

Improving GSM-based Sheep Tracking Using a Cluster-based Multi-hop Approach

Ragnar Stølsmark and Erlend Tøssebro

Institute of Electrical Engineering and Computer Science, University of Stavanger, Stavanger, Norway

Keywords: Sheep Tracking, Wireless Sensor Networks, GSM Coverage.

Abstract: This paper presents a novel sheep tracking algorithm. The current market leading solution, Telespor, performs poorly in situations with low GSM coverage. The algorithm presented here tries to improve the performance of Telespor in low GSM coverage scenarios by using a multi-hop approach for data retrieval. It also uses a cluster-based technique to improve cost and energy consumption. To test the algorithm, simulations were performed based on data collected from a flock of more than 450 sheep. These simulations show that the new algorithm outperforms Telespor in scenarios with low GSM coverage. It also shows that it is possible to get a good ratio of successful updates even with a relatively small amount of full feature nodes. This indicates that the algorithm is an improvement also in terms of cost and energy consumption.

1 INTRODUCTION

Animal tracking has been a popular research topic for years and the combination of GPS and GSM technologies revolutionized the field. In areas with low GSM coverage this tracking method has an obvious shortcoming. In this paper we study a case where low GSM coverage is common: Sheep tracking in Norwegian mountains.

1.1 Norwegian Sheep Farming

Sheep farming in Norway differs from many other countries in that the sheep are not used as milk producers, but solely kept for their meat. The farming is season-based. In the spring new lambs are born. During the summer the sheep and the lambs are sent into the mountains to graze. There are two main reasons for this. First, getting the sheep away from the farm pastures allows the grass to grow so that it can be harvested and used as winter fodder. Secondly, as shown in (Zervas, 1998) it is important to optimize the grazing. The first spring grass is more nutritious than the grass that grows later in the summer season. As spring progresses, a higher altitude is needed to reach this grass and therefore it is better for the sheep to spend the summer in the mountains. In autumn the sheep and lambs are collected. The lambs that are not kept for breeding

purposes are then slaughtered. The rest are kept on the farm throughout the winter.

1.2 The Need for Sheep Tracking

Collecting the sheep after the summer season is the most difficult part of sheep farming. It is highly time consuming, as it traditionally involves hiking through difficult terrain looking for the sheep in what is, by the farmer's experience, considered likely locations. Finding the last few sheep is also a problem. Almost every year some sheep die. Dead sheep lie down and are therefore notoriously hard to locate. Unless the farmer finds them, it will be impossible to know if they are dead or simply missing. This results in an extensive search relying mostly on luck and experience. These problems have been impossible to solve until recent years.

In the last 10 years, cheaper GPS technology combined with the expansion of GSM networks have made it possible to track the grazing sheep. A commercial solution called Telespor has been available for a few years. Telespor is based on the Electronic Shepherd research project (Thorstensen et al., 2004) and is more closely described under the related work section of this paper. Telespor has managed to become quite popular and currently has approximately 30 000 units fitted on sheep and other domestic animals across the country. The success of Telespor comes from a combination of factors. They

were the first commercial solution and time has become much more valuable to people over the last decade. In Norway there is not much money in sheep farming, therefore many farmers keep sheep as a hobby. They have money from their other job, but not much spare time. They are therefore willing to spend some of that money to save time on their sheep farming hobby. Our goal for this project was to see if we could improve GSM coverage, cost or battery life of sheep tracking networks. Our main focus has been on GSM coverage as this is the most pressing issue. To improve the current solution we have created a new sheep tracking algorithm which combines techniques for multi-hop routing and clustered networks. This algorithm has then been compared with Telespor using a simulator based on real sheep data collected by a sheep farmer with 470 sheep.

The paper is structured as follows: Section 2 lists related work with a special focus on the current Telespor solution. Section 3 covers our new sheep tracking algorithm. Section 4 details the sheep data the simulation was based on, as well as the metrics we have used and the results we achieved. Section 5 consist of a discussion of the new algorithm versus Telespor. Section 6 concludes this paper.

