Abstract—Highly assisted and Autonomous Driving is dependent on the accurate localization of both the vehicle and other targets within the environment. With increasing traffic on roads and wider proliferation of low cost sensors, a vehicle-infrastructure cooperative localization scenario can provide improved performance over traditional mono-platform localization. The paper highlights the various challenges in the process and proposes a solution based on Factor Graphs which utilizes the concept of topology of vehicles. A Factor Graph represents probabilistic graphical model as a bipartite graph. It is used to add the inter-vehicle distance as constraints while localizing the vehicle. The proposed solution is easily scalable for many vehicles without increasing the execution complexity. Finally simulation indicates that incorporating the topology information as a state estimate can improve performance over the traditional Kalman Filter approach.

I. INTRODUCTION

For the success of an autonomous vehicle, precise localization is an important criterion. So far the vehicles relied on fused data, between proprioceptive sensors like Odometry and GPS (representing motion), and exteroceptive sensors like Camera (representing environment), for self localization. However, different types of sensors have different strengths and weaknesses. For example, GPS requires line of sight visibility with the satellites and is susceptible to interference. Thus the GPS should be fused with other sensors such as inertial measurement units [21] in order to maintain a navigation solution in areas of poor sky visibility (such as parking garages, urban canyons etc.). Similar challenges face other sensors, e.g. additional processing is required for vision perception in dim light situations [19] and the operation of laser scanner sensors can be affected by target material.

One solution to the above problems is to use high sensitivity (and hence high cost) sensors that can work in extreme situations. However, vehicle manufacturers may be unwilling to consider this approach as it will likely raise the cost of the vehicle. Other solution in such situations is to fuse data acquired from sensors present outside the vehicle. This can be either from other vehicles or infrastructure sensors, that can provide us with additional information to help the vehicle with self localization. This methodology is called Cooperative Localization (CL). With improvement in Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) [1] communication techniques, sharing such sensor data, including observations and state estimates in real time has become possible. Thus CL is increasingly becoming a viable and beneficial [1] solution in such scenarios.

Many solutions already exist which take advantage of such CL such as Extended Kalman Filter (EKF) [5], [7], Maximum Likelihood Estimation (MLE) [3], Maximum A Posteriori Estimation (MAP) [4], Markov Localization [6], split covariance intersection filter [8], random finite set framework [10] and Symmetric Measurement Equation (SME) Filter [11]. All these solutions provide novel ways of solving the problem, but still lack one or more of the following points:

- Bandwidth Limitations: Most of the solutions require not only states but also corresponding covariances to be sent on the network. This requirement becomes profound soon as the number of vehicles increase as all of them contend for the bandwidth. Hence minimum data exchange is the key to solve this issue. Howard, Andrew et al. [3] used MLE which uses minimum bandwidth but relies on identification of other participants for CL.
- Data Association Uncertainties: Correct identification of vehicles is required for multiple vehicle localization. Although V2V and V2I technologies support this feature, if there is no measurement to track association, it is difficult to associate the state estimates. Also the real environment is far from perfect and often contains clutter, making the correct data association task challenging. Therefore a solution should address this uncertainty. Montemerlo, M. et al. [20] used FastSLAM to perform localization in case of unknown data association.
- Coordinate Transformations: All sensors provide the measurements with their own frame of reference. For example GPS measurements are in global coordinates system, Odometry gives relative measurements and Camera has its own frame of reference. Fusing all the data together requires the coordinate transformations to one common frame of reference. This becomes difficult with highly dynamic infrastructure environment, in particular when the location and orientation of external sensors is unknown. Zhang et al. [11] used SME filter to avoid such coordinate transformation by converting to symmetric equations independent of coordinates.
- Scalability: As the number of participants for CL increase the solution becomes difficult to manage. Although the SME filter addresses quite a number of issues, it does not scale linearly as the degree of equations increase with the number of participants. Distributed

*This work was supported by BMWi IKT III SADA Project http://www.projekt-sada.de/.

1Dhiraj Gulati and Feihu Zhang are with the fortiss GmbH, München, {gulati, zhang}@fortiss.org
2Daniel Clarke is with The Centre for Electronic Warfare, Cranfield University, Defence Academy of the United Kingdom, Shrivenham, SN6 8LA. d.s.clarke@cranfield.ac.uk
3Alois Knoll is with Technische Universität München, Garching bei München, Germany, knoll@in.tum.de
Conjugate Gradient (DCG) algorithm in MAP is at most quadratic to the number of robots [4]. But it requires synchronous communication with all the participants, which is difficult in a highly dynamic environment.

