its capacity to maintain an end-to-end connection in the presence of unreliable links or node failure []. It is also associated with the level of partitioning the network suffers when some nodes fail or exhaust their batteries. Despite being a critical feature, often it is not clear how it can be expressed in quantifiable terms. We assert that Katz's *Status Index* can be used as an adequate measure of reliability. Originally proposed for sociometric analysis, it assigns a value of significance to each node in a network. Some of the underlying assumptions made and the intermediate steps taken to construct the *Status Index* are not applicable for our case, but we shall demonstrate that their effect on the final result is marginal.**

***Index Terms*—Reliability, status index, multi-hop communication, wireless sensor networks**

I. Introduction

Reliability is a critical aspect of wireless networks, even though expressing it in measurable or quantitative terms is not straightforward [1], [2]. In general purpose networks, reliability is often associated with fault-tolerance and node failure and defined in terms of these properties [3]. But when considering networks consisting of resource-limited sensor nodes, other inherent aspects play equally important roles in affecting reliability. For example, in wireless sensor networks, communication between any two neighbour nodes may not be successful for various reasons: The wireless link may be unreliable, the receiver may be sleeping to save energy, or the channel may be busy [4]. On the other hand, a transmitter may not differ communication arbitrarily lest its buffer becomes full and it is forced to overwrite packets. The effect of all these aspects on the operation of the network can be considerable, having both short and long-term impacts, depending on the position and role of the nodes and the criticality of the sensing assignment.

In graph theory, the degree of nodes is often equated with their relevance in the network [5], [6] and rightly so. Consider, for example, Figure 1 where nodes 2 and 6 are of degree one and four, respectively. If we consider their information dissemination capacity in the presence of uncertain links and node failure, node 2 and 6 are markedly different: Node 6 can reach its peers using 4 single hop, 7 two-hop, and 9 three-hop alternative routes whereas node 2 can reach its peers using 1 single-hop, 4 two-hop, 6 three-hop and 6 four-hop routes. If taken individually, however, the node degree may not adequately express the contribution of nodes to the overall reliability of the network. For example, despite their considerable difference in node degree, the impacts of the independent failure of node 2 and node 6 on the network appear to be comparable. The same can be said of node 3, whose independent failure affects only itself and node 2. It is, therefore, apparent, that an expression of reliability should take different aspects into consideration. In the context of social networks, additional parameters such as closeness, betweenness, ties, and link importance (weight) [7], [5], [8] are taken into account to quantify the relevance (or centrality) of nodes, however, with the assumption that these parameters are readily available. In most wireless sensor networks, such assumptions cannot be made.

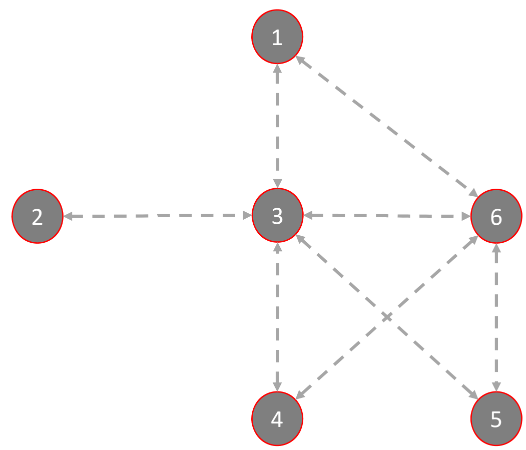

Fig. 1: A wireless network with uncertain or unreliable links.

We assert that Katz's *Status Index* [9] can be applied to measure the contribution of nodes to the overall reliability of a wireless sensor network. Originally proposed for sociometric analysis (to rank the status of people in a social circle), it assigns a value of significance to each node in the network. Some of the underlying assumptions made and the intermediate steps taken to construct the *Status Index* are invalid for our case, but we shall demonstrate that their effect on the final result is marginal. The remaining part of this paper is organised as follows: In Section II, we introduce Katz's *Status Index* and highlight the underlying assumptions. In Section III, we provide the mathematical justification to use the *Status Index* as a measure of network reliability. In Section IV, we

Manuscript first submitted on 07 May 2019. Manuscript accepted on 27 July 2019.

W. Dargie is with the Technische Universität Dresden, Faculty of Computer Science, 01062, Dresden, Germany.
E-mail: waltenegus.dargie@tu-dresden.de

This work has been partially funded by the German Research Foundation (DFG) in the context of the RoReyBan project (DA 1121/7-1).

provide quantitative evaluation of different topologies. Finally, in Section V, we provide concluding remarks.