2 RELATED WORK

2.1 Animal Tracking

Animal tracking and habitat monitoring has received a lot of research attention. In Oxford, a team of scientists have been monitoring badger behaviour (Dyo et al., 2010). They equipped the badgers with magnets to track them inside their burrows. Sensors above ground measured disturbance in magnetic fields caused by the badger magnets. Researchers in Maine (Polastre et al., 2004) applied wireless sensor networks techniques when they monitored the behaviour of seabirds on a remote island. They installed sensor nodes in the bird nests and returned data via a central base station placed on the island. This technique is a static version of the technique applied in the algorithm described in this paper.

In (Stølsmark and Tøssebro, SENSORCOMM 2012), we presented an algorithm for sheep tracking. This algorithm has a cluster-based approach allowing some of the sheep to carry cheaper nodes and still report their location through full feature nodes. They found that this solution reduced the average energy consumption of the sheep nodes since only the leader in a cluster needed to use GPS.

This approach is also used in the algorithm presented in this paper. In our system, we improve on the cluster-based approach by adding a multi-hop position retrieval method to extend GSM coverage. This multi-hop approach is inspired by the promising results we achieved in (Stølsmark and Tøssebro, ECUMICT 2012). We also studied the possibility of using RSSI triangulation as an alternative way to locate sheep (Stølsmark and Tøssebro, SENSORNETS 2012). The results were disappointing and that is the reason we have chosen not to pursue this approach in our sheep tracking algorithm.

2.2 Telespor

Telespor is the current market leading sheep tracking solution. In the Telespor system, the sheep carry nodes with GPS and GSM-capabilities. The nodes find their own position using GPS and report this position to the Telespor server via the GSM network. The sheep farmer can then watch the location of their sheep on a web application. Through this application, the farmer also has access to advanced features such as adjusting the update interval. Lately Telespor has added a cheaper short (a few meters) range lamb node. This node reports via the sheep node and has no GPS or GSM functionality. The purpose of the lamb node is to let the farmer know that lambs are with their sheep mothers. Fig. 1 illustrates the Telespor system. In step 1 the sheep node receives a GPS position from the GPS satellites. In step 2 the node sends that position to the Telespor server via the GSM network.

2.2.1 Problems with the Telespor Solution

The Telespor system is far from perfect. In a research project (Haugset and Nossun, 2010) farmers tested the system during a summer. The biggest problem was the lack of GSM coverage as the sheep were grazing in rural mountainous areas. A few of the farmers reported that sheep could be outside GSM coverage for weeks at a time. Another issue is with the cost of the units. One full GPRS node costs approximately € 200. This is so expensive that few farmers take the cost of equipping all their sheep with tracking units. They put them on a few sheep instead and hope they will be representative for the rest of the flock. This helps in locating part of the flock but the problem of finding the last few sheep persists. Battery life could also be improved. The longer the battery life, the

more location updates a tracking unit can report. Especially in the collection phase frequent updates is important.

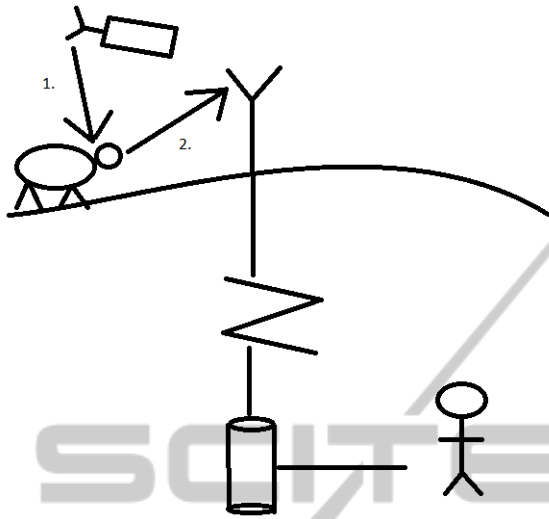


Figure 1: The Telespor system.

