pc | Pervasive Computing Lab
浙江大学普适计算实验室

Ada-MSHyper: Adaptive Multi-Scale Hypergraph Transformer for Time Series Forecasting

Zongjiang Shang, Ling Chen*, Binqing Wu, Dongliang Cui
College of Computer Science and Technology, Zhejiang University, China

Presenter: Zongjiang Shang

Motivations

Two key challenges limit the performance of Transformer in multi-scale temporal pattern interaction modeling.

- **Semantic information sparsity.** Individual time points contain less semantic information, and **pair-wise interactions** may cause the information utilization bottleneck.

- **Temporal variations entanglement.** Multiple inherent temporal variations (e.g., rising, falling, and fluctuating) entangled in temporal pattern, bringing challenges for time series forecasting.

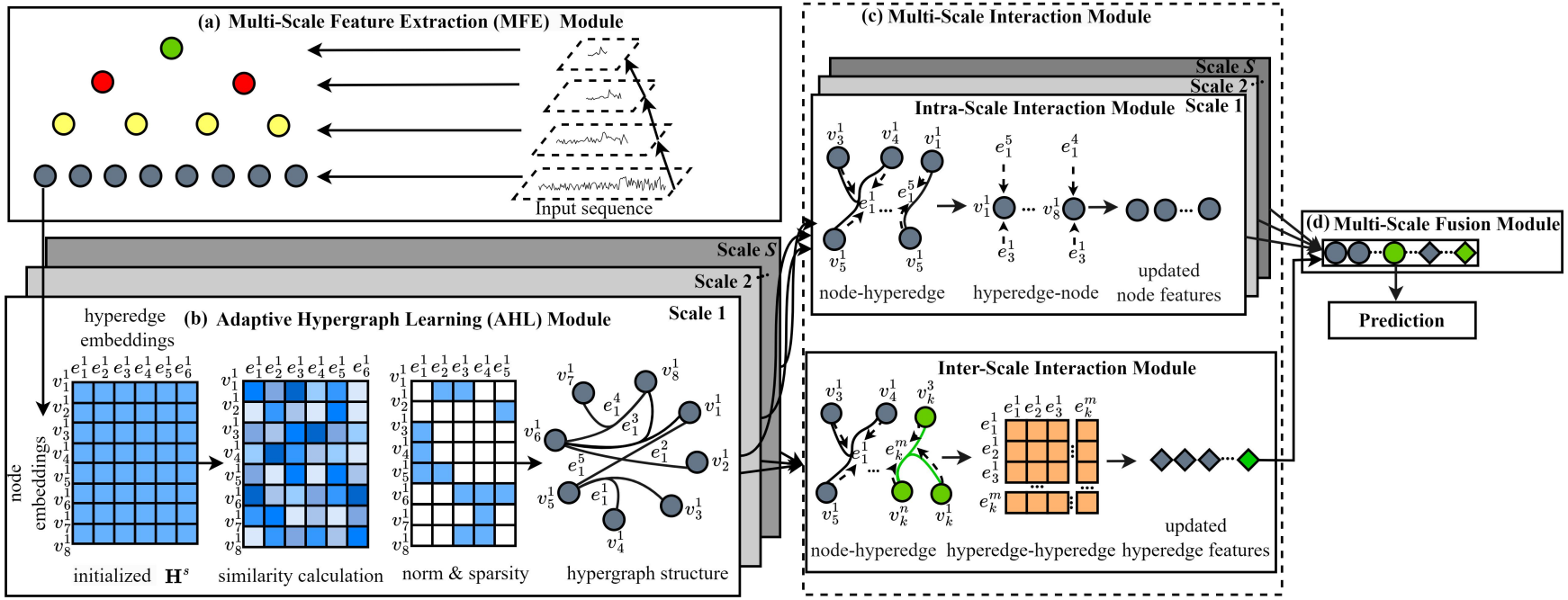
Contributions

Ada-MSHyper is **the first work** that incorporates adaptive hypergraph modeling into time series forecasting.

