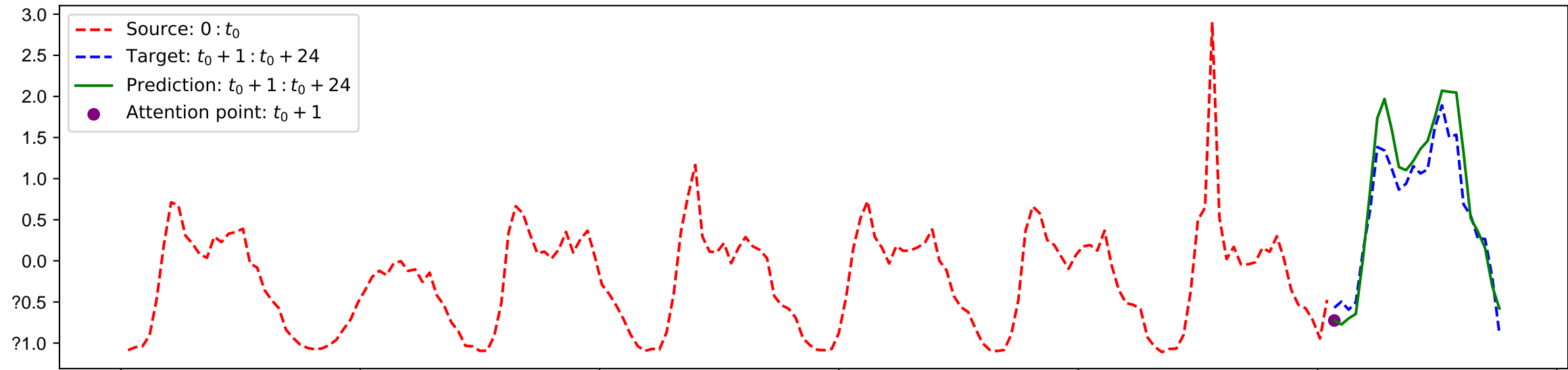Traffic dataset, SE_kernel, 24-steps non-overlapping, 12-th time series, total epochs: 300

Attention pattens for $t_0 + 1$