3 SHEEP TRACKING ALGORITHM

3.1 Problem Definition and Goals

The purpose of a sheep tracking network is to inform the farmer of where his sheep are or have been at a given time. It should be as cheap as possible so that the farmer can afford it. The network needs to work in areas with varying GSM coverage. It also needs to be sturdy enough to handle the rough treatment it gets from hanging around the neck of a sheep for a few months. Waterproofing is a must since the sheep tracking nodes are exposed to the elements, which in a typical Norwegian summer includes a lot of rain. The network must be able to work unattended for at least 100 days since it should not require the farmer to visit his sheep mid-season. This makes energy consumption important, as the farmer should not have to walk to his sheep to change batteries. The algorithm should be as energy-efficient as possible since the energy savings can be used to improve the sheep tracking nodes in three ways:

- Smaller batteries could be fitted, decreasing the weight and size of the nodes, thereby increasing animal comfort. This would also reduce the price of each unit.
- The tracking season could be prolonged. This could make it attractive for tracking of other

animals.

- The update frequency could be increased, making it more useful for the farmer.

A sheep tracking network needs a localization method. The industry standard is GPS. However, GPS comes with a few disadvantages. It has a high energy consumption and it adds cost to the nodes. Therefore, it could be beneficial to combine GPS with another solution so that not all of the sheep needs to be equipped with GPS or that not all of them need to use it for every position update.

To save energy the algorithm must have duty-cycling between a low power consumption sleep state and an active position update state. If the network should have any internal interaction between nodes, this duty-cycling needs to be synchronized. Synchronization is easily achievable through GPS and is an argument for using GPS in at least some of the nodes.

A typical sheep flock range from 10 – 1000 sheep per farmer. Therefore an algorithm needs to be scalable up to those numbers. It also needs to handle the range between sheep. They typically spread out in small flocks over a large area. It is not unusual for the sheep to graze over a 10 x 10 km area. This area is typically covered by mountains making it far from ideal wireless conditions. The sheep are mostly at a high altitude making the network susceptible to fog and further decreasing the achievable transceiver range. These are the conditions and environment a new sheep tracking algorithm has to deal with. When designing our algorithm we had these application-specific issues in mind.

3.2 Proposed Solution

Fig 2 show an illustration of the new sheep tracking algorithm we have created. When sheep are alone it works in a similar fashion as the Telespor algorithm, with the addition of a small delay. This delay is part of the network discovery process where the sheep nodes wait for messages from other sheep in the vicinity. After this delay, a lone sheep will send a position message over the GSM network to the sheep server. It will then sleep until the next update time. As we see it, it is hard to make any algorithm improvements to the solitary sheep situation. Therefore, we have focused on improving the algorithm in situations where there are a flock of sheep within a small area. Since sheep often travel in small flocks this is a realistic scenario. This has lead us to a cluster-based multi-hop solution, more closely described in section 3.2.1 and 3.2.3.

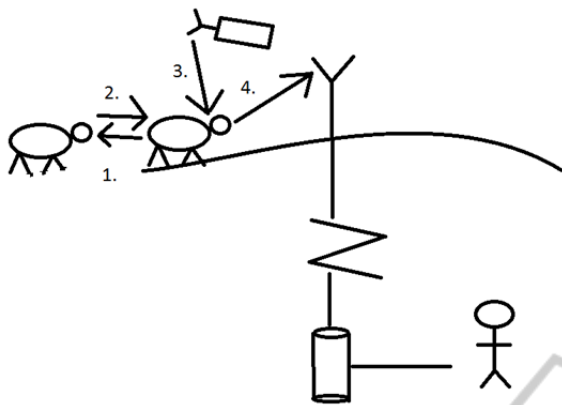


Figure 2: Distributed Algorithm.

Legend:	
1.	Broadcast cluster-head message
2.	Send identity to cluster-head
3.	Cluster-head receives GPS-position
4.	Send identities and GPS-position to server via GSM

3.2.1 Cluster-based Localization

The algorithm divides the sheep into clusters. This is based on the observation of the way that sheep farmers think when speaking of sheep’s locations. They are not interested in the position of each individual sheep, rather they think in terms of clusters. The farmers want to know the position of each cluster and which sheep are in the different clusters.

To take advantage of this cluster-based thinking our algorithm will report only one position per cluster. Together with this position the id of all sheep currently in that cluster are also transmitted back to the server. This way the farmer can get the information he is interested in while at the same time creating a more efficient network. The efficiency comes in the form of energy-saving, since only a few sheep has to use GPS to calculate its position. It also reduces cost since some of the nodes can be without GPS and GPRS modules. These nodes will not be able to report their position on their own, only as part of a cluster. As there is no point in only having the non-GPS nodes, there exists a trade-off between network cost and having a functional network. The denser the sheep flock is, the higher percentage of non-GPS nodes one can have in the network without risking too many lost updates.