- An adaptive hypergraph learning module is designed to model abundant and implicit **group-wise node interactions** at different scales.
- A node and hyperedge constraint mechanism is introduced to cluster nodes with **similar semantic information** and **differentiate the temporal variations** within each scales.
- Experimental results on 11 real-world datasets demonstrate that Ada-MSHyper achieves **state-of-the-art** performance.

Method: Ada-MSHyper

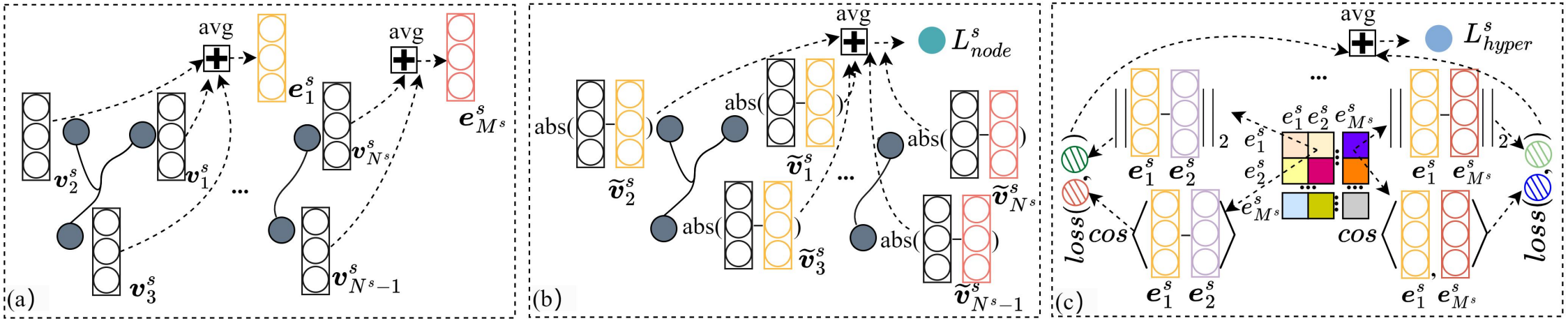
Framework



(a) The MFE module maps the input sequence into subsequences at different scales. (b) The AHL module provides foundations for modeling group-wise interactions. (c) The multi-scale interaction module models group-wise pattern interactions at different scales.

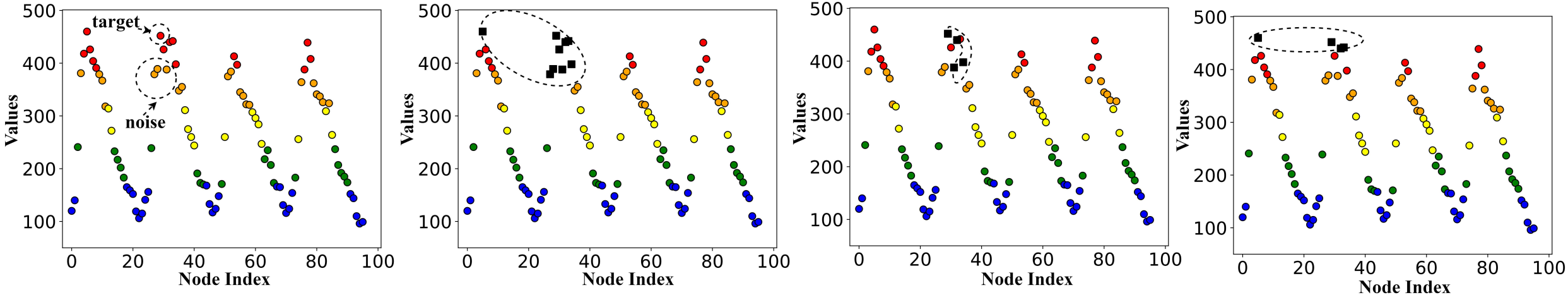
Method: Ada-MSHyper

Node and hyperedge constraint (NHC) mechanism



(a) The generation of hyperedge features. (b) The generation of node loss. (c) The generation of hyperedge loss.