II. Katz's Status Index

According to Katz, the normalised number of direct as well as all possible indirect links people establish with their peers in a social network expresses their relative significance in the network and, therefore, should be taken as a measure of their social standing or status. The following implicit assumptions are made to compute the *Status Index*:

- Direct links are predominantly asymmetric.
- The channels are statistically independent.
- All multi-hop routes of the same order or length are equally valid.
- There can be infinite number of multi-hop routes.

When considering typical wireless sensor networks, some of these assumptions do not necessarily hold. For instance, most wireless links are symmetrical for short-distance communications. This does not mean, of course, asymmetric links do not exist. We shall, however, put aside Katz's assumptions for the time being and follow his argument. Accordingly, one can use a square matrix to express the number of direct (single-hop) links persons (nodes) can establish with their neighbours[1]. For the network depicted in Figure 1, the links between the nodes can be represented as follows:

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \quad (1)$$

$c_{ij} = 1$ if there is a direct link between node i and j, otherwise $c_{ij} = 0$. The values of c_{ii} in $\mathbf{C}$ as well as in all its higher powers are set to zero or are simply ignored. Squaring $\mathbf{C}$ results in the number of two-hop links the nodes establish with their peers:

$$\mathbf{C}^2 = \begin{bmatrix} 0 & 1 & 1 & 2 & 2 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 3 \\ 2 & 1 & 1 & 0 & 2 & 1 \\ 2 & 1 & 1 & 2 & 0 & 1 \\ 1 & 1 & 3 & 1 & 1 & 0 \end{bmatrix} \quad (2)$$

An interesting aspect of Equation 2 is that it is computationally tractable even for a large network, since $c^2_{ij} = \sum_k c_{ik} c_{kj}$ and each multiplication operation yields either 0 or 1. Likewise, $\mathbf{C}^3$ produces the number of three-hop links each node establishes with its peers. But here we encounter a problem. Katz regards as valid all multi-hop links resulting from the higher powers of $\mathbf{C}$ whereas this is not the case for us. For example, according to Equation 2, there are 5 three-hop routes connecting node 2 and 3. This can be achieved, however, if we consider as valid routes that can go back and forth through node 3 (Figure 2), which does not make sense.

[1]The rows and the columns of the matrix refer, respectively, to the source and destination nodes. The order does not matter when the network is symmetric.

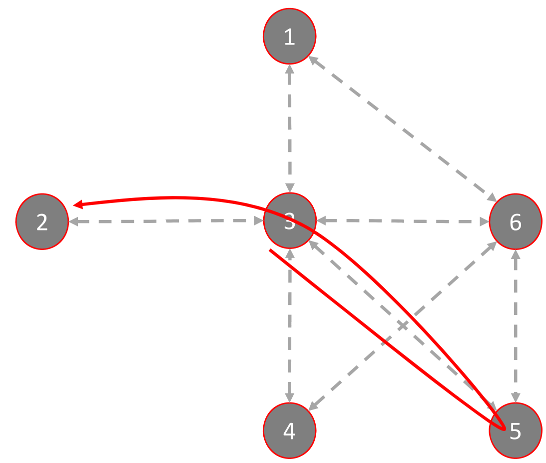

Fig. 2: An invalid three-hop route.

$$\mathbf{C}^3 = \begin{bmatrix} 0 & 1 & 8 & 2 & 2 & 7 \\ 1 & 0 & 5 & 1 & 1 & 3 \\ 8 & 5 & 0 & 8 & 8 & 8 \\ 2 & 1 & 8 & 0 & 2 & 7 \\ 2 & 1 & 8 & 2 & 0 & 7 \\ 7 & 3 & 8 & 7 & 7 & 0 \end{bmatrix} \quad (3)$$

With a slight modification, we can still make use of Equation 3 to determine the exact number of three-hop links a node establishes with its peers in Fig. 1:

$$n^3_{ij} = \begin{cases} c^3_{ij} - (d_i + d_j - 1), & \text{if } (d_i + d_j - 2) > 3 \\ c^3_{ij}, & \text{otherwise} \end{cases} \quad (4)$$

where n^3_{ij} is the number of three-hop routes which can be established between node i and j and d_i and d_j are the node degree of nodes i and j, respectively. The expression becomes more complex for determining the number of valid hops in the higher powers of $\mathbf{C}$. The sum of the j-th column of $\mathbf{C}$ yields the number of single hop links node j establishes with its peers. Similarly, the sum of the j-th column of $\mathbf{C}^2$ yields the number of two-hop links node j establishes with its peers, and so on.