The algorithm does the clustering in a distributed manner. Each cluster is within a single network hop from the sheep that is the cluster head and has a cut-

off based on the received signal strength (RSSI). Messages coming from sheep with a low RSSI are ignored. This is because we do not want to make the clusters too big. RSSI is used as an indicator of distance. Our previous research (Stølsmark and Tøssebro, SENSORNETS 2012) has shown that there is a far from perfect relationship between RSSI and distance in hilly terrain, however for this application it is sufficient.

The cluster head will be determined for each position update by the following method:

1. Each sheep will calculate a delay based on the formula in listing 1. This gives a delay that is longer for those sheep with low battery power and those without GSM coverage will get an even longer delay.
2. If no cluster head message is received by a node before the delay expires, they will become a cluster head.
3. On becoming a cluster head they will send out a cluster head message to the other sheep in their vicinity making them part of their cluster.

Once a cluster head is chosen the followers will send their id to the cluster head and the cluster head will find its own position using GPS and transmit it, using the multi-hop approach described in section 3.2.2, along with the id of the cluster members to the server.

```

If(Sheep has no GSM coverage){
    Delay = (1 - battLevel) * maxDelay
            + maxDelay
}Else{
    Delay = (1 - battLevel) * maxDelay
}
    
```

Listing 1: Calculate cluster head delay

3.2.2 Multi-Hop Transmittal of Position

The transmittal of the sheep position back to the server uses a multi-hop approach to maximise GSM coverage. This occurs in a three-step process.

First every cluster head that does not have GSM coverage broadcast their position and cluster members to everyone within their radio distance. This could be to more sheep than are in their cluster, as there is no specific RSSI or distance limit to this broadcast.

The next step is the wait and forward phase. In this phase that occurs for a predetermined time, every node in the network that is not a cluster head or does not have GSM coverage, forwards the messages they receive. Those cluster heads that have GSM coverage stores the messages they receive.

The third step is that all cluster heads within GSM coverage send the position messages they have stored (including their own) to the server. It is important to note that the way cluster heads are selected, via the delay function, ensures that a cluster cannot have a cluster head without GSM coverage unless all cluster members are without coverage.

3.2.3 Complete Solution

Algorithm 1 gives an overview of the complete solution.

```

On Every Update {
    delay = Calculate delay
    Sleep(delay)
    If(No leader message received){
        Notify Neighbours of leadership
        Leader = true
    }Else{
        Send identity to leader
        Leader = false
    }
    If(Leader){
        Find GPS Position
    }
    While(time < Synch Time){
        If(Leader && has GSM Coverage){
            Store received messages
        }Else if(Leader){
            Send own position and
            follower identities to other
            nodes.
            Forward new received
            messages to other nodes
        }else{
            Forward new received
            messages to leader
        }
    }
    If(Leader && has GSM Coverage){
        Send stored messages to base station
        Send own position and follower
        identities to base station
    }
    Sleep until next update
}

```

Algorithm 1: The Distributed Sheep Tracking Algorithm

3.2.4 Hardware Requirements

We have focused on the software side of the solution in this project. However our software has some hardware requirements. Each sheep node must have

- Wireless transceiver
- Processor
- Ram
- Duty-cycling capabilities

In addition each leader node must have:

- GPS receiver
- GSM transceiver

4 SIMULATION SETUP AND RESULTS

To compare our algorithm with Telespor we ran a series of simulations measuring metrics in a realistic scenario.

4.1 Simulation Data

The simulation is based on data collected by a farmer using Telespor on his sheep flock during the 2010 season. The flock consisted of 473 sheep and was located outside the town of Steinkjer in Trøndelag, Norway. The area the sheep grazed in can be seen on the map in figure 4. This area represents a typical Norwegian sheep grazing area. It is rural, mountainous and has few trees due to its elevation.