Visualization



(a) Input sequence (b) -w/o NHC (c) -w/o NC (d) Ada-MSHyper

With the NHC mechanism, Ada-MSHper can not only cluster nodes with similar semantic information but also reduce noise interference.

Evaluations

The results of long-range time series forecasting under multivariate settings.

Models		Ada-MSHyper (Ours)		iTransformer (2024)		MSHyper* (2024)		TimeMixer* (2024)		MSGNet* (2024)		CrossGNN* (2023)		PatchTST (2023)		Crossformer (2023)		TimesNet (2023)		DLinear (2023)		FiLM* (2022)		FEDformer (2022)		Autoformer (2021)	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Weather	96	**0.157**	**0.195**	0.174	0.214	0.170	0.223	0.163	0.210	0.163	0.212	0.159	0.218	0.177	0.218	0.158	0.230	0.172	0.220	0.196	0.255	0.199	0.262	0.217	0.296	0.266	0.336
	192	0.218	0.259	0.221	0.254	0.218	**0.253**	0.212	0.257	0.212	0.254	0.211	0.266	0.225	0.259	**0.206**	0.277	0.219	0.261	0.237	0.296	0.228	0.288	0.276	0.336	0.307	0.367
	336	**0.251**	**0.252**	0.278	0.296	0.269	0.300	0.263	0.292	0.272	0.299	0.267	0.310	0.278	0.297	0.272	0.335	0.280	0.306	0.283	0.335	0.267	0.323	0.339	0.380	0.359	0.395
	720	**0.304**	**0.328**	0.358	0.347	0.343	0.341	0.343	0.345	0.350	0.348	0.352	0.362	0.354	0.348	0.398	0.418	0.365	0.359	0.345	0.381	0.319	0.361	0.403	0.428	0.419	0.428
Electricity	96	**0.135**	**0.238**	0.148	0.240	0.176	0.261	0.153	0.247	0.165	0.274	0.173	0.275	0.181	0.270	0.219	0.314	0.168	0.272	0.197	0.282	0.198	0.274	0.193	0.308	0.201	0.317
	192	**0.152**	**0.239**	0.162	0.253	0.173	0.260	0.166	0.256	0.184	0.292	0.195	0.288	0.188	0.274	0.231	0.322	0.184	0.289	0.196	0.285	0.198	0.278	0.201	0.315	0.222	0.334
	336	**0.168**	**0.266**	0.178	0.269	0.195	0.297	0.185	0.277	0.195	0.302	0.206	0.300	0.204	0.293	0.246	0.337	0.198	0.300	0.209	0.301	0.217	0.300	0.214	0.329	0.231	0.338
	720	**0.212**	**0.293**	0.225	0.317	0.219	0.315	0.225	0.310	0.231	0.332	0.231	0.335	0.246	0.324	0.280	0.363	0.220	0.320	0.245	0.333	0.278	0.356	0.246	0.355	0.254	0.361
ETTh1	96	**0.372**	**0.393**	0.386	0.405	0.392	0.407	0.385	0.402	0.390	0.411	0.382	0.398	0.414	0.419	0.423	0.448	0.384	0.402	0.386	0.400	0.438	0.433	0.376	0.419	0.449	0.459
	192	0.433	**0.417**	0.441	0.436	0.440	0.426	0.443	0.430	0.442	0.442	0.427	0.425	0.460	0.445	0.471	0.474	0.436	0.429	0.437	0.432	0.493	0.466	**0.420**	0.448	0.500	0.482
	336	**0.422**	**0.433**	0.487	0.458	0.480	0.453	0.512	0.470	0.480	0.468	0.465	0.445	0.501	0.466	0.570	0.546	0.491	0.469	0.481	0.459	0.547	0.495	0.459	0.465	0.521	0.496
	720	**0.445**	**0.459**	0.503	0.491	0.508	0.493	0.498	0.476	0.494	0.488	0.472	0.468	0.500	0.488	0.653	0.621	0.521	0.500	0.519	0.516	0.586	0.538	0.506	0.507	0.514	0.512
