Sweeping Unnecessary Switching Using Weather Data to Improve Resilience in Large Scale Wireless Sensor Networks

Jose Manuel Linares
Lancaster University, Lancaster
Computing Department
Lancaster, UK
Email: j.linares@lancaster.ac.uk

Utz Roedig
Lancaster University, Lancaster
Computing Department
Lancaster, UK
Email: u.roedig@lancaster.ac.uk

***Abstract*—Water utility companies face increasing economic and environmental pressures to optimise their infrastructure, in order to save energy, mitigate extreme weather events, and prevent water pollution. One promising approach consists in using smart systems. However, a smart infrastructure requires reliable communication links which are difficult to provide. In particular, communication links are highly affected by changing weather conditions. Multiple communication transceivers are used to mitigate this issue and to enable nodes to switch to reliable links. Short-term link quality estimators are used to decide which link to use which often leads to the situation where a link switch is initiated which does not prove helpful in the long term. It is not beneficial to switch a link and associated routing for only a brief duration. In this paper we show that long-term link quality estimators based on weather information can be used to augment short-term link estimation to perform more stable link selection. We use experimental data collected over a 6 month duration to show that weather observations combined with classical short-term link quality estimation can provide a more stable network. Our experiments show that the number of unnecessary link switches can be reduced by 98% when incorporating weather data into link selection procedures.**

I. Introduction

Efficiently functioning water infrastructures are essential for modern living standards. They are well known for delivering drinking water, but they also play a key role in handling waste-water and mitigating extreme weather events such as floods and droughts. Managing, maintaining, and updating our water infrastructures is costly and complex. Water infrastructures are heavily distributed (over several 10,000 km^2 in some instances), include a wide range of equipment (pipes, pumps, sewers, vanes, treatment plants, controllers), and have often been constructed over several decades, sometimes going back as far as Victorian times.

Water companies are under increasing pressure to update and improve their infrastructure as environmental regulations are becoming more strict, water consumption surges (mainly due to population growth), extreme weather events occur more frequent and energy costs rise. One important direction of work to address these challenges is the transformation of existing water infrastructures into "smart" systems. The infrastructure is augmented with networked sensors and actuators which are used to implement fine grained automated control. It is expected that a smart water system can provide new capabilities, improve services and reduce operational costs. However, despite promising starts [1], [2], few smart water systems are in production today. One reason, we argue, is the lack of an appropriate integration platform (a "middleware") tailored towards the water industry.

Therefore, we need a communication platform for this middleware that can address these difficulties that current systems have at the moment. We can address this issue by utilizing GSM radio, but this would prove costly on the long term. Another solution would be to use cheaper long range radio transceivers (for e.g. 868Mhz range) to communicate with sensor nodes. Unfortunately this too face challenges as meteorological environments do impact radio quality.

Link Quality Estimators (LQE) can be ctegorised as either hardware or software based schemes. Software based LQE's can provide more fine grain control and able to monitor link quality for certain applications [3]. Some software based schemes such as Kalman filter [4] estimates PSR by calibrating PSR-SNR relationships which give an indication to the link quality. Some schemes rely on RSSI packet drops within a short time frame such as eXtended Transmision Time (ETX) [5] which can give a picture of how the network is performing between two hosts but is considered as reactive and unstable [3]. Other LQE's used involve the use of fuzzy logic [6] in determining link quality, apply four bit in hardware implementations that draw estimations from network, link and physical layers [7] or use mapping methods [4] to determine link quality. In applications where computing power is limited, this alternative is not viable for small Wireless Sensor Networks (WSN) as it is costly in terms of computing power and energy consumption. We therefore need a long range predictor (Weather data) alongside a short range hardware based predictor (dropped packets) to be able to predict when outages might occur, but by understanding the environmental data and implementing a model to carry out simulations.

In understanding the effects of meteorology on radio transceivers, heat was observed as a factor in affecting RSSI strength negatively when the radio transciever was exposed to high levels of heat caused by an increase in resistance within the circuitry of the transceiver [8] but does not analyze other meteorological factors together with humidity and air pressure which is beneficial to better understand this issue holistically [9]. We shall also demonstrate our claim for the need to have multi radio links with an enhanced Link Quality Estimator to predict packet drops and prevent network switching which can result in extra costs which we describe later. We then describe a mathematical formula as to how our method can predict packet drops and prevent unnecesary switching between transceivers.

Our contribution are the following three points:

1) To better understand how meteorological aspects affect radio link quality, explore how it is possible to reduce radio interface switching and suggest how a new Link Quality Estimator can be built utilizing online meteorological data which no existing LQE utilize to this day. This contribution would provide a better understanding on the relationships that meteorological factors play when carrying out switching.

2) Design a Link Quality Estimator (LQE) based on the data provided from the first contribution to reduce costly short term switching. This contribution is a software simulation that was developed and run using real data and use this to better predict when switchin should happen. Running this data has taken several months to reach a result on a modern personal computer.

3) Using the LQE developed to make smart link switching according to our mathematical model. Our mathematical model developed based from findings in contribution 1 and the LQE developed in contribution 2 should predict accurately when switching should take place. We also present results from this final contribution on how our LQE performed.

This paper has 5 sections, section 2 gives an overview picture of how waste water infrastuctures are organised, followed by section 3 giving a description of our testbed and observations on our results. Section 4 explains our weather condition based link selection and we illustrate our mathematical methods based on our observations. In section 5, we give the results of our experiments with tables illustrating our results. We end the paper with our conclusions on our weather based LQE and provide an appendix with tables displaying our results.

II. Waste-Water Infrastructures

Waste-water networks are complex infrastructures that combine civil engineering works (sewers, basins, reservoirs), hydraulic actuators (pumps, gates, valves), sensors (water levels and flows, toxins, gasses), and control devices (for example, Programmable Logic Controllers). The control logic used in waste-water networks is often very simple, relying on fixed threshold values to trigger behaviours (e.g. switching a pump on or off), but more advanced control techniques are now being considered in order to improve infrastructure capabilities [2], [10].

In the following paragraphs, we describe in detail the structure and constituents of existing waste-water networks.

