Privileged Estimate Fusion With Correlated Gaussian Keystreams

Marko Ristic
Autonomous Multisensor Systems Group (AMS),
Institute for Intelligent Cooperating Systems (ICS),
Otto von Guericke University (OVGU),
Magdeburg, Germany
Email: marko.ristic@ovgu.de

Benjamin Noack
Autonomous Multisensor Systems Group (AMS),
Institute for Intelligent Cooperating Systems (ICS),
Otto von Guericke University (OVGU),
Magdeburg, Germany
Email: benjamin.noack@ovgu.de

***Abstract*—Providing cryptographic privacy guarantees in a distributed state estimation problem has been a growing topic of research since the rise in ubiquity of public networks. One such guarantee is having different levels of estimation performance achievable by trusted and untrusted users within a sensor network. In the presence of multiple sensor measurements, guaranteeing better estimation performance by the usual means of adding removable noise to measurements is complicated by an alternative for untrusted users to improve their performance; fusing more measurements. Our novel method adds correlated noise at different sensors, restricting the performace gained from fusing additional measurements, while guaranteeing better performance to those that can remove it. We extend a cryptographic framework for defining estimation privilege and use this to prove the scheme's security goals, while simulations demonstrate the effects of parameters in a concrete estimation problem. A scheme that can ensure such differences in estimation performance between types of estimators can find applications in priority-based or subscription-based performances in environments where more than one sensor is present.**

I. Introduction

Sensor data processing and state estimation have long been active areas of research and continue to find applications in modern systems [1], [2]. In the context of distributed sensing environments such as decentralised autonomous vehicles or distributed weather stations, estimation methods relying on Kalman filters and derivatives [3] are particular prevalent due to their recursive, often optimal, estimation properties and their suitabilty to modeling measurement cross-correlations typically required for data fusion [4], [5]. In recent years, the ubiquity of distributed public networks has seen the additional requirements of preserving algorithm participants' privacies, such as their individual contributions or identifying information, become increasingly relevant and has led to an active field of research [6], [7].

While hiding transmitted information to eavesdroppers can be achieved using common encryption schemes that make ciphertexts indistinguishable [8], distributed estimation tasks that preserve participants' privacies require information to remain hidden during partial or complete processing of the task and often justify some leakage [9], [10]. The security goals of these problems are context specific and have produced a variety of solutions. In [11], non-Bayesian localisation is performed using homomorphic encryption such that individual sensor information and measurements remain private, while in [12], similar goals are achieved in a Bayesian setting by fusing linear measurements when sensors form a heirarchical network. A combination of homomorphic and order-revealing encryption schemes are used in [9] to solve conservative Gaussian estimate fusion while leaking only the ratios of estimate covariances, and in [10], [13], cryptographic aggregation schemes are introduced and used to leak only total consumptions in an energy grid while hiding individual participant power-usages. In addition to these examples of privacy, achieved through hiding and leaking quantifiable information, optimal estimation performance itself can be considered leakage, which introduces a privilege in estimation where leakage defines the difference in performance between trusted and untrusted estimators. In the original Global Positioning System (GPS) [14], this was achieved with a secondary encrypted channel that allowed better estimation performance to parties which held an encryption key. Similarly, in [15], a synchronised chaotic system is used to add noise to measurements which can only be removed by estimators knowing its properties. This idea of estimation privilege is further explored in [16], where a formal cryptographic definition is given and a scheme for a single sensor presented. Here, a synchronised psuedorandom Gaussian keystream adds measurement noise only removable by estimators holding the stream key.

In this work, we consider this definition of estimation privilege, presented in [16], and introduce a modified scheme suitable for an environment of multiple sensors, where both holding the secret key and fusing additional measurements can lead to better estimation performance. Our contribution consists of a generalised notion of estimation privilege, the introduction of security requirements for privileged estimation in a well defined multisensor environment and a scheme that satisfies them. Along with a cryptographic proof sketch, simulation results are provided to demonstrate the effects of scheme parameters and how they can be chosen to provide varying amounts of privilege. Use-cases for varying performance in this way include subscription or priority models where some users are provided better results than others. For example, subscription-based weather forcasts using measurements from

spacially distributed stations or modular mass-produced sensors that differ in accuracy dependent on cost.

