

Supplementary Document for Identifying influential pandemic regions using graph signal variation

Sudeepini Darapu*, Subrata Ghosh*, Abhishek Senapati †, Chittaranjan Hens*,
Santosh Nannuru*

* *International Institute of Information Technology, Hyderabad, India*

† *Center for Advanced Systems Understanding, Görlitz, Germany*

This is the supplementary document for ICASSP'23 submission “Identifying influential pandemic regions using graph signal variation” [1]. In [1], due to space limitation we are only able to present partial simulation results. Here we include results for additional simulations and show raw (unprocessed) data behaviour, evolution of disease, and identification of influential pandemic regions using methods discussed in paper. Figure 1(a) shows the nodes generated using uniform random sampling and Figure 1(b) is the graph used for epidemic data generation.

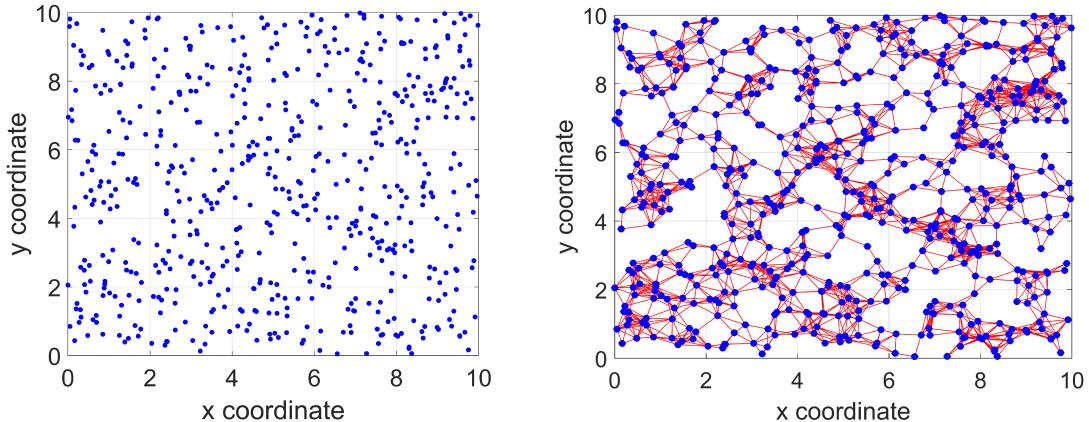


Figure 1: **a):** Nodes generated using uniform random sampling **b):** Graph with nodes connected using a distance based threshold.

1 Epidemic Data

In Section 4.2 of [1], we discussed two scenarios for generating synthetic graph signal data X based on the network model. The robustness of our approach was also tested using two different coupling strengths (low and high). For lower and higher κ , raw data for single and double perturbations at various time stamps are shown in Figures 2-5. Following are our observations:

- The evolution of disease at various time steps is shown in Fig. 2 and Fig. 4 for single and double perturbation using raw data. The infected source node has the highest infection severity at $t = 1$. Since κ is low, infection spread is slow across the network.
- From the data in Figures 3 and 5, for high value of κ , we see that the infection spread is faster and wider.

- add a point on single vs double perturbation data
- From the figures when κ is high, since most nodes have high infection, it is difficult to identify nodes playing an influential role in spreading the pandemic. Hence we develop methods in Section 3.2 of [1] to identify influential nodes in order to understand disease evolution.
- Our methods are capable of providing critical information on the spread of diseases that the unprocessed raw data could not.