III. Measure of Reliability

Reliability becomes an issue when the success of establishing a wireless link becomes uncertain. Suppose the probability of establishing a link between any two neighbour nodes is p. For a network of an appreciable size, the total number of links (direct as well as indirect) a node can successfully establish with its peers can be expressed as:

$$\mathbf{T} \approx \sum_{k=1}^{\infty} (p\mathbf{C})^k = (\mathbf{I} - p\mathbf{C})^{-1} - \mathbf{I} \quad (5)$$

The summation term in Equation 5 is a geometric series. The reason why we say "approximated" is that some of the links are invalid in the higher powers of $\mathbf{C}$, as we already pointed out in the previous section. The error arising in Equation 5 is minimised, however, on account of p. We do not need to worry about reliability when p is large, because shorter routes (such as single-hop or two-hop routes) will guarantee reliable communication. In case p is small, we need longer routes. But $\mathbf{C}^k$ encodes inaccurate number of alternative routes for large k. At the same time, however, p^k approaches zero, thereby minimising the error introduced in $\mathbf{C}^k$.

Consequently, despite starting off with different premises, it seems that we have arrived at the same result as Katz did. The column sums of $\mathbf{T}$, depicted as a column vector $\mathbf{t}$, indicate how strongly connected the nodes are. Hence, $\mathbf{t}$ can be regarded as a measure of the contribution of the nodes to the overall reliability of the network. We can obtain $\mathbf{t}$ by multiplying Equation 5 with a row vector of unit elements:

$$\mathbf{t}^\intercal = \mathbf{u}^\intercal \left[(\mathbf{I} - p\mathbf{C})^{-1} - \mathbf{I} \right] \tag{6}$$

Multiplying both sides of the above equation by $(\mathbf{I} - p\mathbf{C})$ yields [9]:

$$\mathbf{t}^\intercal (\mathbf{I} - p\mathbf{C}) = \mathbf{u}^\intercal - \mathbf{u}^\intercal (\mathbf{I} - p\mathbf{C}) = p\mathbf{u}^\intercal \mathbf{C} \tag{7}$$

From which we have:

$$\mathbf{t} = \left(\frac{1}{p}\mathbf{I} - \mathbf{C}^\intercal \right)^{-1} \mathbf{s} \tag{8}$$

where $\mathbf{s} = \mathbf{u}^\intercal \mathbf{C}$ is a column vector resulting from summing the columns of $\mathbf{C}$. In other words, $\mathbf{s}$ encodes the number of single-hop links a node establishes with its peers. Equation 7 is not normalised to the potential number of available routes. The normalisation term is proposed by Katz himself, which is given as:

$$m = \sum_{k=1}^{\infty} p^k (n-1)^k \cong (n-1)! p^{n-1} e^{1/p} \tag{9}$$

where n is the total number of nodes in the network. Consequently, the contribution of each node to the reliability of the network can be expressed as:

$$\mathbf{r} = \frac{1}{m}\mathbf{t} \tag{10}$$

Katz's approach breaks in case $1/p$ is not greater than the largest characteristic root of $\mathbf{C}$. This can be verified from Equation 8. In most wireless sensor networks the probability of establishing a link between any two neighbour nodes is independent of the topology of the network. Hence, we modify Katz's *Status Index* by directly normalising the number of sing-hop links a node establishes with its neighbours by $(n-1)$:

$$\mathbf{H} = \frac{\mathbf{C}}{(n-1)} \tag{11}$$

Notice that for a fully meshed network, the normalisation yields 1 for the sum of each column of $\mathbf{H}$. Similarly:

$$\mathbf{H}^k = \frac{\mathbf{C}^k}{(n-1)^k} \tag{12}$$

Thus, the normalised number of single as well as multi-hop links the nodes establish with their peer can be expressed as:

$$\mathbf{T} \approx \sum_{k=1}^{\infty} (p\mathbf{H})^k = p\mathbf{H}(\mathbf{I} - p\mathbf{H})^{-1} \tag{13}$$

		1	2	3	4	5	6
P = 0.2	K	0.221	0.116	0.402	0.221	0.221	0.354
	R	0.096	0.049	0.221	0.096	0.096	0.180
P = 0.7	R	0.570	0.296	1.114	0.570	0.570	0.955

TABLE I: The *Reliability Index* for the network of Figure 1. **K**: The original *Status Index*. **R**: The modified *Reliability Index*.