ETTh2	96	**0.283**	**0.332**	0.297	0.349	0.300	0.351	0.296	0.347	0.328	0.371	0.309	0.359	0.302	0.348	0.745	0.584	0.340	0.374	0.333	0.387	0.322	0.364	0.358	0.397	0.346	0.388
	192	**0.358**	**0.374**	0.380	0.400	0.384	0.400	0.376	0.394	0.402	0.414	0.390	0.406	0.388	0.400	0.877	0.656	0.402	0.414	0.477	0.476	0.404	0.414	0.429	0.439	0.456	0.452
	336	0.428	0.437	0.428	**0.432**	0.443	0.438	0.434	0.443	0.435	0.443	**0.426**	0.444	**0.426**	0.433	1.043	0.731	0.452	0.452	0.594	0.541	0.435	0.445	0.496	0.487	0.482	0.486
	720	0.413	**0.432**	0.427	0.445	**0.412**	0.441	0.464	0.464	0.417	0.441	0.445	0.464	0.431	0.446	1.104	0.763	0.462	0.468	0.831	0.657	0.447	0.458	0.463	0.474	0.515	0.511
ETTm1	96	**0.301**	**0.354**	0.334	0.368	0.348	0.369	0.318	0.356	0.319	0.366	0.335	0.373	0.329	0.367	0.404	0.426	0.338	0.375	0.345	0.372	0.353	0.370	0.379	0.419	0.505	0.475
	192	**0.345**	**0.375**	0.377	0.391	0.392	0.391	0.366	0.385	0.376	0.397	0.372	0.390	0.367	0.385	0.450	0.451	0.374	0.387	0.380	0.389	0.389	0.387	0.426	0.441	0.553	0.496
	336	**0.375**	**0.397**	0.426	0.420	0.426	0.410	0.396	0.404	0.417	0.422	0.403	0.411	0.399	0.410	0.532	0.515	0.410	0.411	0.413	0.413	0.421	0.408	0.445	0.459	0.621	0.537
	720	**0.437**	**0.435**	0.491	0.459	0.483	0.448	0.454	0.441	0.481	0.458	0.461	0.442	0.454	0.439	0.666	0.589	0.478	0.450	0.474	0.453	0.481	0.441	0.543	0.490	0.671	0.561
ETTm2	96	**0.165**	**0.257**	0.180	0.264	0.183	0.267	0.175	0.258	0.177	0.262	0.176	0.266	0.175	0.259	0.287	0.366	0.187	0.267	0.193	0.292	0.183	0.266	0.203	0.287	0.255	0.339
	192	**0.230**	0.307	0.250	0.309	0.257	0.313	0.241	0.304	0.247	0.307	0.240	0.307	0.241	**0.302**	0.414	0.492	0.249	0.309	0.284	0.362	0.248	0.305	0.269	0.328	0.281	0.340
	336	**0.282**	**0.328**	0.311	0.348	0.335	0.361	0.303	0.343	0.312	0.346	0.304	0.345	0.305	0.343	0.597	0.542	0.321	0.351	0.369	0.427	0.309	0.343	0.325	0.366	0.339	0.372
	720	**0.375**	0.396	0.412	0.407	0.410	0.402	0.391	**0.394**	0.414	0.403	0.406	0.400	0.402	0.400	1.730	1.042	0.408	0.403	0.554	0.522	0.410	0.400	0.421	0.415	0.433	0.432
Traffic	96	**0.384**	**0.248**	0.395	0.268	0.413	0.272	0.473	0.288	0.605	0.344	0.570	0.310	0.462	0.295	0.522	0.290	0.593	0.321	0.650	0.396	0.647	0.384	0.587	0.366	0.613	0.388
	192	**0.401**	**0.258**	0.417	0.276	0.422	0.274	0.473	0.296	0.613	0.359	0.577	0.321	0.466	0.296	0.530	0.293	0.617	0.336	0.598	0.370	0.600	0.361	0.604	0.373	0.616	0.382
	336	**0.423**	**0.261**	0.433	0.283	0.438	0.292	0.508	0.312	0.642	0.376	0.588	0.324	0.482	0.304	0.558	0.305	0.629	0.336	0.605	0.373	0.610	0.367	0.621	0.383	0.622	0.337
	720	**0.453**	**0.282**	0.467	0.302	0.457	0.292	0.512	0.318	0.702	0.401	0.597	0.337	0.514	0.322	0.589	0.328	0.640	0.350	0.645	0.394	0.691	0.425	0.626	0.382	0.660	0.408

The results of short-range time series forecasting under multivariate settings.