The data consisted of 389622 measurements. The area the sheep grazed in measured 33 km between the two points that was the furthest from each other. Each measurement included the following data:

- Time and date of the measurement
- Sheep id
- Measured position (latitude and longitude)
- GSM signal strength

We checked the GSM coverage in the area using the sheep measurements to create coverage maps. The maps can be seen in Fig. 3, 4 and 5. These maps were made by colouring the areas where the Telespor nodes measured a GSM signal over a certain threshold. The threshold was low in fig. 3, medium in fig. 4 and high in fig. 5. These maps correspond to the low, medium and high GSM coverage scenarios we have used in the simulations.

4.2 Simulation Setup

We wrote a custom simulator in Java specifically for this project. The simulator simulates both GSM coverage and sheep movement. GSM coverage is simulated based on the values reported by the sheep nodes in the data set. For a given position the GSM coverage value would be calculated as a weighted

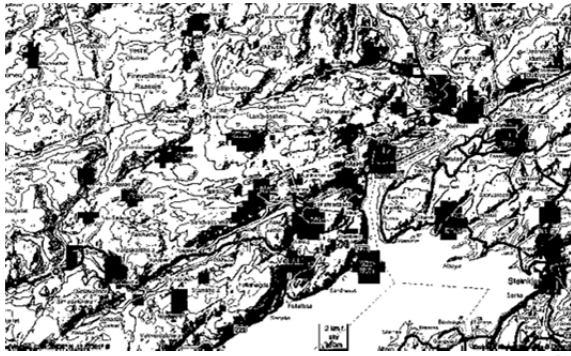


Figure 3: Low GSM coverage map.

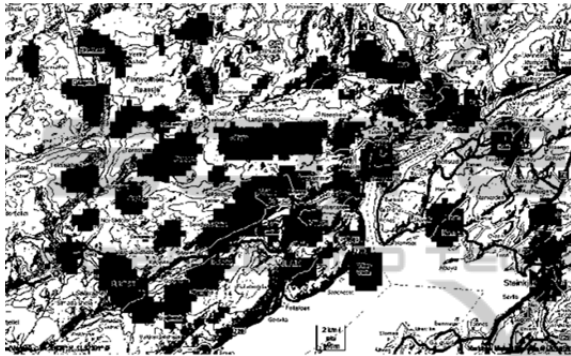


Figure 4: Medium GSM coverage map.

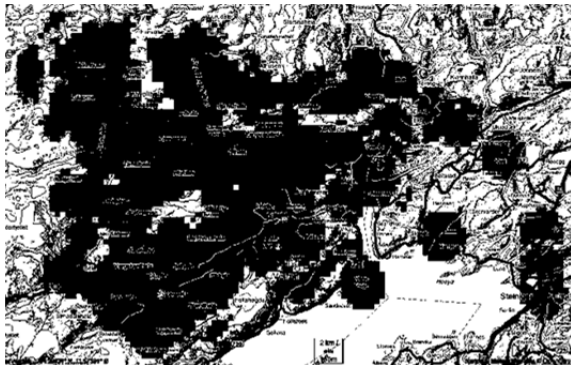


Figure 5: High GSM coverage map.

average of the GSM signal measurements recorded in the dataset within a 2000 m radius. The measurements would be weighted based on how close they were to the actual position. A position would be considered as having GSM coverage if this weighted average was above the coverage threshold defined in the simulation scenario.

Sheep movement is simulated based on the dataset we received. The sheep report their position once per day in our simulator. Their position at the update time is calculated by interpolating the sheep

position between the previous and next position recorded in the dataset.

For every update the sheep ran either Telespor or the new distributed algorithm. When the algorithm was finished, energy was removed from the battery based on which role (follower or leader) the sheep had in the algorithm. Those sheep that ran out of energy would be removed from the next algorithm iterations. Simulation metrics were recorded at the end of each iteration. Time would then advance one day to the next update. This would continue until the end of the recorded sheep grazing season. The simulation parameters are listed in table 1.

Table 1: Simulation parameters.