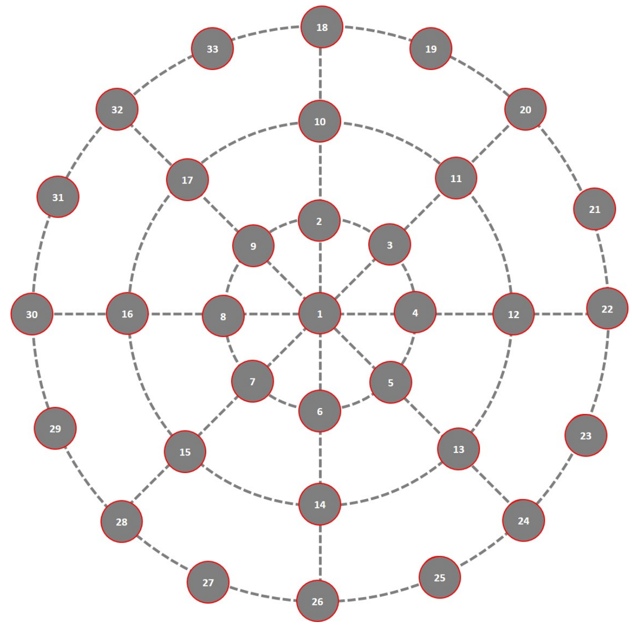

Fig. 3: A moderately large-scale wireless sensor network.

Finally, the modified version of the contribution of each node to the reliability of the network is expressed as:

$$\mathbf{r} = \mathbf{u}^\intercal \mathbf{T} \tag{14}$$

A legitimate question to ask (and answer) at this point is: does the approach scale? In terms of computational complexity, as can be judged from Equation 13, the approach scales. In terms of properly measuring the rank of nodes despite the complexity of the network's topology, referring once again to Equation 13, the approach yields an outcome as long as $\mathbf{H}$ has an inverse, which is the case for wireless sensor networks.

IV. Numerical Evaluation

The normalised – normalised to $(n-1)$ – column sum of $\mathbf{C}$ for the network of Fig. 1 results in a column vector:

$$\mathbf{s}^\intercal = [0.4\ 0.2\ 1.0\ 0.4\ 0.4\ 0.8] \tag{15}$$

The *Reliability Index* we computed for the same network using the original *Status Index* and its modified form are given in Table I. For $p = 0.2$, both approaches yield comparable results and ranked the nodes in the same fashion. For $p > 0.2$, the original *Status Index* breaks, but its modified version does not.

Figure 3 displays a moderately large wireless sensor network wherein nodes are deployed in a polar coordinate system. In order to maintain the connectivity of the network, additional nodes are added in the third tier – assuming that the node in the middle is the base station – of the network. In this

Node's Position	Node ID	Node Degree	R-Index
Base Station	1	8	0.198
First tier	2, 3, 4, 5, 6, 7, 8, 9	4	0.098
Second tier	10, 11, 12, 13, 14, 15, 16, 17	4	0.095
Third tier	18, 20, 22, 24, 26, 28, 30, 32	3	0.070
Connecting Nodes	19, 21, 23, 25, 27, 29, 31, 33	2	0.047

TABLE II: The *Reliability Index* for the network of Figure 3.

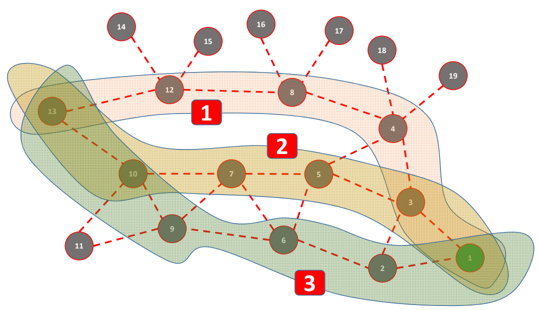

Fig. 4: Determining the most reliable route in a flat-topology network.

topology, the nodes in the first and the second tiers are of degree four whereas some nodes in the third tier are of degree three and some of degree two. Since the predominant traffic flow is towards the base station, the nodes located near the base station are more relevant than those which are farther away. Our *Reliability Index* captures this aspect accurately, as can be seen in Table II.