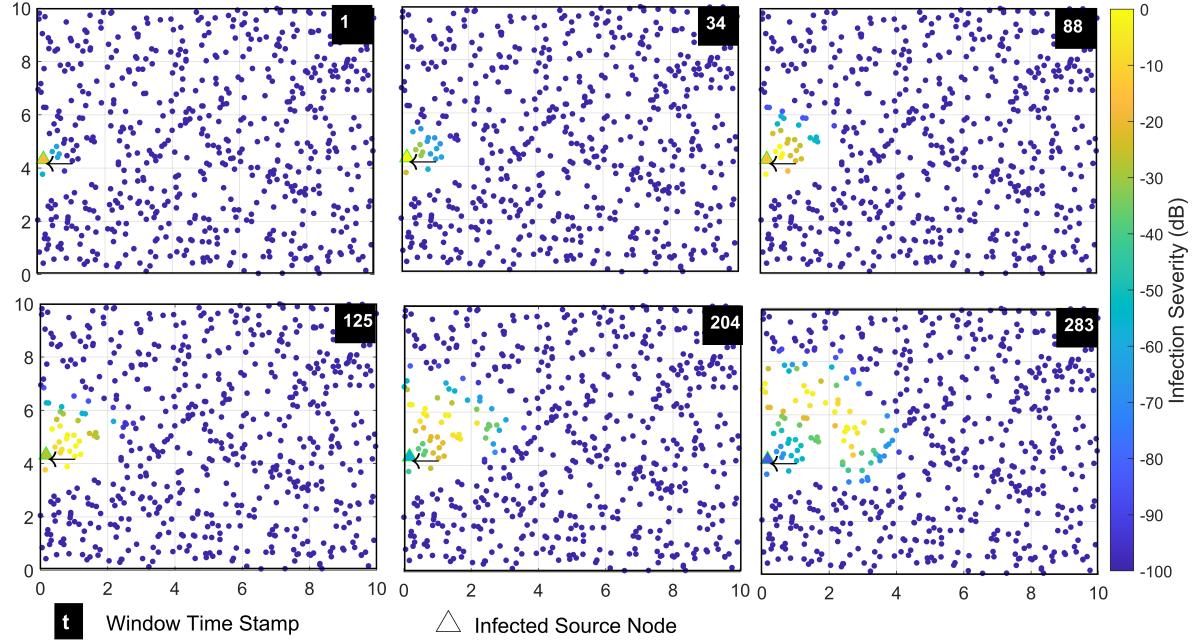


Figure 2: Raw data of single perturbation for $\kappa = 0.0001$

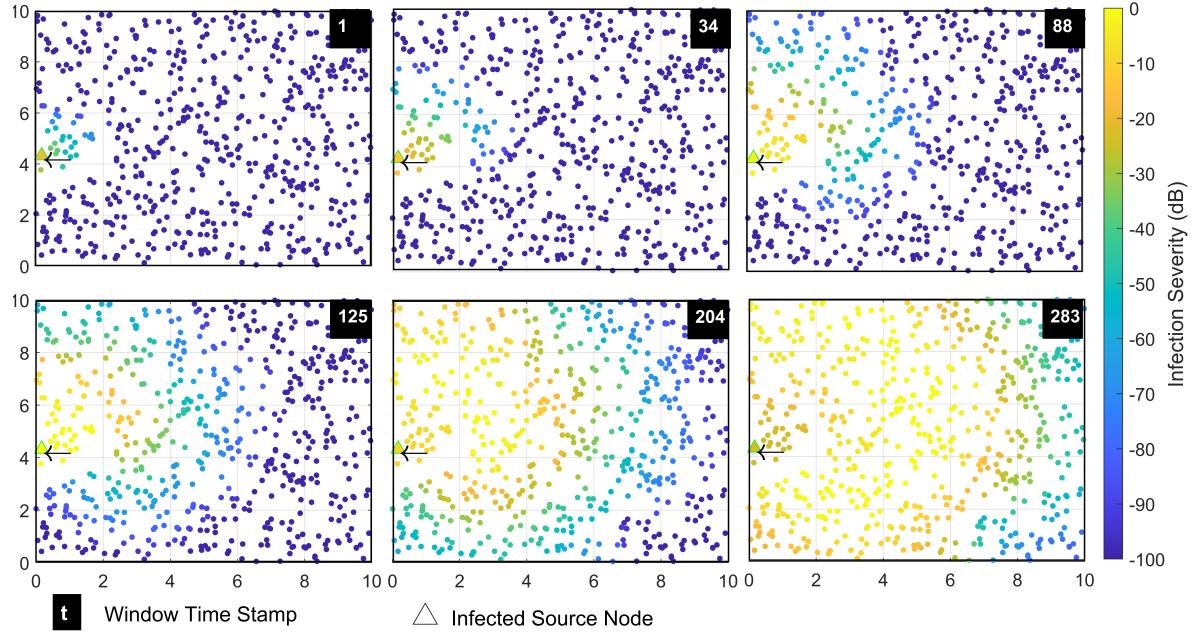


Figure 3: Raw data of single perturbation for $\kappa = 0.1$

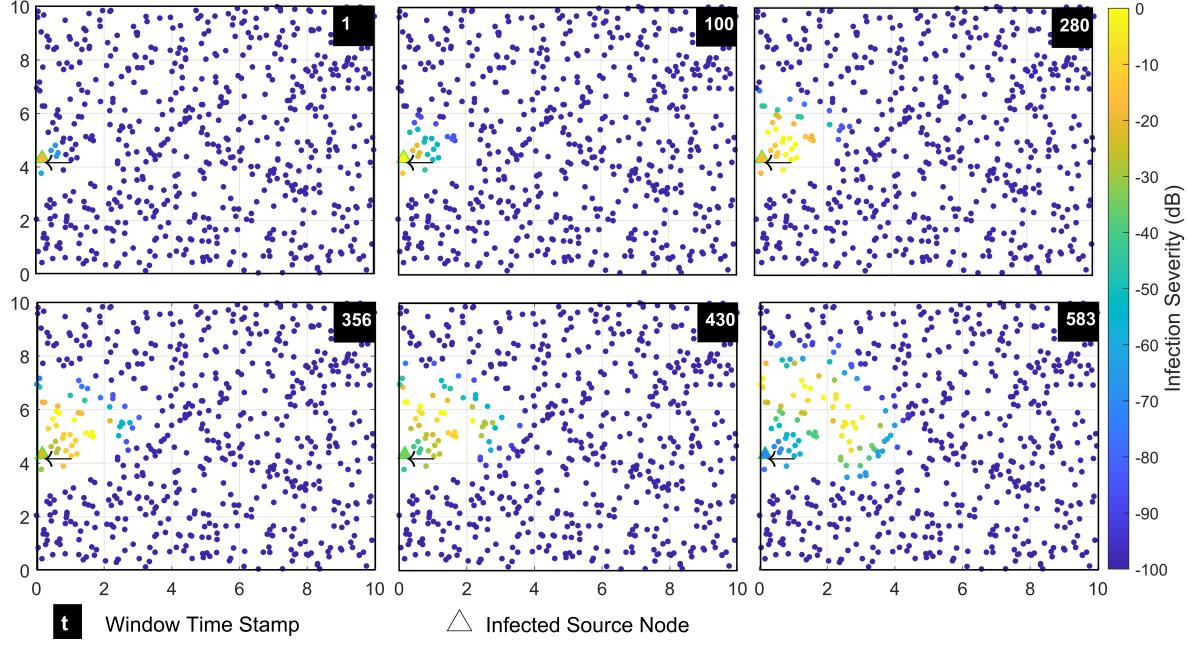


Figure 4: Raw data of double perturbation for $\kappa = 0.0001$

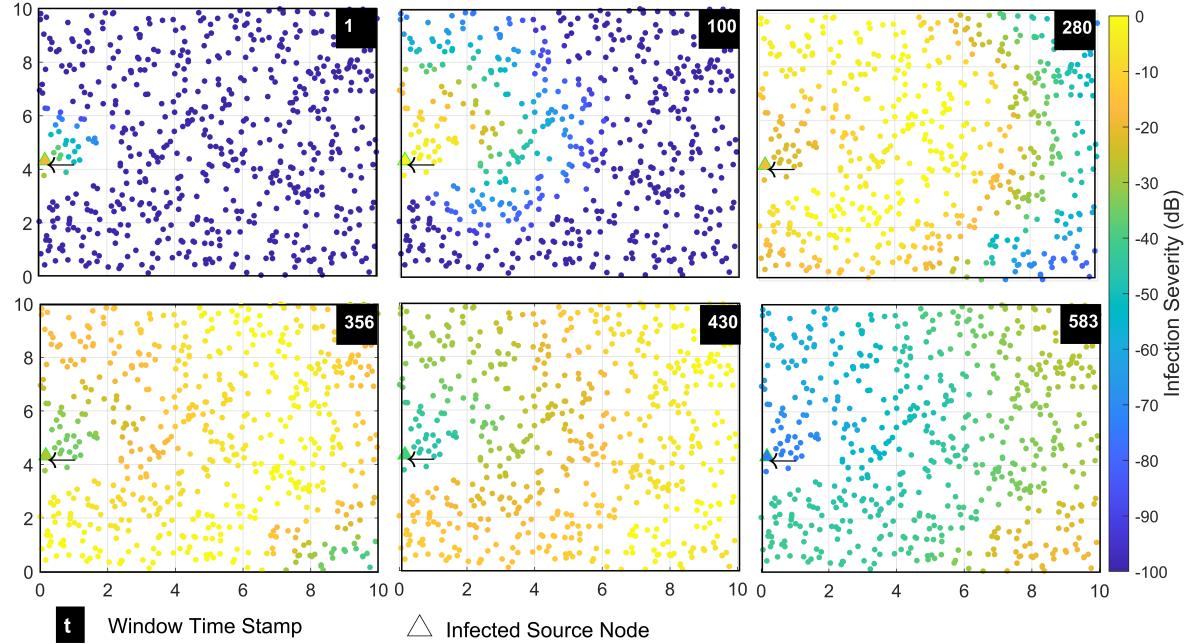


Figure 5: Raw data of double perturbation for $\kappa = 0.1$

2 Results

Figures 6-9 shows the evolution of disease from an infected source node with single and double perturbations for $\kappa = 0.0001, 0.1$ with raw data, HPF, LV and TLV processing.

- Figures 7, 9 show that all the identified influential nodes for higher κ using HPF, LV and TLV are a subset of the influential nodes that are identified by raw data processing across all time stamps.
- After some time stamps of higher κ , the infection at the source node decreases, yet raw data

processing incorrectly identifies source node as influential node. This can be seen in Figure 7 ($t = 170$) and Figure 9 ($t = 379$).

- After the disease spreads (at the end of time stamps), the source node and its neighbors should have less infection. Supporting this claim, the initially identified influential nodes using TLV show less infection severity at the end of the disease.
- In contrast, the initially identified influential nodes using raw data processing, including source nodes, always have higher infection severity even after the disease spread. This can be seen at $t = 288$ in Figure 7 and at $t = 583$ in Figure 9.
- From Figures 6-9, we can observe that TLV can capture a variety of graph signal variations as indicated by differently colored nodes. Disease spread is driven by the influential spreader nodes highlighted in red.