A waste-water infrastructure is usually organised in catchments. Each sewer catchment consists of a connected network of sewer pipes that collect sewage in an area and pumps it to a treatment plant or a discharge point. The number of catchments managed by a water company can be substantial and, taken together, can cover an extensive area. Anglian Water for instance collects waste-water from about 6 millions customers through 1,100 waste-water catchments over an area 27,500 km^2 in the East of England. Many catchments use a *combined sewer system* which collects both waste-water from households and water during rainfall. A combined sewer system must have enough capacity to prevent flooding in cases of heavy rainfall.

Mere gravity is usually insufficient to transport water in a sewer catchment. A catchment is therefore often equipped with a set of pumping stations that transport waste-water over an elevation, so that it can continue to flow under the effect of gravity. A pumping station is built around a wet well; an underground reservoir that acts as a buffer for incoming water (sewage and rain water in a combined system). A wet well is usually equipped with a number of pumps (duty pump, assist pump and storm pump; see [10] for more details of pumps).

The pumps of a wet well can be switched on and off, and must be controlled to process the incoming water, while minimising energy consumption, and optimising the pumps' lifetime. Currently, decisions on pump operations are made locally within one pumping station. Wet well filling levels are monitored and pumps are controlled such that energy consumption is minimised while pump lifetime is maximized [10]. The available water buffer of the wet well is used to bundle and shift times of pump activity. However, the use of the water buffer in the wet well must be managed carefully as capacity must remain to deal with heavy rainfall. The opportunities for energy savings by pump management are substantial. Anglian Water for instance spends about £60 million pounds in energy yearly, with £32 million spent on waste-water operations. Operation costs would be less than the stated figure if communications were to be more effective between pumping stations [10].

III. Radio interface link quality test bed

In this section, we highlight the need for having a multi-radio interface link so that middleware can use the optimal link when weather conditions threaten communication links. We highlight the importance that we want to limit rapid switching between transceivers as much as possible as this can be costly. The map on figure 1 shows the locations of the transcievers; to the south, an xbee node and to the north, a box carrying an xbee with a GSM modem.

A. RSSI strength and effects of meteorological conditions

We carry out this testbed in order for us to guage on the effects that weather environments have over radio propagation; more specifically, we want to know how the 868Mhz band (non licensed band) and GSM bands (licensed) perform in different weather scenarios. During May to December 2014, we organised a test bed where we placed an xbee node 1 Km apart with no line of sight (obstacles such as trees and buildings interfere line of sight) that would send a data packet which is one character in length and would send one transmission; no retries were allowed. The nodes on each side would log a timestamp and the RSSI on each xbee node. One of the nodes had a GSM modem (Huawei E220) that is connected to a GSM carrier. A module was coded to fetch the GSM RSSI

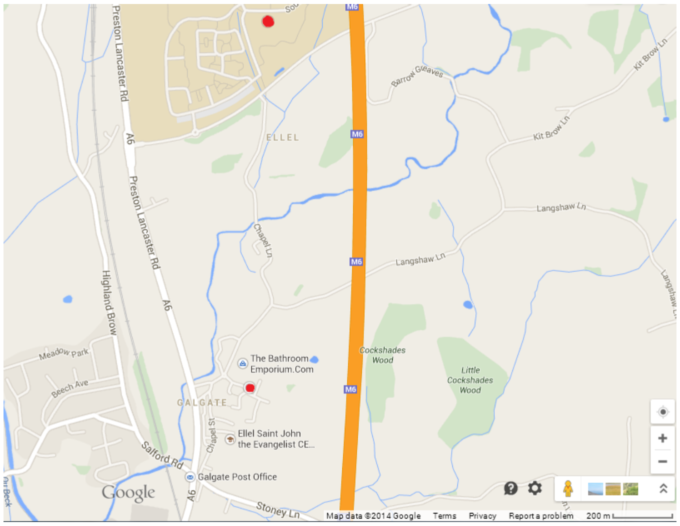

Fig. 1. Map showing locations of 868Mhz sender node in the south and 868Mhz receiver on the north with GSM receiver more than 1Km apart.

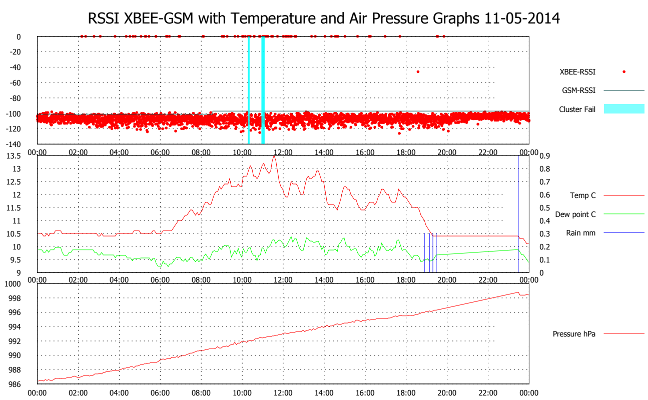

Fig. 3. May 11th 2014

levels which normally receives signals from the carrier every 2 seconds which updates its RSSI level. The 868Mhz radio (Xbee Pro Series 5) RSSI is fetched using local and remote AT commands and the RSSI level is logged together with a timestamp that collects RSSi data once every 30 seconds

During the experiment, various challenges were faced which affected our results. Such errors such as erroneous RSSI fetched from the GSM, hardware bugs present in raspberry pi R1 with power issues, water penetration and equipment failure was experienced due to harsh weather conditions which slowed the experiments. We selected random days during our investigation and these days show significant packet losses as shown in clusters and where the data was reliable to draw conclusions. We observe that there have been days (we include two days in May and several days of June 2014) where xbee RSSI worsened during the day (RSSI -100 or less) and GSM signal was acceptable between -100 and -60 which crosses the boundaries between ”Marginal”, ”Acceptable” and ”OK”.

B. Observations

1) Data Collection and Graphs: We collected data that is split into three separate graphs; the first shows 868Mhz RSSI, GSM RSSI and cluster fail from 868Mhz packet drops if failure rate is more than 30% during 5 minutes. In our setup, no more than 30% failures (this is later represented as 0.3 in our score threshold in section 4) within a predefined period (in our experiments, this was dependent on the frequency that our weather station could produce and this ranged from 5 minutes to 30 minutes). In 2014, we had readings with weather data for every 5 minutes and in 2015, we had weather data once every 30 minutes.

