1.7 統計的手法の物理解析への応用とシミュレーション

柴田 章博

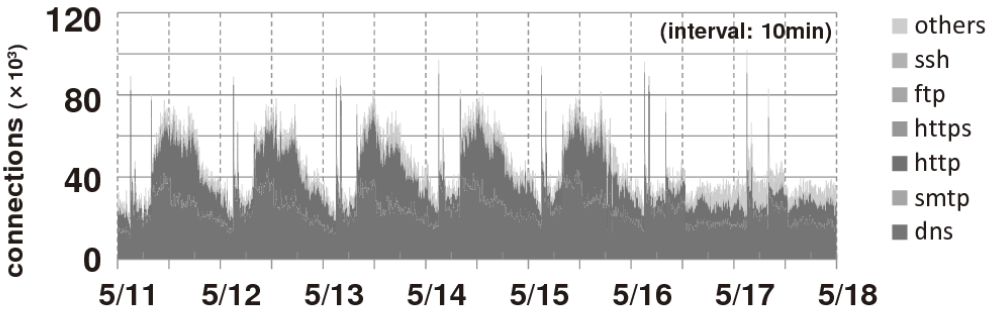
概要

ベイズ統計に基づく数値統計やデータマイニングの技術の手法的研究とともに、主として数値シミュレーションの解析や高エネルギー物理や構造解析など加速器実験におけるデータ解析で必要とされるデータ解析法について具体的テーマを取りあげてその解析法を開発・改善を行っている。

系列データの解析

DFA(Detrended Fluctuation Analysis) による時系列データ解析を拡張し、共分散解析による因果律を分析法を定式化した。前年度に引き続き時系列データの事例として、インターネット通信のデータを用いて、データ間に潜む因果律や相関の検出を検証している。インターネット通信のトラフィックは、自律的な成長と調整の機能を持った集中制御の仕組みを持たない自己駆動系の流動であり、物理的な解析対象としても興味深い。

KEK のファイアウォール(FW)の通信のログを利用して、ＬＡＮとＷＡＮの間の通信に潜む関係を検出できるかについて、定点のルーターを通過するパケットの観測のみから検出できるかについて検討を進めた。図は 2009 年の KEK LAN からインターネットへの TCP/IP 通信のプロトコル別の接続数を示している。



共分散解析の結果、プロトコル別の接続数のみのデータから、SMTP,Web の通信と DNS によるアドレス解決の通信を揺らぎ相関として検出できることを確認した。通信プロトコルや通信の向きなどさまざまな組み合わせにおいて検証を進めた。現在、その成果を論文にまとめている。