Parameter	Value
Number of sheep	473
Antenna range	505 m ($\sigma = 170$ m)
Cluster range limit	300 m ($\sigma = 100$ m)
Update interval	1 day
Simulation duration	61 days

4.3 Simulation Metrics

Our simulation metrics are based on what we consider the most important aspects of a sheep tracking system. The primary factor we have looked at is improving GSM coverage. This is one of the biggest problems with the current solution and also the main focus we had when designing our algorithm. As a metric for coverage we have used the number of missed position updates divided by the total number of updates. A position update is considered as missed when using the Telespor algorithm if the sheep is outside GSM coverage at the time of transmittal. When using our algorithm an update is considered as missed if there is no way to transmit the update directly, or via other sheep.

Full feature nodes are more expensive than Telespor nodes since they also need to have a wireless transceiver. The transceiver is used for communication with other sheep. In our system, we also have a simpler node type that is unable to send any position updates without being near a full feature node. These simpler nodes are cheaper than Telespor nodes since they do not have GSM and GPS. The full feature node ratio is therefore important to determine the most cost effective system. In the simulations, we have looked at how small this ratio can be without causing too many missed updates from the simple nodes in different scenarios. We have also looked at what our nodes must cost to have a similar overall cost as Telespor.

Table 2: Missed update ratio (percent).

GSM coverage	Algorithm					
	Telespor	Distributed Algorithm (100% full feature nodes)	D. A. 80%	D. A. 60%	D. A. 40%	D. A. 20%
Low	71.89	0.00	0.00	0.00	0.00	0.02
Medium	43.08	0.00	0.00	0.00	0.00	0.02
High	33.77	0.00	0.00	0.00	0.00	0.02

The final metric we have looked at is energy consumption. This should be as low as possible to allow for the maximum number of updates. Just like the cost metric, the full feature nodes in our system use more power than Telespor and the simpler nodes use less power. The main saving point for the simpler nodes is that they do not use GPS. GPS positioning is one of the main energy consumers in our system and therefore the energy saving potential is substantial. The full feature nodes will use slightly more energy than a normal Telespor node due to the communication with the other sheep. We use a previous analysis of a similar system to make a prediction of the energy consumption of the new solution.

4.4 Simulation Results

4.4.1 Coverage

Table 2 displays the missed update ratio for the different scenarios. Please note that these numbers are based on the simplified assumption that all GSM nodes within GSM coverage can send their updates back to the server. This is a fair assumption as the sheep could resend their messages if packet loss occurred. As the results show, the coverage was vastly improved by the distributed algorithm. The reason for this improvement is the multi-hop message forwarding. Since the sheep forward their position using the full antenna range of 505 meters, a message can travel long distances using multiple hops. Even with just 20 % full feature nodes, only a few updates were lost. This is possible because the cheaper nodes also forward position messages. This means that in a normal sheep flock, it might be enough to have a single sheep with GSM coverage. As an example, Telespor has between 33 and 72 % missed updates with the same sheep positions. This means that in the worst scenario, 28% of the sheep had GSM coverage, with 20 % full feature nodes there will be an average of 26 sheep with GSM coverage. This is more than enough to provide excellent coverage. The improvement was greater

than expected from our results in (Stølsmark and Tøssebro, ECUMICT 2012). This is probably because the real world data has better clustering characteristics than the random placement algorithm used by the simulator in the other paper. That simulator placed sheep either randomly or in the exact same position (cluster). The truth is that sheep walk in clusters, but individual sheep in a cluster still spread out over an extensive area.

4.4.2 Cost and Energy Analysis

We have performed a cost analysis based on the price of the Libelium Wasp mote nodes we have previously used for sensor network testing. A node with the same features as Telespor cost 235 €. A full feature distributed algorithm node cost 250 €. A simple node cost 150 €. Given these prices, the distributed algorithm would have the same price as Telespor with a ratio of 80 % full feature nodes. With 20% full feature nodes the cost would be reduced by almost 30%. This means that the farmer can choose to save money at the expense of localization accuracy. These numbers are meant as an illustration, in a real setting, mass production of standardized nodes would decrease their price significantly.

We performed an energy analysis for a similar system in (Stølsmark and Tøssebro, SENSORCOMM 2012). There we found that the energy consumption could be reduced by as much as 50 % by using cluster-based localization. The drastic reduction is possible because GPS localization is an energy-intensive activity. In a n sheep cluster, each sheep only needs to use the GPS on $1/n$ of the position updates. The cluster-based part of the algorithm therefore saves energy. The multi-hop part will use more energy than Telespor due to the way messages are delivered. The power consuming broadcast function is only used in scenarios with low GSM coverage. The distributed algorithm will therefore in most situations outperform Telespor when it comes to energy consumption.