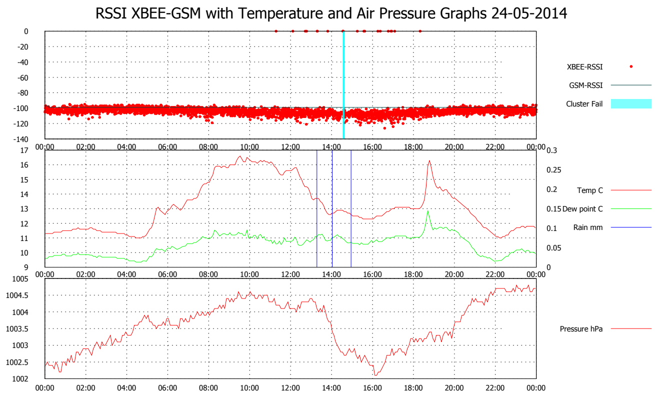

Fig. 4. May 24th 2014

The weather data collected is as follows: the middle graph shows air temperature in degrees centigrade, dew point (the temperature given when water droplets in the air starts to condense hence reaching saturation point) in degrees centigrade and rain in mm. The third graph shows barometric pressure in hectapascals. What we aim to find is if there is some correlation between the RSSI collected on out 868Mhz radio and also observe if meteorological data has any effect to it. If there is such a correlation, we can then use this data in our model to better predict when failures are likely to occur and; 1) pre-empt unnessessary switching (switches or short switching occuring within a short time frame) and 2) saving time, computation, real time actuations and energy costs from making those switches.

The data collected during May 2014 showed that there is no effect on RSSI when it rained; see figures 3, 4, 5 and 7. However, we do observe an improvement on RSSI levels

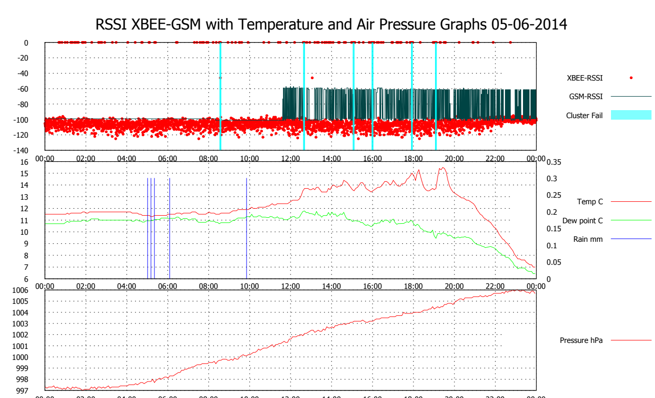

Fig. 5. June 5th 2014

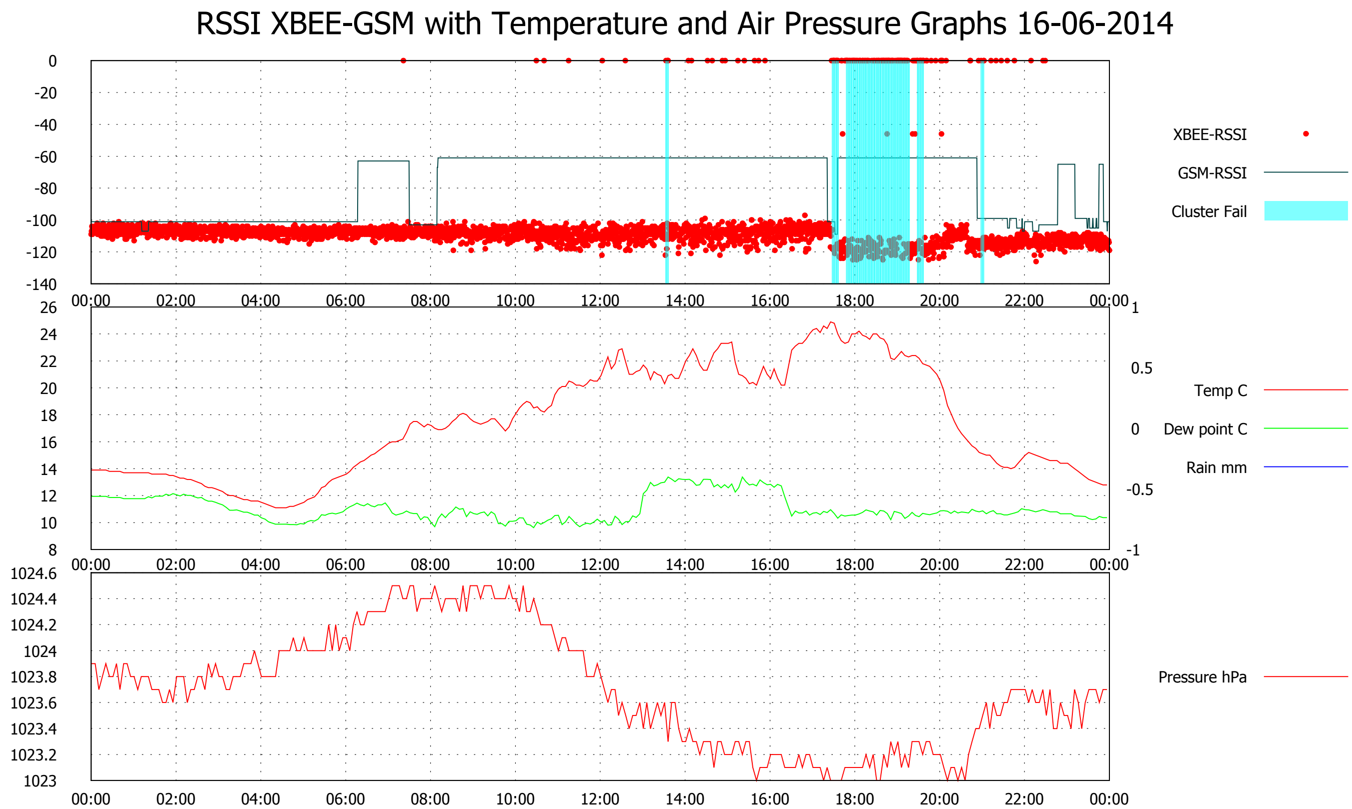

Fig. 2. June 16th 2014

during the evening when the air has cooled down and humidity is rising (The ratio between air temperature and dew point). On this observation, we included this factor which is explained in the next section that takes into account that high humidity levels seems to improve radio propagation at the 868Mhz band. We observed similar occurences in the following months and we noticed that GSM seems to be largely unaffected from environmental factors.