In section II, the multisensor estimation privilege problem is presented. Relevant preliminaries are introduced in section III and the estimation privilege fusion scheme itself in section IV. A cryptographic analysis and the simulation results are then given in sections V and VI, respectively, before the concluding remarks in section VII.

A. Notation

- Matrices, vectors.
- Positive definitness.
- Pseudorandom distribution.
- Estimator in cryptographic sense.
- Negligible fuction and negligible covariance.

II. Problem Formulation

- We are considering an environment where multiple sensors are present and required for the greatest estimation accuracy of the system they are measuring.
- We want to provide multiple levels of privilege to estimators, such that estimators with a higher level of privilege can achieve better estimates given the same measurements.
- We consider linear system and measurement models, given by the usual equations, which will make proving relevant cryptographic privileges straightforward. Each sensor i holds its own secret key sk_i, which can be made available to an estimator of a suitable privilege.
- In addition to linear systems, we will make the assumption that the sensors are synchronised in time k to simplify meaningful cryptographic evaluation when recieving measurements. In practice, this would restrict the estimation scheme to scenarios where measurements are taken infrequently and timestep synchronisation is easier to guarantee.
- While we are interested in a cryptographic difference in estimation between estimators who do and do not hold sensor keys, respectively, the involvement of multiple sensors means that access to additional sensors and the fusing of measurements from sensors whose key is not known also need to be considered.
- We want to provide a scheme for estimation privileges that guarantees two types of estimation differences.
- Firstly, we want a lower bound on the difference between estimation performance of an estimator that holds no sensor keys (an unprivileged estimator) but has access to all measurements, and an estimator holding a subset of sensor keys (a privileged estimator) using only measurements from sensors to which they hold a key. This construction means the bound remains a lower-bound when the unprivileged estimator has access to fewer sensors or when the privileged estimator has access to more and exhaustively captures the benefits of knowing sensor keys.
- Secondly, we want an upper bound on the additional estimation performance a privileged estimator can achieve by fusing measurements from sensors to which they do not hold a key. The motivation behind this guarantee is that fusing additional measurements from sensors whose keys are not known should not provide as much estimation benefit as acquiring another sensor key, thus preserving the order of possible estimation performance across privileges.
- The construction of the resulting scheme should be such that two free parameters can be chosen to control the values of these two bounds, respectively.

III. Preliminaries

- When defining our scheme and discussing its cryptographic guarantees, we will make use of Gaussian keystreams and the notion of cryptographic estimation privilege. These have been summarised below.

A. Gaussian keystreams

- A Gaussian keystream is a sequence of pseudorandom Gaussian samples generated using a random key. The sequence is indistinguishable from a truly random Gaussian sequence to any observer that does not hold the key, while it is exactly reproducible if they do hold the key.
- A Gaussian keystream can be constructed from any cryptographic bitstream cipher when making the floating-point randomness assumption.
- Give the assumption about floating-point numbers and why they are reasonable to use in place of truly random real numbers in cryptography proofs.
- Define the equations for turning a bitstream cipher into a multivariate Gaussian stream with multivariate covariance denote S.

B. Cryptographic Estimation Privilege

- The formal definition of a privileged estimation scheme was introduced in (previous paper) and captures a reduction in estimation performance that can always be achieved for an estimator not knowing the scheme key compared to one knowing it and unmodified measurements.
- We will use a slight generalisation of this definition to capture an arbitrary difference in estimation that can always be achieved between two estimators given the measurements they have access to.
- Define a privileged estimation scheme (Setup and Noise) but with modified Noise to allow the case that neither estimator has access to true measurements.
- Define a privileged estimation scheme where series D_k can be positive or negative definite.
- We note that the difference in the generalised definition above is that the estimator for which we bound estimation performance, considered unprivileged, may perform better than the privileged estimator. That is, the difference series D_k may be negative-definite. This

feature will be useful when discussing the additional estimation performance a privileged estimator can achieve by fusing measurements from sensors to which they do not hold a key.

IV. Privileged Fusion

- The idea behind the privileged fusion scheme is to add correlated Gaussian keystreams to measurements from each sensor, which can be removed by estimators holding the respective sensor secret keys.
- As we assume that the sensors are synchronised, we can exactly capture the correlation between the modified measurements by considering the stacked measurement model when having access to j sensors (stacked eq with modified $H^{(j)}$ and correlation matrix $C^{(j)}$, $j \leq n$).