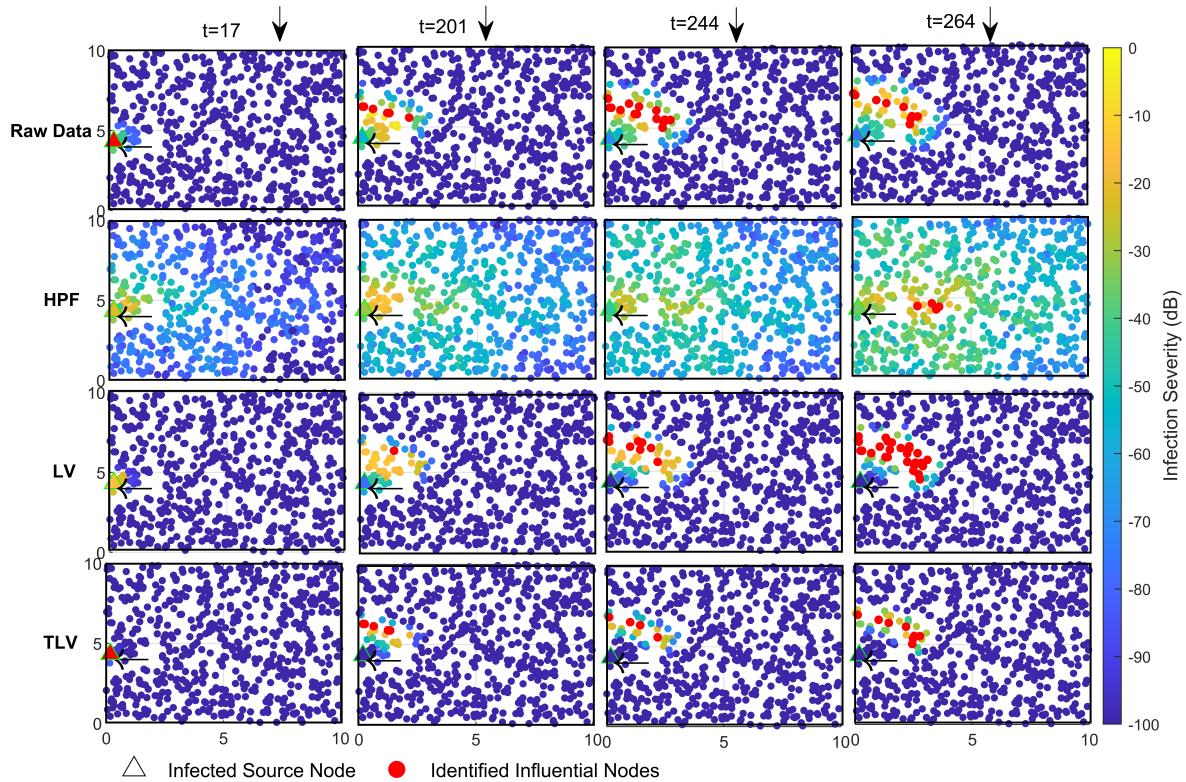


Figure 6: Case 1: Illustration of Raw data, HPF, LV, and TLV for single perturbation, $\kappa = 0.0001$