We noticed in June (see figure 2) that there was a large group of cluster failures and shows a drop in RSSI levels whilst temperature was high, humidity was low (dry air) and air pressure dropped. This pattern of a wide RSSI range seems to improve and narrow further, and improve, during the evening which may correlate with what Thelen stated in his paper [11] on improved RSSI levels when humidity was high during the night. We also observe that 868Mhz failure tends to be most prominent during the day when air temperature peak (most notably on figure 2). We do suspect however that changes in air pressure may have a negative effect as we observe that packet failure seem to occur when there are sudden disturbances in air pressure which is an indication that there is a quick shift in air masses that flow from a region of high pressure to a region of low pressure that are either cold or warm [12]. However, during the winter in December 2014, we observed many packet failures and we noticed that air pressure dropped very rapidly and temperatures oscillated rapidly during the day, see figure 7. This behaviour followed for a few days as there was a storm in the north of Scotland which also affected parts of North England. Hence, we were not able to observe significant cluster failure when small changes in air pressure were occuring, however, rapid changes in air pressure does seem to play a role which merits further investiagation to back this claim. On this basis, we need variables assigned for temperature and air pressure and within our formula, express that a high level of temperature, dry air and rapid drops in air pressure, does affect negatively 868Mhz propagation. These variables will be expressed as L which will be our long term predictor.

2) Temperature: We observe that RSSI levels improve when approaching night time during the summer month and we hypothesize that it could be either falling temperatures which correlates with Boano et al findings [8] at the transmitter level. However, we are also looking at the propagation level which heat alone cannot explain hence we look at humidity and air pressure next.

3) Humidity: This observation correlates well with findings from Thelen's investigation [11] with his potato fields in which it was found that humidity seems to improve RSSI levels. During the day when temperatures rise and air becomes drier, we observe that RSSI levels become scattered such as in figure 3. One explanation that we see such improvement is in tropspheric enhancements when we see a lesser inversion relating to a warm ground and cooler humid air that occurs during the summer months. This has the effect of extending signals in the evening and early hours of the morning until the warming of the sun negates this effect. Rain does not seem to have a detrimental effect on the 868Mhz band. As we can observe on figure 4 that there was precipitation and there was no significant packet losses.

4) Air Pressure: In data that we collected during 2014 and 2015, we have observed that pressure also plays a role in determining weather patterns as this ultimately governs meteorological factors. We observe that changes in pressure (especially low pressure and when falling) seems to have a negative effect on the 868Mhz radio link and the algorithm's behaviour is to switch when changes occur. We have seen for instance when the UK suffered from storms and weather bombs (an explosive cyclogenesis where a pressure drop of 24hPa within 24 hours is observable) causing severe gales and long periods of heavy rain often with sleet as shown in this news article [13]. This is visualized in a graph (see fig 7) where there have been packet failures during the whole day. We also observed that changes in pressure and whether the pressure was rising or falling contributed to the effectiveness of the algorithm which is illustrated in table I where pressure was 993.7hPa and pressure was falling.

In the graph, we demonstrate that although it did rain, precipitation alone does not explain why there have been failures during the day.

IV. Weather Condition Based Link Selection

In the previous section we have shown how weather conditions influence link conditions. In this section we look first at how links are usually selected within outdoor communication systems. Thereafter we describe how weather data can be used to organise the link selection process more efficiently.

A. Preliminaries

In many outdoor deployments communication devices are equipped with multiple transceivers of different types. The different transceivers are used to ensure nodes can maintain network connection in case one link (or link type) becomes unavailable. Different link types will have different cost and to minimise cost, nodes will aim to use the cheapest link. In practice, a short range wireless transceiver will incur no cost while a GSM radio will cost depending on the transmitted data volume; in this case the optimal strategy is to use the short range wireless transceiver whenever this is possible. In such a simple two radio system with a static pricing model the cost efficiency can be described by the percentage of time spent using the cheapest link E_{cheap}.