A method to address all the above challenges using Factor Graph is proposed in this paper. The topology, i.e. the inter-vehicle distance, is used as a constraint between factor graphs. The outline of the proposed methodology is as follows: Measurements from internal sensors, like GPS, for the observed vehicles are plotted as factor graphs. Then the topology constraints (created from infrastructure sensor, like RADAR) are added between the factor graphs for various vehicles. A smoothing algorithm like the Levenberg Marquardt Optimizer is applied on the resulting graph to obtain the fused states.

The advantages of using factor graphs are as follows: (a) Bandwidth requirement remains optimal as constraints are built on a simple global coordinate system with minimal data for covariances; (b) Using topology information helps to avoid the required data association; (c) Topology or inter-vehicle distance is independent of coordinate transformations; (d) It is scalable for any number of vehicles as the maximum degree of equations is quadratic.

Additionally, using an infrastructure sensor as part of CL reduces the costs for vehicle manufacturers, as then vehicles can be equipped with minimum number of off-the-shelf sensors.

II. PROBLEM DESCRIPTION

Fig. 1(a) depicts a simple localization scenario which we would like to solve with Vehicle-Infrastructure CL. The assumptions for the problem are:

- Each vehicle is equipped with some internal sensors to localize itself in absolute reference. Here the 2D global coordinate system is chosen as frame of reference. These values can be obtained by sensors like GPS.
- The external or the infrastructure sensor can measure the relative positions of the vehicles in its own local 2D coordinate system. These values can be obtained by infrastructure sensors like RADAR. Also there is no prior information about the configuration of this sensor, i.e. its location and orientation is unknown.
- The Vehicle-to-Infrastructure bi-directional communication mechanism to exchange data is present. Further the communication is error free and there is no delay.
- The communication mechanism and/or the protocol are not used to identify the individual vehicles.
- There is no clutter in the environment and no false detection originates from the vehicles.

The goal of vehicle-infrastructure CL is to improve the precision of position estimates. As mentioned in introduction, various solutions have advantages and disadvantages. Broadly these solution have been classified as centralized solutions or decentralized solutions.

In centralized solutions, the whole system is considered as a single system with one Fusion Center (FC). All the participating vehicles or robots send their measurements to this FC. EKF [5], [7], Particle Filter, MLE [3] are some of the techniques used by these FC to fuse data.

On the other hand, within the de-centralized framework, each node in the network generates its own fusion result based on its own measurements and state estimates generated by other nodes within the network. Distributed MAP estimation [4], split covariance intersection filter [8] are some of the techniques used in decentralized solution. Martinelli [13] proposed a solution based on a hierarchy of EKFs.

All of the above solutions assume that both the FCs and individual nodes know in advance the transformation between global measurement and the relative measurement. CL thus becomes a function of rigid architecture, while the goal is to work in a dynamic environment. In dynamic environment, the configuration of exteroceptive sensor is often unknown, and scalability, with the increasing number of vehicles, becomes a challenge.

To the best of our knowledge, solving all the problems simultaneously remains an active theme of research in the academic community. The next section describes the concepts of Factor Graphs and how it can be used to address the identified problems to achieve a feasible solution for vehicle infrastructure CL.

III. FACTOR GRAPHS

A. Overview of Factor Graphs

A factor graph is a bipartite graph that explains the connection between the complex functions with many variables and its factors of simpler functions. Fig. 1(b) represents an example of a simple factor graph with variables \( x, y \) and \( z \) and functions \( f \) and \( g \) with factorization:

\[
h(x, y, z) = f(x, y) * g(y, z).
\]

A Factor Graph also represents Probabilistic Graphical Model (PGM). Therefore, the complex function can be seen as the full joint distribution over all the variables and simpler functions can be seen as conditional distributions over the subsets of the variables. The above example can also be seen as:

\[
P(x, y, z) = P(x, y) * P(y, z)
\]

Now to obtain a precise position of a target using the infrastructure sensor, respective target must be tracked. In
the case of multiple targets, additional data association is required to map multiple tracks to respective targets. Various solutions exist to solve this problem like Joint Probabilistic Data Association [22], Probability Hypothesis Density (PHD) filter [23] and Multi Hypothesis Tracking [24]. Within the frameworks presented by these techniques, the computational complexity increases exponentially as the number of targets. Furthermore, these solutions don’t take into account the topology of the vehicle group.