Models		Ada-MSHyper (Ours)		iTransformer* (2024)		MSHyper* (2024)		TimeMixer* (2024)		WITRAN* (2023)		PatchTST* (2023)		Dlinear* (2023)		Crossformer* (2023)		FEDformer* (2022)		Pyraformer* (2022)		Autoformer* (2021)	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	1080	**0.534**	**0.509**	0.562	0.521	0.557	0.517	0.682	0.569	0.602	0.660	0.549	0.512	0.593	0.564	0.877	1.204	0.699	0.615	1.015	0.798	0.695	0.626
	1440	**0.616**	**0.498**	0.620	0.556	0.667	0.578	0.793	0.625	0.705	0.878	0.619	0.553	0.661	0.607	0.863	1.175	0.621	0.567	1.075	0.833	0.876	0.696
	1800	**0.689**	0.627	0.780	0.631	0.758	0.624	0.877	0.643	0.775	**0.623**	0.775	**0.623**	0.746	0.658	0.849	1.163	0.806	0.649	1.111	0.844	0.852	0.704
	2160	**0.779**	**0.635**	1.102	0.736	0.998	0.721	1.007	0.686	0.852	1.171	0.851	0.665	0.783	0.667	1.095	0.821	0.935	0.717	1.129	0.847	–	–
ETTh2	1080	**0.426**	**0.461**	0.486	0.488	0.464	0.469	0.483	0.480	0.432	0.474	0.453	0.468	0.730	0.617	1.481	0.918	0.514	0.526	3.224	1.458	0.559	0.547
	1440	**0.465**	**0.437**	0.512	0.507	0.524	0.506	0.547	0.510	0.472	0.443	0.513	0.501	1.144	0.770	1.901	1.044	0.578	0.546	3.254	1.548	0.638	0.708
	1800	**0.503**	0.505	0.565	0.529	0.522	**0.496**	0.606	0.544	0.626	0.610	0.517	0.503	1.327	0.840	3.109	1.486	0.645	0.584	3.328	1.565	0.776	0.689
	2160	**0.527**	0.515	0.600	0.546	0.542	**0.510**	0.616	0.557	0.657	0.619	0.547	0.519	1.670	0.919	3.630	1.485	0.762	0.639	3.246	1.465	–	–
ETTm1	1080	**0.460**	**0.445**	0.534	0.483	0.520	0.465	0.502	0.465	0.464	0.459	0.494	0.459	0.514	0.479	2.588	1.236	0.513	0.499	1.071	0.793	0.651	0.551
	1440	**0.473**	**0.449**	0.556	0.495	0.542	0.477	0.523	0.488	0.543	0.467	0.508	0.467	0.534	0.491	2.946	1.349	0.511	0.494	1.136	0.834	0.602	0.542
	1800	**0.492**	0.475	0.571	0.501	0.564	0.490	0.526	0.487	0.550	0.497	0.504	**0.434**	0.556	0.507	4.113	1.602	0.514	0.496	1.111	0.812	0.641	0.558
	2160	0.510	0.483	0.555	0.499	0.550	0.487	0.542	0.491	0.569	**0.481**	**0.507**	**0.481**	0.556	0.515	4.574	1.743	0.551	0.516	1.054	0.804	–	–
ETTm2	1080	**0.404**	**0.416**	0.463	0.438	0.464	0.439	0.450	0.432	0.415	0.434	0.449	0.432	0.559	0.519	2.587	1.236	0.501	0.468	4.879	1.733	0.527	0.489
	1440	**0.413**	**0.429**	0.475	0.452	0.475	0.449	0.471	0.452	0.442	0.442	0.475	0.452	0.699	0.593	2.946	1.349	0.495	0.480	4.429	1.708	0.519	0.489
	1800	**0.435**	0.432	0.468	0.453	0.454	0.449	0.464	0.452	0.479	**0.410**	0.456	0.449	0.721	0.612	4.113	1.602	0.477	0.474	4.502	1.780	0.503	0.496
	2160	0.449	0.457	0.467	0.454	0.463	0.451	0.473	0.459	**0.447**	**0.449**	0.466	0.457	0.639	0.572	4.574	1.743	0.473	0.477	4.454	1.758	–	–