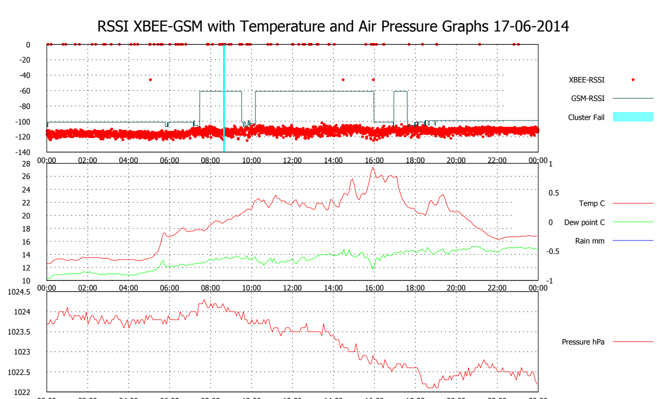

Fig. 6. June 6th 2014

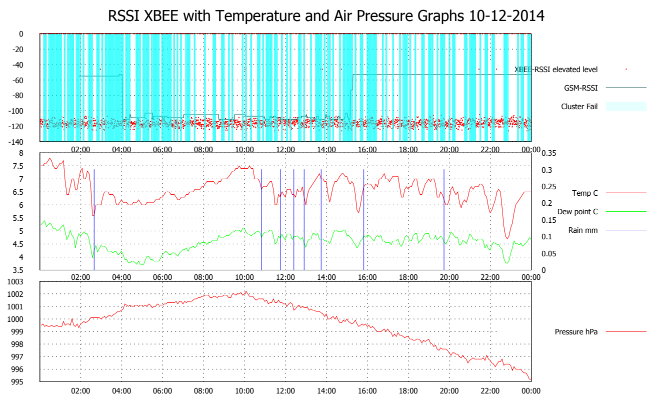

Fig. 7. RSSI severely affected by an explosive cyclogenesis

In addition to link cost, there is as well a non monetary cost associated with link switching. When a node switches from one link (i.e. transceiver type) to another the network must be informed and re-organisation must take place to ensure reachability of the node. Routing tables must be updated or/and all other nodes in the network must be informed of this change. It takes some time to propagate the change during which nodes may be not reachable. It is therefore desirable to reduce the amount of link switches and in particular switches which are of short duration. We define a *short switch* as a link switch operation in which has to be reversed within time T_{short}. Short switches are to be avoided as the decision to use the new link has to be reversed before the link change is fully propagated and effective. T_{short} depends on the design of the communication system and the duration will vary but can be in the order of minutes. The number of short switches observed in a given time period is dented as S_{short}.

In our work we aim to use weather data to reduce S_{short} while maximising E_{cheap}. It is our aim to reduce the number of short switches while still utilising the cheapest communication link as much as possible.

B. Short Term Link Quality Estimators

Link Quality Estimators (LQE) are used to describe how the quality of a link is developing over time. Generally, LQE describe short term link developments. For example, the success rate of recent packet transmissions or Received Signal Strength Indicator (RSSI) of recently received packets is used to estimate link quality for the near future on sender and/or receiver side. LQE are a useful tool to decide if a link is still available for the next packet transmission but they generally do not capture long-term developments of a link.

In this work we use a particular LQE to estimate quality of the short range wireless link. Nodes transmit a probe packet every t_S seconds. The transmitter records if the transmission was successful or not. The LQE (referred to as sort term LQE S) describes the percentage of probe transmissions lost within a time window T_S. A threshold τ_S is used to decide at which point the link is considered unusable. If data is transmitted these can be used as probe transmissions and dedicated link probing is not required.

In this work we used the aforementioned LQE metric, however, other methods could be used to decide if the short range wireless link is useable.

C. Long Term Link Quality Estimators

Link selection based on short term link quality estimation does not necessarily lead to selection of links stable over longer duration. Especially in situations where a link is in a transition phase short term LQE often leads to oscillation between links. We argue that this can be avoided by including long term link quality estimation in the decision. In particular we have shown that weather conditions have an impact on link quality and, therefore, weather information can be used as long term LQE.

As our experiments have shown there is not one single environmental parameter that describes how the link quality is developing over long time periods. Temperature, humidity and pressure have to be analysed in order to estimate in a meaningful way how link quality will be affected. We construct a link quality estimator based on each of the parameters and then combine the result into a single link quality estimator which is subsequently used to decide if a link can be used or not. L_t is the temperature LQE, L_h is the humidity based LQE and L_p is the pressure based LQE. All three LQE are represented as binary value of 0 or 1. We decided to use this binary notation instead of a sliding value between 0 and 1 as we found that this simpler representation is sufficient. This way a weather parameter is either taken into account when making decisions or left out depending on weather conditions. We then use α, β and γ to weigh these three binary LQE values and then combine them to create a single LQE.

Symbol	Purpose
c	Current moving time
t	Temperature value in degrees Centigrade
h	Humidity value in RH%
p	Pressure value in hectapascals (hPa)
T_{trs}	Temperature Threshold in degrees Centigrade
H_{trs}	Humidity Threshold in RH
P_{trs}	Pressure Threshold in hPa
Q_t	Quality Estimator result
T_Q	Time to remain on the expensive transceiver
T_t	Current Temperature
H_t	Current humidity
P_t	Current Air Pressure in hectapascals (hPa)
T_L	Time when S+L algorithm started
α	Truth value determined for L with temperature either 0 or 1
β	Truth value determined for L with humidity either 0 or 1
γ	Truth value determined for L with air pressure either 0 or 1
$S(pf)$	Short term predictor packet failure
$L_t(t)$	Long term predictor for temperature assigned a weighting value if true
$L_t(h)$	Long term predictor for humidity assigned a weighting value if true
$L_t(p)$	Long term predictor for pressure assigned a weighting value if true

We define the short term packet failure in the following manner:

$$S(pf) = \begin{cases} 0.3 > pf \leq 1, & \text{if } pf \geq 0.3 \\ 0, & \text{else} \end{cases} \tag{1}$$

If the packet failure for S_{pf} is greater than 0.3, then it would activate the L component of the algorithm, thus the resulting equation for our LQE can be given as:

$$L(t) = \alpha \cdot L_t(t) + \beta \cdot L_h(t) + \gamma \cdot L_p(t) \tag{2}$$

The temperature related LQE L_t is defined as follows:

$$L_t(t) = \begin{cases} 1, & \frac{T(t-t_t)}{T(t)} < T_{trs} \vee T(t) \geq T_{trs} \\ 0 & \text{else} \end{cases} \tag{3}$$

Here $T(t-t_t)$ is the temperature value recorded t_t seconds in the past. If the ratio of past and current temperature value is above the set threshold T_{trs} the value of L_t is set to 1 as in this case temperature is likely to have an impact on the communication link (see previous section). If the temperature is above a maximum temperature T_{max} is also set to 1 as in this case an impact on the communication link will be noticed.

Similar the value considering humidity is calculated:

$$L_h(t) = \begin{cases} 1, & \frac{H(t-t_h)}{H(t)} < H_{trs} \vee H(t) < H_{trs} \\ 0 & \text{else} \end{cases} \tag{4}$$

$H(t-t_h)$ is the humidity value recorded t_h seconds in the past. H_{trs} is a humidity ratio threshold and H_{max} is the maximum humidity threshold above which an effect on the communication link is to be expected.

The pressure contribution L_p is calculated slightly different. Here we take into account the calculated temperature value L_t as pressure changes are only an indicator for link disturbances if temperature is as well within the considered range:

$$L_p(t) = \begin{cases} 1 \cdot L_t(t), & \frac{P(t-t_p)}{P(t)} > P_{trs} \vee P(t) \leq P_{trs} \\ 0 & \text{else} \end{cases} \tag{5}$$

$P(t-t_p)$ is the pressure value recorded t_p seconds in the past. We only consider pressure if there is a significant change in pressure over time (determined by threshold P_{trs}) or exceeds a predefined threshold.