Within the framework of PGM, the topology can be considered to be a factor within the graph, thus representing an additional state within the joint state estimate. One of the earliest usages of factor graphs involved the calculation of sum-product algorithm [14]. Indelman et al. [15] used factor graphs for multi-sensor information fusion for navigation. Cunningham, A. et al. introduced the concept of application constraints as factors between nodes [16].

A specific contribution of this paper is to apply the concept of constraints originating from the topology of the vehicles on factor graphs. The task of smoothing through optimising the resultant graph provides an estimate of the topology state.

B. Factor Graph formulations

Mathematically, a factor graph is a bipartite graph $G_k = (F_k, V_k, E_k)$ with two types of nodes: factor nodes $f_i \in F_k$ and variable nodes $v_j \in V_k$. Edges $e_{ij} \in E_k$ can exist only between factor nodes and variable nodes, and are present if and only if the factor $f_i$ involves a variable $v_j$ [14]. A factor graph for graph $G$ can also be expressed as:

$$f(X) = \prod_i f_i(X_i),$$

where $X_i$ is the set of all variables $x_j$ connected by an edge to factor $f_i$.

Each factor $f_i$ also has an error function that represents the error between the predicted measurement and the actual measurement. This error function is minimized by adjusting the estimates of the variables $X$ using the non-linear least square optimizer like Levenberg Marquardt Optimizer. The optimal estimate $\hat{X}$ is then obtained by optimizing the complete graph $G$ as:

$$\hat{X} = \arg\min_X \left( \prod_i f_i(X_i) \right)$$

In comparison, the EKF uses a measurement model $h(.)$ which predicts a sensor measurement from a given state estimate. The factor, in factor graphs, then represents this measurement model. For a Gaussian noise model, a measurement factor can be written as:

$$f_i(X_i) = d[h_i(X_i) - z_i],$$

where $h_i(X_i)$ is the measurement model as a function of the state variables $X_i$; $z_i$ is the actual measurement and the operator $d(.)$ represents a cost function. The process model can be similarly represented as factor graph (more detail is provided in [17]).

C. Factor formulations for factor graph

In this section we show how factors are generated for different sensor measurements. For our experiment we have factors from Odometery, GPS and RADAR Sensors.

1) Odometery Measurements: We use a constant velocity model while formulating the odometery factor. The odometery measurement equation is given by

$$x_{i+1}^o = x_i^o + v^o \delta t + n^o,$$

where $n^o$ is the measurement noise, $v^o$ is the constant velocity and $\delta t$ is the time interval for which we want to calculate the distance.

The current odometery measurement depends on past measurement, therefore for $x_i$, we can write:

$$x_{i+1}^o = h^o(x_i^o),$$

where $h^o$ is the function which calculates the odometery measurement at time $t_i+1$. Therefore the odometery connects states (represented by nodes in the graph) at two consecutive time intervals, so it is formed as a binary factor.

An odometery factor $f^{ODOM}$, of a given measurement $x_i^o$, is defined by the error function as follows:

$$f^{ODOM}(x_{i+1}^o, x_i^o) \triangleq d(x_{i+1}^o - h^o(x_i^o))$$

Fig. 2(a) illustrates a factor graph with three state nodes and two odometery factors.

(b) Factor graph with three state nodes, two odometery factors and three GPS factors.

(c) Factor graph for $n$ vehicles with three state nodes, two odometery factors, three GPS factors and three topology factors each.
This factor is only connected to the node \( x_i \) that represents the state at time \( t \). Fig. 2(b) illustrates a factor graph with three state nodes, two odometry factors and three GPS factors.

3) Topology Measurements: RADAR is used as the exteroceptive/infrastructure sensor for our topology factor. As mentioned earlier, the configuration (i.e. orientation and location) of this sensor is unknown, the topology information (the distance between the vehicle) at time \( t \) can be calculated as follows:

\[
y_{it}^2 = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (p_{x,y}^{i,t} - p_{x,y}^{j,t})^2 + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (p_{y,t}^{i} - p_{y,t}^{j})^2
\]

\[
(9)
\]

where \( p_{x,y}^{i,t}, p_{y,t}^{i} \) represents the \( x \) and \( y \) position of \( i^{th} \) vehicle as observed by RADAR.