A. Privileges

- Before we introduce the means of modifying measurements with sensor keys to affect the estimation performance for different privileges, we must also define the considered privilege levels in terms of access to sensor keys.
- In this work, we will consider privileges as knowing sequential sensor keys, in order to make correlated noise generation simpler.
- That is, in the presence of n sensors, we will consider exactly n possible privilege levels, where each privilege j corresponds to holding the sequential sensors' secret keys $sk_1, ..., sk_j$ while being unprivileged corresponds to holding no secret keys.
- As is intuitive, each successive privilege level, knowing exactly one additional sensor key, will correspond to better estimation results than the previous levels or the unprivileged case.

B. Noise Generation

- Similarly to the Gaussian keystream introduced earlier, pseudorandom standard normal samples can be correlated when transformed together even when generated using different keys.
- To capture a correlation between additive noises from each sensor, we introduce the fully correlated component Z and the uncorrelated component Y and given n sensors we define the generated noise cross-correlation matrix as S (give equation).
- The generation of the additive noise can then be computed with m standard normal pseudorandom samples generated from keys of each sensor with the following equation (using $S^{1/2}$) and added to the measurements at every timestep k as follows (give measurements equation with two noises).
- In the equation above, computing $S^{1/2}$ requires more than the simpler requirement in single sensor pseudorandom Gaussian keystream ($S^{1/2}S^{1/2\top} = S$)), as the form of the matrix $S^{1/2}$ will affect which subsets of keys are sufficient for partially computing the correlated noises. That is, privileged estimators that only hold a subset of sensor keys should still be able to generate the correlated pseudorandom Gaussian keystream for the subset of sensors for which they hold keys, but not for the remaining sensors.
- To achieve this and support the privileges introduced in the previous subsection, a lower-triangular form of $S^{1/2}$ is required, for which we will use the Cholesky decomposition. In this form, computing $S^{1/2}$ for a subset of sensors $1, ..., j$, $j \leq n$ satisfies the following (equation showing that this matrix is the top left block matrix of the equivalent decomposition for all n sensors).
- Now, an estimator of privilege j (that holds the keys $sk_1, ..., sk_j$, $j \leq n$), can correctly generate the noises for sensors $1, ..., j$ by computing (give partial generation equation up to sensor j).
- The subset of computed noises can then be removed from received sensor measurements $z'_1, ..., z'_j$ by subtracting it.
- Since we are using a cryptographically sound stream cipher, pseudorandom Gaussian samples are indistinguishable from truly random ones to an estimator not holding any relevant keys. We can then consider three estimation problems.
- First, an estimator with access to j sensors and no keys (unprivileged estimator) will have measurements indistinguishable from those following the measurement model given by (model but with additional measurement noise Y and Z).
- Second, an estimator with privilege j and access to the first j sensors $1, ..., j$ can remove all added pseudorandom noise from observable measurements and follows the measurement model given earlier exactly.
- Lastly, an estimator with privilege j and access to the first j sensors $1, ..., j$ as well as some additional l remaining sensors. In this case, noises can be removed from the first j measurements, while the remaining l measurements have conditional Gaussian noise based on the first j pseudorandom samples. This can be captured by the measurement equation (give slightly more complicated conditional measurement equation that will not have zero-mean noise).
- Intuitively, and from the measurement models observed for the different estimators above, we can tell that the parameters Z and Y will control the difference in estimation performance between estimators.
- More specifically, the fully correlated term Z will control the difference in estimation between an unprivileged and privileged estimator, while the uncorrelated term Y will control the difference in estimation between privileged estimators with and without access to additional measurements from sensors to which they do not hold keys.
- These two differences correspond to the two guarantees we want to show for the presented scheme (introduced in the problem formulation section).
- The effects of choosing these parameters will be formally exploring in the cryptography section.

C. Noise Distribution

- While we have described a method for generating noise for n sensors such that different estimators observe different measurement models being followed, we have not discussed where noise is generated and how it distributed to the relevant sensors.
- Due to the correlation between sensor noises in the method above, generation of one sensor's noise is dependant on generated noise from others. To handle this, they can be generated either centrally and distributed to sensors or sequentially by the sensors 1 through n.
- To compute them centrally, equation (generation of n noises given before) is computed by a central processor and each noise $p_{i,k}$, $1 \leq i \leq n$, is sent to sensor i to add to their measurement at time k.
- Sequentially, each sensor can generate its own noise $p_{i,k}$ based on the noises generated before it, that is, given $p_{j,k}$, $1 \leq j < i$, the conditional Gaussian can be computed as (equation for generating sequential correlation). However, a downside of this method is the increasing size of data being sent to each sensor as the noise is generated and can result in later sensors requiring more processing power than may be available.
- In both of the cases however, the computation of noises $p_{i,k}$ can be done offline and in advance, such that sensors have sufficient addative noises already computed prior to commencement of measurement.