D. Link Selection Using Short and Long Term LQEs

Using only short term link quality estimators to select a link is not desirable. As we will show this leads to a situation where a link is chosen and shortly after the decision is reverted. As a consequence, links frequently change and no stable routing topology can be established. To avoid this situation we combine short term link estimators and long term estimators as described in the previous section. We give both estimators an equal weight and create an overall link estimator $Q(t)$:

$$Q(t) = \frac{S(t) + L(t)}{2} \tag{6}$$

If $Q(t)$ is above a set threshold the link is selected. The link is then used for a time period T_Q before the expression is re-evaluated to decide if the link should not be used anymore. The time T_Q is calculated using a decaying Euler function which avoids frequent link switching from occuring and is responsible of blocking any decisions to be taken and remain on a reliable radio communication until the time expires:

$$T_Q = e^{\frac{(\frac{-(T_L - c)}{1000})}{Q_t \cdot 10}} \tag{7}$$

The function T_L - x needs to be converted to a positive number by multiplying this result by -1 as subtracting a past time with a future time yields a negative number.

E. Evaluation

During the month of July 2014, we observe several meteorological patterns; 1) During the day, it is typically warm with temperatures over 15 degrees centigrade and colder during the night and 2) Relative humidity is low during the day (dry) and high during the night (air is almost saturated with water vapour). Pressure oscillates and periods when the pressure is high indicates good weather and sudden drops indicates a storm arriving. We then use this data to calculate the factor L.

From the period 01 of July and 05 of July, we observe that S (the metric used for short term switching) is below 30%. We chose this figure as a suitable value between a more aggresive policy which activates GSM more often and opt for a reliable connection or a lenient policy which sacrifices reliability and limits switching. This policy translates that if failures within a 5 to 30 minute period occur, then we calculate L and use both values to determine for how long we should switch to the more expensive and reliable transceiver (GSM). We observe a few instances during the first day of July that factor L was activated because the failure rate for S was greater than 30%. This is drawn on the "xbee On S+L" graph where the coloured lines indicates when the xbee was activated whilst using S+L factors. We also observe that in the graph "xbee On S value" shows when the S value exceeded our value of 30%. On the lower two graphs, we observe no short term switching for S+L but switching does occur for the S factor alone.

During the period of 05 July to 12 July 2014, we observe a more extreme period when switching takes place. The 8th, 9th and 10th of July are days when the S value was high with periods reaching 100% of packet failures occuring within a short time frame of 5 minutes (in 2015, we extend the time frame to be less than 30 minutes). This caused the S value to be greater than 30% and switch from the cheap transciever (868Mhz) to GSM radio. This behaviour is reflected on the "xbee On S+L" and "xbee On S" graphs. We can also appreciate how the S+L graph normalises the S values and thus the decision was taken to use the more expensive radio transceiver as this was a time of instability. We sacrifice more time spent on the cheaper radio and switch to the expensive radio transceiver, but we save on possible network switches taking place unnecessarily within a short period of time (less than 30 minutes).

A high value of Q_t would thus cause the curve to be steeper and cause a long interval which prevents the node from switching back to the more expensive and reliable transceiver. Once the time has elapsed, it would cause the node to re-check the status and re-evaluate if the conditions for the cheaper radio are favourable, or remain on the expensive transceiver (GSM). A lower value $Q_t$would mean that conditions are favourable for 868Mhz and the checking interval is reduced (T_Q).

V. Results and optimization problem

In this section, we describe how we predefine values for our simulation and present an optimization problem as to find the best values that would bring our efficiency close to 100%. In the following tables, we give details of the parameters which we get varying results. The simulation was run for the month of July 2014:

In table I, we will explain what the definitions of each line on the table. Total Traversed readings is the amount of readings that was used as we need to discard data that either has no weather data available, no RSSI data or was skipped by the L blocking algorithm (different parameters can change this value). Used traversed readings is the amount of readings used that was either switched on 868Mhz or GSM and the packet was not dropped.We coined the term "short term switch" as an Evil switch, hence we used Evil Switch S which is the amount of switching that would have taken place on the following scenario; in t1, we received a good packet, t2 was a packet drop and t3 we received a good packet. Evil switching S+L is the algorithm used to determine how many switches there are on the following condition where Ts is the score obtained on a particular time frame and $Score_{trs}$ is the predefined score threashold used:

$$(T_s \geq Score_{trs}) \tag{8}$$

The above statement determines if the L score has exceeded the threshold that was predetermined. If so, it will stop using the 868Mhz radio in exchange for a more expensive radio transceiver.

$$(T_s - 1 < Score_{trs}) \wedge (Ts \geq Score_{trs}) \wedge (T_s + 1 < Score_{trs}) \tag{9}$$

The above equation detects if there is an evil switch in L. This means that the previous time frame, the calculated threshold must be below the set threshold St and that the 868Mhz radio was on and the current time frame (T_s) threshold should be greater than or equals to the set threshold and the future time frame (T_s + 1) should be less than the set threshold. If all conditions pass, then it is considered an evil switch under S+L. The higher the figure between 0 and 1, the longer it will block the decision to either switch or not. We also have times used in S and S+L times expressed in seconds. The bottom part of the table highlights which thresholds and weights were used for the month which are self explanatory. We then display the amount of time in seconds spent on S and S+L respectively.

In our tests, we find that in some months, it can significantly reduce switches taking place - in May 2015, we noticed a reduction from 1407 switches to just 34 in table III which is a reduction of 98.23% switches potentially taking place in May, by using optimal values as shown in table I. We conducted experiments to show that it is possible to profile the next month by running results on the previous month which will yield results that is close to optimal values. We also ran similar tests using different weights in 2014 and 2015 and we found that changing weights can affect the effectiveness of how efficently switching occurs.