Likewise, Figure 4 displays a flat-topology network consisting of 19 nodes. Suppose node 13 wishes to send packets to the base station (node 1) using the shortest and the most reliable path possible. We have identified three different routes which have the same hop distance from the base station. The nodes in question and their reliability indices (assuming $p = 0.6$) are given in Table III. In order to identify the most reliable route, we can describe the reliability of a route in different ways. One of them is expressing it in terms of the average *Reliability Index* of the route. In this respect, the first route (the one on top) is the most reliable, as can be seen from Table III. Another option is expressing it in terms of the nodes having the minimum *Reliability Index* (RI) in the routes, because these nodes can be weak links. This metric can be expressed as:

$$\texttt{Best Route} = \texttt{max}\left(\texttt{min}\left(\mathbf{RI}\right)\right) \tag{16}$$

This metric compares the minimum indices of the routes and declares the route having the maximum value as a winner. In this respect, the second route (the one in the middle) is the most reliable route.

V. Conclusion

In this paper we investigated the usefulness of Katz's *Status Index* in ranking the contribution of nodes to the overall reliability of a wireless sensor network. Even though some of the assumptions originally made by Katz are not valid for our case, their impact on the final result, as we demonstrated, is marginal. One of these assumptions is the dependency of p on the overall topology of the network (i.e., the adjacency matrix). We relaxed this assumption by modifying the normalisation factor.

Routes	Relay Nodes with RI	Average Route RI	Total Route RI	min(RI))
R1	12 (**0.143**), 8 (**0.147**), 4 (**0.184**), 3 (**0.151**)	0.156	0.625	0.143
R2	10 (**0.149**), 7 (**0.154**), 5 (**0.155**), 3 (**0.151**)	0.152	0.609	0.149
R3	10 (**0.149**), 9 (**0.151**), 6 (**0.152**), 2 (**0.113**)	0.141	0.565	0.113
max(min(**RI**)) = max(0.143, 0.149, 0.113) = 0.149 : R2				

TABLE III: Routes qualifying metrics expressed in terms of the *Reliability Index*.

The different types of flat as well as hierarchical networks we considered clearly demonstrate the applicability of the approach to a wide range of networks. One of the drawbacks of the proposed approach is that the nodes' physical location does not have any bearing on their status. In most practical wireless sensor networks, however, there is typically a single node (the base station) to which sensed data is transmitted. The base station occasionally sends commands or queries but the predominant traffic flow is from the child nodes to the base station. Since nodes cooperate with one another to support in-network processing and multi-hop communication, the computational and communication burden of those nodes which are placed near the base station is heavier than the burden of those which are farther away. The drawback can be observed in Figure 4 where node 4 is ranked better than nodes 2 and node 3 whereas, in reality, the contribution of these two nodes to the reliability of the overall network is more significant than that of node 4. The expression power of our *Reliability Index* will improve if it encodes this aspect. This will be the focus of our future research.

References

[1] W. Dargie, "Dynamic power management in wireless sensor networks: State-of-the-art," *IEEE Sensors Journal*, vol. 12, no. 5, pp. 1518–1528, 2011.

[2] H. Zhang, C. Jiang, R. Q. Hu, and Y. Qian, "Self-organization in disaster-resilient heterogeneous small cell networks," *IEEE Network*, vol. 30, pp. 116–121, March 2016.

[3] O. Kabadurmus and A. E. Smith, "Evaluating reliability/survivability of capacitated wireless networks," *IEEE Transactions on Reliability*, vol. 67, pp. 26–40, March 2018.

[4] W. Dargie and C. Poellabauer, *Fundamentals of wireless sensor networks: theory and practice*. John Wiley & Sons, 2010.

[5] T. Opsahl, F. Agneessens, and J. Skvoretz, "Node centrality in weighted networks: Generalizing degree and shortest paths," *Social networks*, vol. 32, no. 3, pp. 245–251, 2010.

[6] J. Gao, B. Barzel, and A.-L. Barabási, "Universal resilience patterns in complex networks," *Nature*, vol. 530, no. 7590, p. 307, 2016.

[7] L. Freeman, "The development of social network analysis," *A Study in the Sociology of Science*, vol. 1, 2004.

[8] H. Kim and R. Anderson, "Temporal node centrality in complex networks," *Physical Review E*, vol. 85, no. 2, p. 026107, 2012.

[9] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.