V. Cryptographic Privilege

- To give a sketch proof of the cryptographic privilege provided by the presented multisensor scheme, we first recall two important assumptions. We consider cryptographically generated random floating-point numbers to be sufficiently close to real numbers that real number proofs can be used, namely the optimality of the Kalman filter, and that all sensors are synchronised, such that noises $p_{i,k}$ are always added to measurements $z_{i,k}$ from each sensor and that the measurement stacked model (given a the start of the scheme section) is correct for all k.
- The sketch proof will rely on the optimality of the linear Kalman filter to produce series' of covariances that are the best achievable (smallest possible) for a given estimator, and take the difference between estimators in question to bound their difference and achieve cryptographic estimation privilege.
- Similarly to the previous paper, the bounding series can be used in a cryptographic sketch proof which shows that the existence of an estimator violating the bound would imply the existence of a better linear estimator than the Kalman filter, known not to exist. This then guarantees the bound by contrapositive.
- We consider two types of unprivileged estimation which we want to bound, namely estimators holding no keys and estimators holding only a subset of keys, as specified in the desired guarantees.

A. Unprivileged Adversaries

- If we assume an unprivileged estimator can access all n sensors, then their stacked estimation model can be described by the appropriate equation at each timestep k.
- We can then write the combined Kalman predict and update equation as the equation with params from the previous one.
- Due to the KF preserving error covariance order, by setting $P_0 = 0$ we get a series of covariance such that no unprivileged estimator can estimate with error covariances less than or equal to the series (lower-bound).
- Similarly we can do the same for an estimator to which we want to lower-bound the estimation difference with the unprivileged estimator.
- A privileged estimator of privilege j (access to the first j keys) and that can only access the first j sensors, has the appropriate measurement equation. Similarly, setting the initial covariance to zero gives the lower bound series on the best possible estimation error achievable by the privileged estimator using only the measurements from the sensors to which it holds keys.
- Taking the difference of the two estimator bounding series' produces the difference series.
- In the context of cryptographic privileged estimation scheme, the Setup and Noise algorithms are given accordingly. The optimality of the KF can then be used to achieve the security notion.
- In the above, we assume the unprivileged estimator has access to all n sensors, while the j privilege estimator has access to only the first j sensors. In the case when the unprivileged estimator has access to fewer sensors or the privileged one to more, their difference in estimation can only increase, thus keeping the computed lower-bound a lower-bound and the cryptographic guarantee does not change (albeit the definitions of Setup and Noise will, to capture the now available sets of measurements).

B. Privileged Adversaries

- We can use a similar approach to seperately guarantee the largest possible benefit in estimation available to an estimator of privilege j (access to first j keys) when fusing measurements from sensors to which they do not hold keys to get a better state estimate.
- We can again write stacked estimation models for the two estimators. The estimator of privilege j with access to the first j measuremets still follows the model (given in previous subsection).
- The estimator with access to the additional l measurements, however, does not have zero-mean measurement noise as seen in (equation for the addional measurements estmiator given before) required for the Kalman filter. As the mean is known at each k, from the known noises $p_{i,k}$, $1 \leq i \leq j$, remaining measurements can be offset to produce the equivalent measurement model with zero-mean noise (given here).

- Again, using the combined predict and update equations of the KF and setting $P_0 = 0$, we can use the optimality of the KF to give the best possible performances of the estimators as a series of covariances.
- Taking the difference of the bounds now gives a bound on how much better an estimate can become when unprivileged measurements are fused with privileged ones.
- In the context of a cryptographic privileged estimation scheme, the Setup and Noise algorithms can be given as follows. KF optimality can then be used to achieve the security notion.
- We note that unlike in the unprivileged adversary case, or the previous paper, here we do not use unmodified measurements for the privileged estimates and the resulting algorithms provide better estimation for the adversary than the privileged estimator. In this form, we bound how much better an adversary can estimate when using unprivileged estimates resulting in series D_k consisting of negative-definite rather than positive definite matrices. This makes use of the generalised cryptographic definitions earlier.