One optimization problem we have encountered is that certain value combination can yield a different efficiency result, for example, changing any parameter will either increase or decrease 868Mhz switching efficiency. Another problem we face is that we cannot use the same values throughout the year, although it may be possible to find best values for certain types of weather conditions. We also observe that by altering the values in, we can limit the amount of S+L switching taking place as seen in figures 8 and 9. To improve the reliability of the algorithm, we have run the algorithm to test two months in 2015 which are May and June. It is possible to use the values from the previous month and use them on the next month with a small penalty of 75600 seconds as can be seen in tables I, II and III.

TABLE I. MAY 2015 OPTIMUM PARAMETERS

Output for May 2015	Value
Total Traversed Readings:	88699
Usable Traversed Readings:	68114
Evil Switches on S	1297
Evil Switches on S+L	23
Time spent with S 868Mhz:	1860439
Time spent with S+L 868Mhz:	1606490
Score Threshold:	0.3
Temp Threshold:	19.1
Humidity Threshold:	48.0
Pressure Threshold:	993.7
Temp Ratio:	0.95
Humidity Ratio:	1.02
Pressure Ratio:	1.03
Temp Weighting	0.25
Humidity Weighting	0.25
Pressure Weighting	0.5

TABLE II. JUNE 2015 EFFICIENCY TEST WITH MAY 2015 PARAMETERS

Output for June 2015 exp 1	Value
Total Traversed Readings:	86045
Usable Traversed Readings:	64813
Evil Switches on S	1407
Evil Switches on S+L	39
Time spent with S 868Mhz:	1658949
Time spent with S+L 868Mhz:	1249816
Score Threshold:	0.3
Temp Threshold:	19.1
Humidity Threshold:	48.0
Pressure Threshold:	993.7
Temp Ratio:	0.95
Humidity Ratio:	1.02
Pressure Ratio:	1.03
Temp Weighting	0.25
Humidity Weighting	0.25
Pressure Weighting	0.5

TABLE III. JUNE 2015 OPTIMAL PARAMETERS

Output for June 2015 exp 1	Value
Total Traversed Readings:	86049
Usable Traversed Readings:	65341
Evil Switches on S	1407
Evil Switches on S+L	34
Time spent with S 868Mhz:	1658949
Time spent with S+L 868Mhz:	1325416
Score Threshold:	0.3
Temp Threshold:	32.29
Humidity Threshold:	37.0
Pressure Threshold:	1005.19
Temp Ratio:	0.96
Humidity Ratio:	1.05
Pressure Ratio:	1.03
Temp Weighting	0.25
Humidity Weighting	0.25
Pressure Weighting	0.5

VI. CONCLUSIONS

We have sketched an integration platform for smart waste-water infrastructures and aim to implement the outlined middleware solution and deploy a prototype within the Anglian Water infrastructure. We believe that an integration platform with the capabilities outlined in this paper is necessary to achieve the goal of a smart waste-water infrastructure. Our test bed also proves our claim that there is a need to support our framework with multiple radio links in order to provide resilience when weather conditions worsen and which links are required to work to send commands that operate the duty and storm pumps from the wet well. We then described our mathematical solution to predict S+L and prevent unneccessary evil switching taking place. We demonstrate how parameters in our algorithm can affect the efficiency and time spent on a cheaper radio transceiver, thus saving costs to the user. We also demonstrate that it is possible to deploy this algorithm by profiling the previous month and thus using these parameters on the next month can achieve a reduction in switching costs.

REFERENCES

[1] L. Montestruque et al. Csonet: a metropolitan scale wireless sensor-actuator network. In *Int Workshop on Mobile Device and Urban Sensing (MODUS)*, 2008.

[2] Manfred Schütze, Alberto Campisano, Hubert Colas, Wolfgang Schilling, and Peter a. Vanrolleghem. Real time control of urban wastewater systems, where do we stand today? *Journal of Hydrology*, 299(3-4):335–348, December 2004.

[3] Nouha Baccour, Anis Koubâa, Luca Mottola, Marco Antonio Zúñiga, Habib Youssef, Carlo Alberto Boano, and Mário Alves. Radio link quality estimation in wireless sensor networks. *ACM Transactions on Sensor Networks*, 8(4):1–33, September 2012.

[4] M. Senel, K. Chintalapudi, Dhananjay Lal Dhananjay Lal, a. Keshavarzian, and E.J. Coyle. A Kalman Filter Based Link Quality Estimation Scheme for Wireless Sensor Networks. *IEEE GLOBECOM 2007 - IEEE Global Telecommunications Conference*, pages 875–880, 2007.

[5] Douglas SJ De Couto, Daniel Aguayo, John Bicket, and Robert Morris. A high-throughput path metric for multi-hop wireless routing. *Wireless Networks*, 11(4):419–434, 2005.

[6] Nouha Baccour, Anis Koubâa, Habib Youssef, Maissa Ben Jamâa, Denis Do Rosário, Mário Alves, and Leandro B. Becker. F-LQE: A fuzzy link quality estimator for wireless sensor networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5970 LNCS:240–255, 2010.

[7] R Fonseca, O Gnawali, K Jamieson, and P Levis. Four-Bit Wireless Link Estimation. *HotNets*, pages 0–5, 2007.

[8] Carlo Alberto Boano, Hjalmar Wennerström, Marco Antonio Zúñiga, James Brown, Chamath Keppitiyagama, Felix Jonathan Oppermann, Utz Roedig, Lars-Åke Nordén, Thiemo Voigt, and Kay Römer. Hot Packets: A systematic evaluation of the effect of temperature on low power wireless transceivers. In *Proceedings of the* 5^{th} *Extreme Conference on Communication (ExtremeCom)*, pages 7–12. ACM, August 2013.

[9] Cesar Ortega-Corral, Luis E Palafox, J Antonio Garcia-Macias, Jaime Sanchez Garcia, Leocundo Aguilar, and Juan Ivan Nieto Hipolito. Transmission power control based on temperature and relative humidity. In *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2014 IEEE Ninth International Conference on*, pages 1–6. IEEE, 2014.

[10] S. Ostojin, SR Mounce, and JB Boxall. An artificial intelligence approach for optimizing pumping in sewer systems. *Journal of hydroinformatics*, 13(3):295–306, 2011.