Irrespective of the coordinates systems, global or local, the distance between any two vehicles remains same. Equation (9) is taken as the measurement for topology factor. Also (9) can be formulated as:

\[
(y_{\text{top}}^{i})^2 = h_{\text{top}}(x_0, \ldots, x_N) + n_{\text{top}}
\]

\[
(10)
\]

where \( n_{\text{top}} \) is the measurement noise and \( h_{\text{top}} \) is the new measurement function, relating between measured positions of all the \( N \) vehicles and the new topology measurement. Equation (10) also highlights the fact that the measurement is directly proportional to number of vehicles in the system. Therefore the factor is \( N-\alpha y \) factor, where \( N \) is the number of vehicles in the system. Equation (10) results in factor \( f_{\text{top}} \) written as:

\[
f_{\text{top}}(x_{\text{top}}) = d((y_{\text{top}}^i)^2 - h_{\text{top}}(x_0, \ldots, x_N)).
\]

\[
(11)
\]

This factor is connected to all the states \( x_i \) at time \( t \). Fig. 2(c) illustrates a factor graph for \( n \) vehicles with three state nodes and two odometry factors and three GPS factors each. The corresponding state nodes for the vehicles are also connected to each other with topology factor.

D. Smoothing

The non-linear problem represented by factor graphs is solved by using the linearization algorithm using Levenberg Marquardt Optimizer. Using an initial estimate \( x_0 \) it iteratively finds an update \( \Delta \) from the linearized system:

\[
\arg \min_{\Delta} J(x_0) \Delta - b(x_0)
\]

\[
(12)
\]

where \( J(x_0) \) is the sparse Jacobian Matrix at the current linearization point \( x_0 \) and \( b(x_0) = f(x_0) - z \) is the residual for given the measurement \( z \). The Jacobian matrix is equivalent to a linearized version of the factor graph, and its block structure reflects the structure of the factor graph. After solving (12), the linearization point is updated to the new estimate \( x_0 + \Delta \). Further detail on this process is presented within [15].

In our scenario the jacobian for Odometry is calculated from (5). Each position \( x_n^2 \), in \( 2-D \) plane refers \( x \) and \( y \), i.e. \( h^o(x_n^2) = h^o(x_1, y_1) \) Therefore Jacobian is obtained with \( \partial x \) and \( \partial y \) as:

\[
\frac{\partial (h^o(x_n^2))}{\partial x} = \begin{bmatrix} \frac{\partial (h^o(x_1, y_1))}{\partial x}, \frac{\partial (h^o(x_1, y_1))}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

\[
(13)
\]

The jacobian for GPS from (7) is same as that of odometry. Similarly the jacobian for topology measurement from (9) with \( \partial x \) and \( \partial y \) is:

\[
\frac{\partial (h_{\text{top}}^i(x_n^i))}{\partial x} = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} 2 * (p_{x,y}^{i,t} - p_{x,y}^{j,t})\]

\[
(14)
\]

E. Topology Measurement Uncertainties/Covariances

Measurements of Odometery and GPS can be obtained from the sensors directly. The corresponding uncertainties/covariances are provided by sensor manufactures. But the topology measurement is derived from the infrastructure sensor measurement. Assuming \( \sigma_x^2 \) and \( \sigma_y^2 \) as the \( X \) and \( Y \) covariances respectively for infrastructure sensor, then the corresponding matrix for the measurements from sensor is a diagonal matrix which can be written as:

\[
\text{Cov}(x, y) = \text{diag} \left[ \sigma_x^2, \sigma_y^2, \sigma_x^2, \sigma_y^2, \ldots, \sigma_y^2 \right]
\]

\[
(15)
\]

Then using (9) and (15), we obtain the covariance for the topology estimate at any time \( t \) as:

\[
\sigma_{\text{top}}^2(x, y) = M * \text{Cov}(x, y) * M^T
\]

\[
(16)
\]

where \( M \) is a \( 1 \times 2 N \) matrix as follows:

\[
M = \begin{bmatrix} \frac{d}{dx_1} (y_1^2), \ldots, \frac{d}{dx_N} (y_N^2), \frac{d}{dy_1} (y_1^2), \ldots, \frac{d}{dy_N} (y_N^2) \end{bmatrix}
\]

\[
(17)
\]

IV. SIMULATION RESULTS

A. System Setup

The simulation was implemented with two vehicles on a ground plane for 250 steps. To implement the factor graphs and the corresponding factors we utilize the Georgia Tech smoothing and Mapping (GTSAM) open source library [18].