VI. Simulation

- In addition to showing how to derive the bounds to the benefits of using unprivileged measurements, we have simulated concrete scenarios to demonstrate the methods.
- We consider the following linear system, and linear measurement models for 4 sensors. Correlated and uncorrelated noise covariances are of the form $Y = \sigma_y I$ and $Z = \sigma_z I$.
- Implementation details.
- First figure shows 4 plots. Each plot shows three traces. The traces are: the trace of the unprivileged estimator (no keys, all measurements), the trace of a privilege j estimator (j measurements) and the trace of a fusing privilege j estimators (all measurements). j is equal to $1, 2, 3$ and 4 for each plot, respectively. Here, the difference between the first two traces in figure 4 and the second two traces in figure 1 are equal to the traces of the two D_k series from the cryptography section, respectively. Values for σ_y, σ_z are fixed.
- The second figure focuses on a single plot from the first figure but varies σ_y and σ_z. Again, 4 plots are shown, with σ_y increased in two and σ_z increased in two, such that they are both increased in only one plot. This will show that while both parameters affect the estimation of unprivileged and additional fusion estimators, σ_y and therefore Y predominantly affects the unprivileged estimator and σ_z, Z, affects the additional fusion estimator.

VII. Conclusion

- Concluding remarks.
- Future work includes exploring key subsets that do not need to be sequential, decentralised methods for multi-key correlated noise generation, and the effects and cryptographic guarantees of sensors falling out of synchronisation.

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Dover Publications.

[2] D. Simon, *Optimal State Estimation: Kalman, H Infinity and Nonlinear Approaches*. Wiley-Interscience.

[3] A. J. Haug, *Bayesian Estimation and Tracking: A Practical Guide*. John Wiley & Sons.

[4] A. G. O. Mutambara, *Decentralized Estimation and Control for Multisensor Systems*. CRC press.

[5] M. Liggins, C. Y. Chong, D. Hall, and J. Llinas, *Distributed Data Fusion for Network-Centric Operations*. CRC Press.

[6] K. Ren, C. Wang, and Q. Wang, "Security Challenges for the Public Cloud," vol. 16, no. 1, pp. 69–73.

[7] M. Brenner, J. Wiebelitz, G. von Voigt, and M. Smith, "Secret Program Execution in the Cloud Applying Homomorphic Encryption," in *5th IEEE International Conference on Digital Ecosystems and Technologies (DEST)*, pp. 114–119.

[8] J. Katz and Y. Lindell, *Introduction to Modern Cryptography: Principles and Protocols*. Chapman & Hall.

[9] M. Ristic, B. Noack, and U. D. Hanebeck, "Secure Fast Covariance Intersection Using Partially Homomorphic and Order Revealing Encryption Schemes," vol. 5, no. 1, pp. 217–222.

[10] E. Shi, T.-H. H. Chan, and E. Rieffel, "Privacy-Preserving Aggregation of Time-Series Data," p. 17.

[11] A. Alanwar, Y. Shoukry, S. Chakraborty, P. Martin, P. Tabuada, and M. Srivastava, "PrOLoc: Resilient Localization with Private Observers Using Partial Homomorphic Encryption," in *16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 41–52.

[12] M. Aristov, B. Noack, U. D. Hanebeck, and J. Müller-Quade, "Encrypted Multisensor Information Filtering," in *21st International Conference on Information Fusion (Fusion 2018)*, pp. 1631–1637.

[13] M. Joye and B. Libert, "A Scalable Scheme for Privacy-Preserving Aggregation of Time-Series Data," in *International Conference on Financial Cryptography and Data Security*, ser. Lecture Notes in Computer Science. Springer, pp. 111–125.

[14] P. D. Groves, "Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems," vol. 30, no. 2, pp. 26–27.

[15] C. Murguia, I. Shames, F. Farokhi, and D. Nešić, "Information-Theoretic Privacy Through Chaos Synchronization and Optimal Additive Noise," in *Privacy in Dynamical Systems*. Springer, pp. 103–129.

[16] M. Ristic, B. Noack, and U. D. Hanebeck, "Cryptographically Privileged State Estimation With Gaussian Keystreams," vol. 6, pp. 602–607.