[11] John Thelen. Radio Wave Propagation in Potato Fields. *In Proceedings of the First Workshop on Wireless Network Measurements - WiNMee 2005*, 2:5, 2004.

[12] N Braithwaite S Ross, S Lewis. *Understanding The Weather - S189*. Open University Press, 2009.

[13] 'url: http://www.bbc.co.uk/news/uk-30407295 - weather bomb' hits power and travel in northern uk - accessed: 01/05/2015.

TABLE IV. JULY 2014 BASED ON FIGURE 8

Output for July 2014	Value
Total Traversed Readings:	89018
Usable Traversed Readings:	77474
Evil Switches on S	510
Evil Switches on S+L	129
Time spent with S 868Mhz:	2303637
Time spent with S+L 868Mhz:	1533459
Score Threshold:	0.3
Temp Threshold:	24
Humidity Threshold:	26
Pressure Threshold:	1027.5
Temp Ratio:	0.95
Humidity Ratio:	0.95
Pressure Ratio:	1.03
Temp Weighting	0.2
Humidity Weighting	0.2
Pressure Weighting	0.6

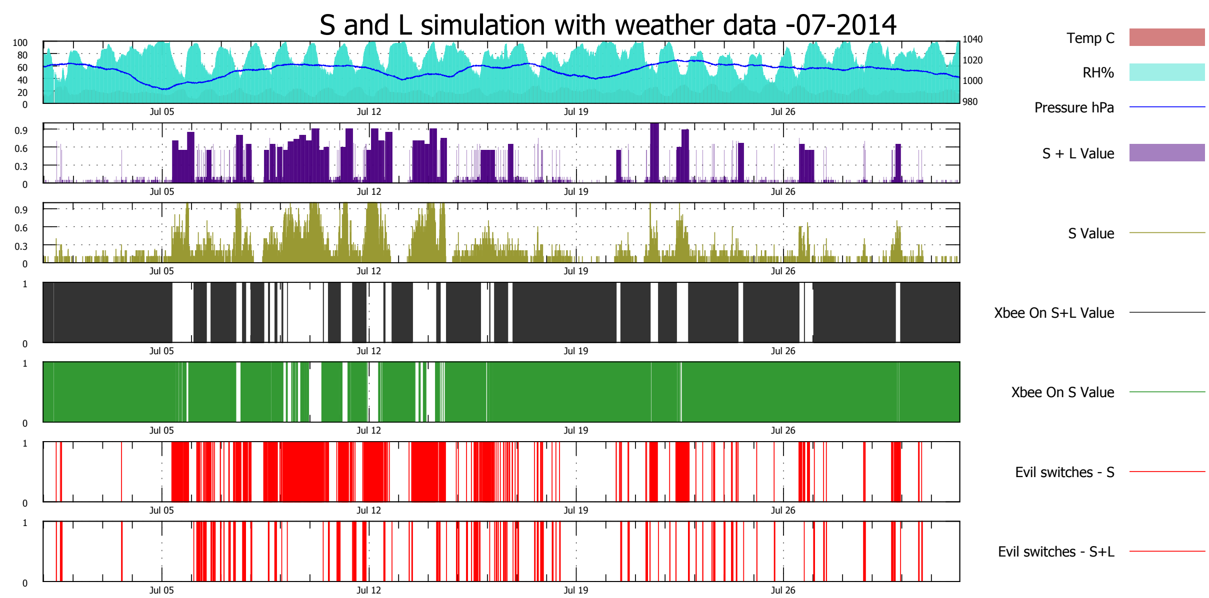

Fig. 8. Simulation run during July 2014 data using values that causes S+L evil switching - table showing details on table IV

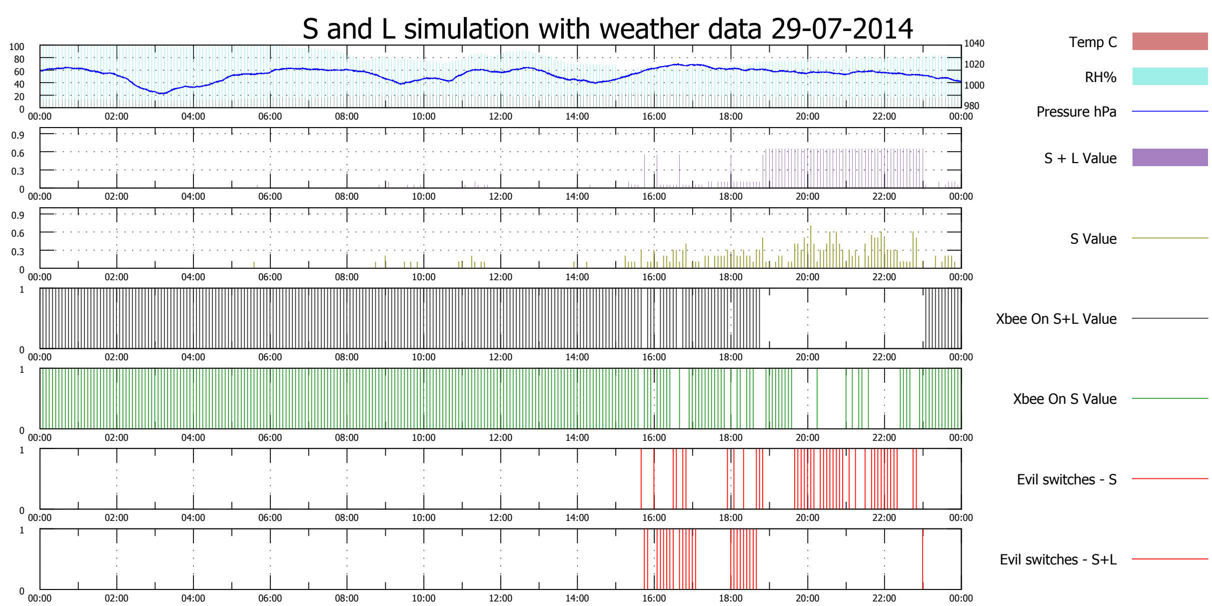

Fig. 9. A day period on the 29th July 2014