Abstract:

Internet of Things (IoT) has gained tremendous popularity with the recent fast-paced technological advances in embedded programmable electronic and electro-mechanical systems, miniaturization, and their networking ability. IoT is expected to change the way of human activities by extensively networked monitoring, automation, and control. However, widespread application of IoT is associated with numerous challenges on communication and storage requirements, energy sustainability, and security. Also, IoT service quality requirements are application-specific. In this dissertation, we have identified novel methodologies to exploit the data-driven IoT framework for optimization of resources and development of context-aware cognitive applications in a massive machine type communication context. Through practical case studies, IoT application specific unique approaches and optimization techniques are proposed to reduce the data handling footprint, leading to communication bandwidth, cloud storage, and energy saving, without compromising service quality, thereby making them viable for wide-ranging adoption.

In the first part of dissertation, we introduce a novel data-driven framework for data pruning in wide area monitoring and control in smart grid, which is an emerging IoT application. Due to stringent latency constraints, packet losses and end-to-end transmission delay exceeding the permitted threshold values may jeopardize the stability of power grid. However, transient occurrences in the grid are relatively sparse, and much of the data is routine monitoring data having high redundancy. The proposed framework exploits the temporal correlatedness in the consecutive samples to dynamically prevent redundant PMU data from being transmitted without affecting the quality of power grid health monitoring. The missing samples are predicted at the receiving end using $\epsilon$-support vector regression model. It is noted that though PMU data has high temporal correlation, it actually characterizes a non-stationary process. Consequently, hyper-parameters of the prediction model are recomputed as necessary to maintain the accuracy and robustness of prediction. Also, for low runtime complexity, some of the hyper-parameters are precomputed using empirical optimization of offline data characteristics. Appropriate performance indices are defined to quantify the performance of the proposed algorithm, and its computational latency is estimated via online execution using Simulink model. It is found that the proposed dynamic prediction algorithm selectively transmits the PMU data, thereby achieving up to 90% reduction in channel bandwidth requirement without affecting the quality of stability monitoring of the system. Further, comparison of the proposed algorithm with the closest competitive scheme demonstrates 73% and 60% better performance, respectively, in terms of power system health monitoring and bandwidth saving.

Subsequently, in the next part intelligent data pruning is investigated for automated electric metering in smart cities, which, similar to PMU data, is expected to increase the volume of network traffic exponentially. However, it does not possess the same nature of dynamics as the PMU data, and is more relaxed in terms of delay tolerance. It may be noted that as the granularity of sampling average power consumption in smart meter increases, compressibility of the data reduces, owing to irregular load profile. Besides, the data appears incoherent in time domain, though it can actually be represented by a sparsifying basis. Thus, adaptively choosing the sparsity over optimum batch size before data transmission can be utilized for substantial reduction in data volume. To this end, considering high resolution data at the smart meter, the problem of smart meter data characterization and reduction is addressed to achieve higher compression gains and reduced bandwidth requirement for data transmission from smart meter to the data collector. A novel Gaussian mixture based model is proposed for the characterization of high frequency
smart meter data, which is used in evaluating the quality of data reduction at the smart meter. Further, an adaptive data reduction scheme using compressive sampling is devised to operate at the smart meter which achieves about 40% bandwidth saving in data transmission to the nearest collection center without any appreciable loss of information. Performance comparison of the proposed data reduction scheme with an existing competitive approach demonstrates noise robustness during data transmission. Additionally, to achieve the same order of RMSE, bandwidth saving with the proposed scheme is 12:8% and 7:4% higher, respectively, for data sampled at 1 second and 30 seconds. Real-time implementation of the proposed system level design is tested on smart meters deployed at IIT Delhi campus.

Finally, in the last part, channel-adaptive transmission strategies based on simple yet efficient channel prediction frameworks using stochastic modeling and data-driven learning of channel variability are proposed for sporadic but time-critical PMU data. The proposed channel prediction frameworks are accompanied with adaptive channel coding that assigns redundant symbols to the packet in accordance with the current channel state. A probing-based transmission scheme is also proposed which is considered as the benchmark for comparing the stochastic model-based and learning-based approaches. Through large-scale simulations, the prediction and packet loss performance is analyzed at varying SNR and fading conditions. The results demonstrate that, for a given channel fading condition, packet loss probability of the proposed learning-based transmission closely matches with the benchmark scheme, while with the stochastic model-based prediction the loss probability is found to be 12:3% higher. However, the respective signaling overhead requirements are 38% and 98% lower with respect to the benchmark.