We appreciate the time and efforts all reviewers and ACs have devoted to the review of this paper. Your comments and suggestions help us a lot to improve the quality of this paper. Basically, all reviewers posted many shared concerns about the writing and vague problem formulation in our previous submission that make all reviewers hard to understand and evaluate. With the help of the suggestions and questions all reviewers provided, we have reorganized the writing of the paper and added more details to illustrate the IP networking backgrounds, working mechanism, the motivation of our problem and methods, as well as more experiments to justify the novelty and superiority of our proposal.

We realized that many of the questions the reviewers raised are related to the traffic modelling problems (e.g., transportation traffic flow, crowd flow) in transportation domain. As pointed out by Reviewer 2, the traffic flow we are talking about in this paper is the internet traffic flows (or more precisely, IP network traffic flows), which have totally different network structure and working mechanism. The novel identifications of the domain-specific data properties, structure and working mechanism make the design of our neural network structure & learning strategy stand out from the existing works. The traffic flow modelling problem in IP network, particularly at the time granularity of subsecond, is a crucial challenge but less exposed in existing deep learning community. Among the very few studies in the literature, most of the authors only explored this problem by harvesting the existing neural network models, e.g., GNN, RNN and their variants, as we reviewed in Section 1.2 of the revised version. We did the pioneering work (if not the first time) to design a customized neural network model by deeply combine the IP network structure and working mechanism. The experimental results also made well justifications of the superiority and robustness of our proposals on both synthetic and realist traffic patterns in multiple practical network applications. We believe this work will stimulate more innovations to both networking and deep learning research communities.

As suggested by Reviewer 2, we presented the revised paper to researchers outside of the project from both networking and deep learning communities to seek their feedbacks in the past two weeks. We believe these revised details will be informative to readers who are even totally new to the IP network background. In response to the major concerns of the reviewers, we make the following clarifications in supplementary to the provided revised paper version:

1. In INTRODUCTION (Section 1), we clearly defined the traffic flow modelling problem in the context of IP networks in Subsection 1.1, the flaws of existing models for this problem and the motivations of our design in Subsection 1.2, as well as the contributions we claim in Subsection 1.3. We hope the added details answers the questions the reviewers raised about the domain background, problem formulation as well as the novelty of our proposals.
2. In SPATIO-TEMPORAL INDUCTION (Section 2), we clearly analyzed the domain-specific data correlations created by the IP network structure and work mechanism and explained why the temporal evolutions of a flow can be induced through the history spatial patterns at its correlated locations. Such data properties are physically created by IP network structure and working mechanism, which is independent of any specific IP network applications.
3. In FLOW NEURAL NETWORK (Section 3), we clearly described the motivations of the structure design of FlowNN. Specifically, the "***Path-level Coding***" is used to extract the similar components for the behavioral patterns manifested by the path-wide time series. Based on the path-level pattern, the "***Node-level Induction***" extracts the individual patterns at different path nodes. This is performed by the "***Contrastive Induction Learning***". With the help of the proposed contrastive induction loss, FlowNN is able to get trained in a self-supervised way without the labels from practical applications. By fine-tuning a Readout layer, the pretrained

FlowNN model successfully generalized to multiple practical network applications and different dataset. This shows the strong model generality and robustness of our approach, as concerned by the reviewers.

4. In EXPERIMENTS (Section 4), we added more datasets and application tasks to test the effectiveness of our model and rewrote the descriptions of result analysis with a clearer structure. In terms of the chosen baselines, most of them are the widely used (e.g, SARIMA, GRU) in the literature or the very recent work (e.g, STHGCN) in networking domain or closely related to the data correlation (e.g., multiGRU). Particularly, we also tested the widely cited STGCN model for transportation domain with the help of its open source codes in https://github.com/VeritasYin/STGCN_IJCAI-18 . In all experiments, we reported very exciting results.
5. Finally, as we highlighted in CONCLUSION (Section 5), with the powerful intelligent traffic model provided by FlowNN, it's possible to stimulate more innovations to solve the diverse downstream networking applications across the network planning, control and management in IP networks. This will also be a new "playground" for deep learning communities to develop and test new models and algorithms.