Simulated vehicles have proprioceptive sensors to measure their location in global coordinates. The exteroceptive sensor located outside the vehicle provides location within its local coordinate system. No configuration information for exteroceptive is available and hence transformation between the two coordinate systems is unknown.

Both proprioceptive and exteroceptive sensors are assumed to have zero mean Gaussian noise. The covariances are assumed as diag[1.0, 1.0], diag[9.0, 9.0] and diag[0.1, 0.1] for the Odometery, the GPS and the exteroceptive sensor respectively. We assume the step interval, \( T \) as 1. It is also assumed that there is no false or missed detections during the whole process.

For Kalman filter, using linear Gaussian Dynamics for constant velocity, the process model is represented as:

\[
F = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\]

\[
(18)
\]
B. Results

Fig. 3 shows the ground truth, the fused trajectory for Odometry and GPS measurements using Kalman Filter and the fused trajectory for Odometry, GPS measurements and Topology Measurements using Factor Graphs for Vehicle 1 and Vehicle 2. Fig. 4 shows the details of part of trajectory for Vehicle 1 during the process. Although the Kalman filter is the industry standard, its clear from Fig. 4 that estimate using the Topology Factor is far superior than Kalman Filter. This is possible because Topology Factors add additional constraints to further reduce the uncertainty of the position.

Fig. 5 analyzes the performance of both methods by calculating Root Mean Square Deviation (RMSE) values for Vehicle 1. The total error is the sum of the RMSE of each vehicle for n steps:

$$E_{\text{Error}} = \sqrt{\frac{\sum_{j=1}^{n} \sum_{i=1}^{2} ((x_{\text{est}} - x_{\text{real}})^2 + (y_{\text{est}} - y_{\text{real}})^2)}{n}}$$ (20)

The initial high total error for Kalman filter is because GPS coordinates have high uncertainty. Although both the methods stabilize almost at the same rate, but the method involving topological factors has lower total error. Hence the Topology Factors perform better than Kalman filter.

The bandwidth requirement is minimal as only observation from exteroceptive sensor are sent on the network. The covariances are not required to be transmitted. If a measurement needs 1 byte for x and y each, then for N vehicles the exteroceptive sensor only needs 2 * N bytes at any step.

The solution is scalable for any number of vehicles because the topology factor (11) remains quadratic for any number of vehicles.

C. Discussion

The results presented here support a solution which meets the initial requirements stated at the start of the paper. The factor graph implementation is low bandwidth and scalable. Further, the introduction of a topology factor reduces the computational and informational burden of data association and allows the coordinate transforms of nodes within the system to be derived as part of the state estimate. It also has the potential of supporting plug-and play paradigm [12].

Overall, the methodology indicates improved performance over the traditional Kalman Filter implementation. However, it is important to note that this improved performance is due to the introduction of the topology factor. The Kalman Filter provides the optimal estimate to the Bayes filtering problem. As all of the nodes estimating the vehicle states are connected with the topology factor, from a Bayesian perspective this has the effect of introducing new common information between nodes and ultimately reduces the overall the uncertainty estimate. Additionally the topology factor avoids the data association issue.

The solution is easily scalable as the Fig. 6 and Fig. 7 show the results for 4 vehicles.

Presently, the solution is implemented as a batch process and all factors influence the joint state estimate. This is not a viable solution at run-time and some form of local smoothing estimate should be used.

Finally, the results presented here assume an ideal environment without clutter, obscuration, false or missed detection
Further work should evaluate the robustness of the solution and without the introduction of new vehicles to the system. Further work should evaluate the robustness of the solution in live scenarios including all of the above challenges.

V. CONCLUSIONS

In this paper, a solution based on factor graphs for vehicle-infrastructure cooperative localization is presented. The solution addresses various challenges for this problem, namely, the bandwidth issue, data association uncertainties, coordinate transformation overheads, and scalability. This has the potential to solve a number of challenges in the highly assisted and autonomous driving communities. The proposed solution is evaluated and discussed using simulated data for constant velocity model. Our simulations indicate improved RMSE performance over the traditional Kalman filter approach. This is achieved through the introduction of a topology factor, which is absent in Kalman Filter, interconnecting all of the nodes within the system.

Future work will focus on the implementation of the presented approach with effects like clutter and obscuration with incremental smoothing, which can be tested in a live scenario incorporating other models like constant turn and constant acceleration. This will also demonstrate the plug and play feature for the dynamic environment.

REFERENCES


[18] https://colab.cc.gatech.edu/borg/gtsam/.