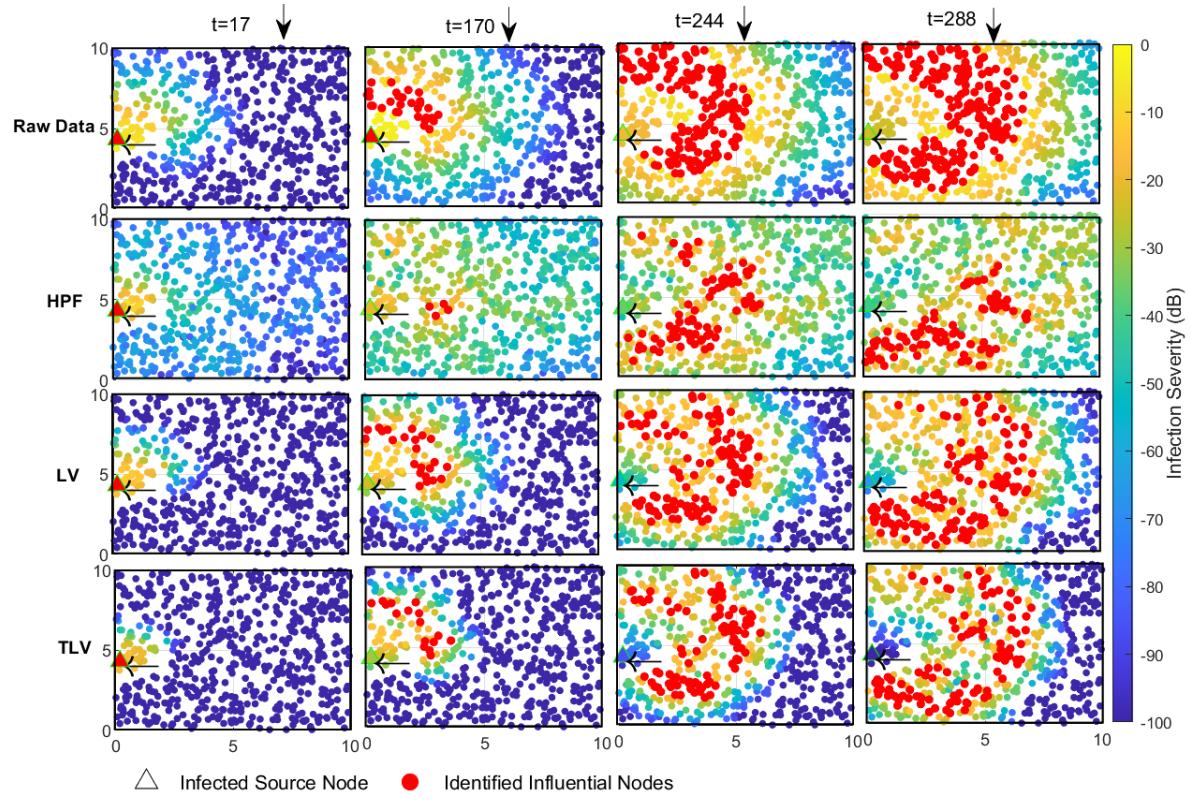


Figure 7: Case 2: Illustration of Raw data, HPF, LV, and TLV for single perturbation, $\kappa = 0.1$

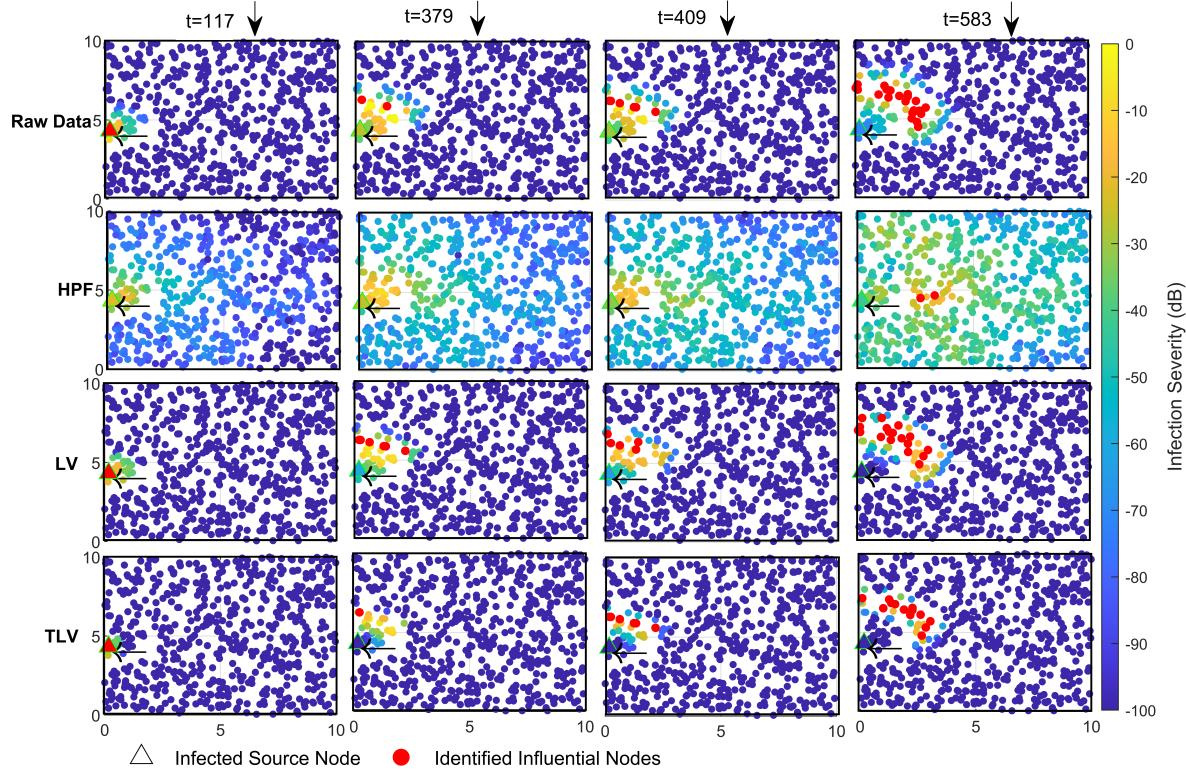


Figure 8: Case 3: Illustration of Raw data, HPF, LV, and TLV for double perturbation, $\kappa = 0.0001$