5 DISCUSSION

The substantial improvement in coverage shown in the simulation results leads to many interesting possibilities. The network can be tailored to suit almost every situation. A farmer can choose to reduce cost at the expense of accuracy. Another example is a farmer who knows his sheep is in an area with poor GSM coverage. He can then choose to increase the number of GSM nodes. This will increase the probability that one of them is in an area with coverage.

The cluster-based approach may seem similar to the Telespor lamb and sheep node solution. The difference lies in the range. The Telespor sheep node has a range of a few meters compared to the cluster range limit of 300 meters used in the simulations presented here. This makes it possible to have much larger clusters, reducing cost and energy consumption significantly. The lamb and sheep node solution does not have any communication between sheep (cluster heads). This makes our algorithm more suitable for low GSM-coverage scenarios, since it benefits from message forwarding.

The simulation results are promising, but there are some scenarios where even the multi-hop approach will not work. The typical situation involves a sheep straying away from the rest of the flock and staying in an area with no GSM coverage. This situation is impossible to improve without a different localization method. A possible solution for these situations could be satellite communication.

The energy consumption is a concern in scenarios with low GSM coverage. Nodes can still benefit from the cluster-based localization, but will spend a lot of energy forwarding messages. We are certain that farmers would still prefer using that extra energy on knowing the location of their sheep. If a sheep is walking into an area with low GSM coverage it could mean it is straying away from the rest of the flock, making it important to know where it is heading.

6 CONCLUSIONS

The new distributed sheep tracking algorithm is a vast improvement over Telespor in low GSM coverage scenarios. In a scenario where only 1 in 4 Telespor updates reaches its destination, the distributed sheep tracking algorithm has almost no missed updates. The cluster-based approach means farmers will still be able to tell where their flocks

are, but makes it possible to reduce energy consumption and cost considerably. This system therefore address the two primary concerns sheep farmers have with the current system, cost and coverage.

ACKNOWLEDGEMENTS

Egil Magnus Pettersen in Telespor for providing sheep testing data.

REFERENCES

- Dyo V., Ellwood S. A., Macdonald D. W., Markham A., Mascolo C., Pásztor B., Scellato S., Trigoni N., Wohlers R. and Yousef K., 2010. Evolution and sustainability of a wildlife monitoring sensor network. *Proc. of the 8th ACM Conference on Embedded Networked Sensor Systems (SENSYS '10)*, 127-140, ACM.
- Haugset A. S., and Nossum G., 2010. Erfaringer med bruk av elektronisk overvåkningsutstyr på sau i 2010. *Notat 2010:17 Trøndelag Forskning og Utvikling AS*.
- Polastre J., Szewczyk R., Mainwaring A., Culler D. and Anderson J., 2004. Analysis of wireless sensor networks for habitat monitoring. *Wireless Sensor Networks*, 399-423, Springer.
- Stølsmark R., Tøssebro E., 2012. Can a multi-hop solution improve GSM coverage for tracking networks *ECUMICT 2012*. 95-106. Nevelland.
- Stølsmark R., Tøssebro E., 2012. Reducing Energy Consumption in a Sheep Tracking Network Using a Cluster-based Approach. *SENSORCOMM 2012*. 129-135. IARIA.
- Stølsmark R., Tøssebro E., 2012. Uncertainty in Trilateration. Is RSSI-based Range Estimation Accurate Enough for Animal Tracking? *SENSORNETS 2012*. 237-241. ScitePress.
- Thorstensen, B., Syversen, T., Bjørnvold, T. A., Walseth, T., 2004. Electronic Shepherd – a low cost, low. bandwidth, wireless network system. *Mobisys '04. Proceedings of the 2nd international Conference on. Mobile Systems, Applications, and Services*. 245-255. ACM.
- Zervas, G., 1998, Quantifying and optimizing grazing regimes in Greek mountain systems. *Journal. of Applied Ecology*, 35, 983-986, Wiley.