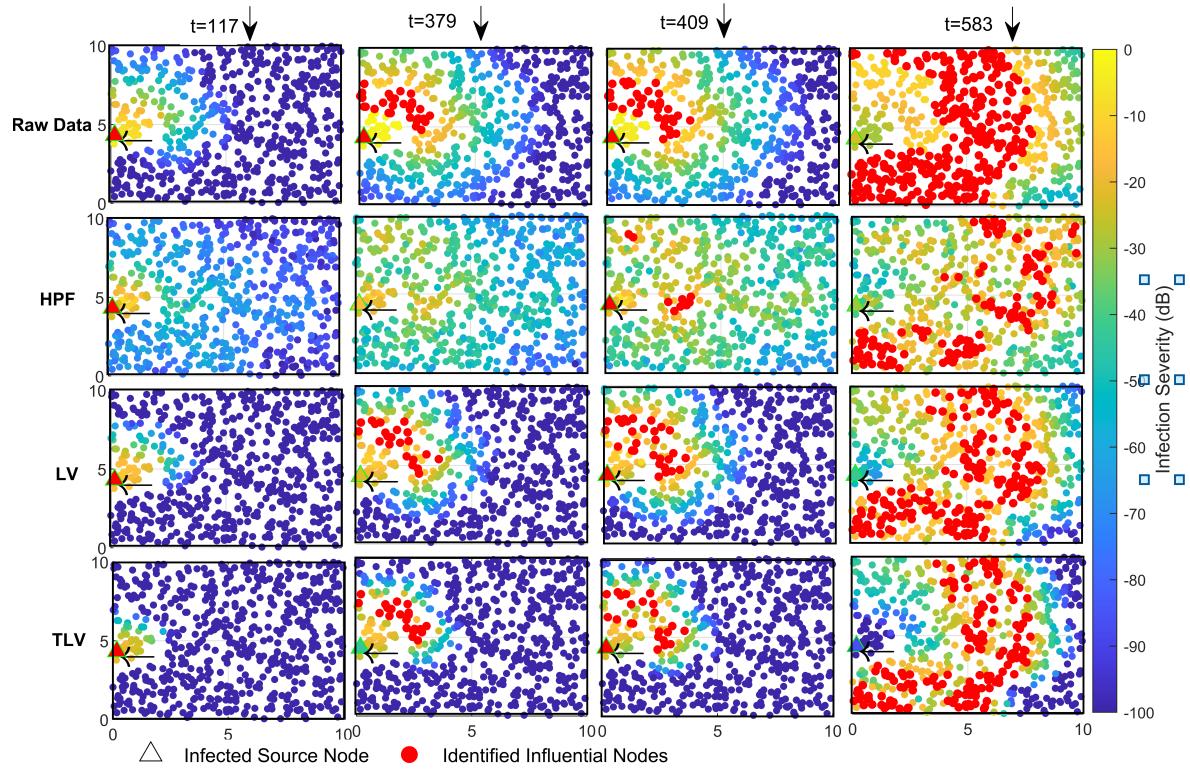


Figure 9: Case 4: Illustration of Raw data, HPF, LV, and TLV for double perturbation, $\kappa = 0.1$

References

- [1] S. Darapu, S. Ghosh, A. Senapti, C. Hens, and S. Nannuru, “Identifying influential pandemic regions using graph signal variation,” <https://doi.org/10.48550/arXiv.2211.05517>, 2022.