The results of long-range time series forecasting under univariate settings.

Models		Ada-MSHyper (Ours)		iTransfomer* (2024)		MSHyper* (2024)		TimeMixer* (2024)		PatchTST* (2023)		DLinear (2023)		Crossformer (2023)		Pyraformer (2022)		FEDformer (2022)		Autoformer (2021)		Informer (2021)	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	**0.057**	**0.173**	0.059	0.185	**0.056**	0.181	0.057	0.181	**0.056**	0.181	**0.056**	0.180	0.076	0.216	0.099	0.277	0.079	0.215	0.071	0.206	0.193	0.377
	192	**0.072**	**0.198**	0.073	0.208	0.076	0.211	0.072	0.204	0.076	0.210	**0.071**	0.204	0.085	0.225	0.174	0.346	0.104	0.245	0.114	0.262	0.217	0.395
	336	**0.070**	**0.213**	0.084	0.223	0.090	0.236	0.085	0.227	0.094	0.242	0.098	0.244	0.106	0.257	0.198	0.370	0.119	0.270	0.107	0.258	0.202	0.381
	720	**0.085**	0.228	0.089	0.236	0.096	0.245	**0.083**	**0.227**	0.101	0.250	0.189	0.359	0.128	0.287	0.209	0.348	0.142	0.299	0.126	0.283	0.183	0.355
ETTh2	96	**0.116**	**0.262**	0.136	0.287	0.117	0.266	0.133	0.283	0.130	0.276	0.131	0.279	0.125	0.273	0.152	0.303	0.128	0.271	0.153	0.306	0.213	0.373
	192	**0.168**	**0.323**	0.187	0.342	0.172	0.325	0.190	0.341	0.181	0.331	0.176	0.329	0.187	0.334	0.197	0.370	0.185	0.330	0.204	0.351	0.227	0.387
	336	**0.177**	**0.350**	0.219	0.374	0.211	0.362	0.226	0.379	0.226	0.379	0.209	0.367	0.227	0.377	0.238	0.385	0.231	0.378	0.246	0.389	0.242	0.401
	720	**0.221**	**0.380**	0.253	0.403	0.248	0.398	0.241	0.396	0.253	0.406	0.276	0.426	0.266	0.410	0.274	0.435	0.278	0.420	0.268	0.409	0.291	0.439
ETTm1	96	**0.027**	**0.118**	0.029	0.127	0.029	0.127	0.029	0.128	0.029	0.126	0.028	0.123	0.035	0.145	0.127	0.281	0.033	0.140	0.056	0.183	0.109	0.277
	192	**0.038**	**0.148**	0.045	0.162	0.044	0.159	0.044	0.160	0.043	0.158	0.045	0.156	0.055	0.180	0.205	0.343	0.058	0.186	0.081	0.216	0.151	0.310
	336	**0.052**	**0.165**	0.059	0.189	0.059	0.186	0.058	0.185	0.056	0.183	0.061	0.182	0.072	0.209	0.302	0.457	0.084	0.231	0.076	0.218	0.427	0.591
	720	**0.071**	**0.206**	0.080	0.218	0.080	0.217	0.081	0.218	0.080	0.217	0.080	0.210	0.097	0.248	0.387	0.485	0.102	0.250	0.110	0.267	0.438	0.586
ETTm2	96	**0.051**	**0.163**	0.071	0.193	0.071	0.194	0.068	0.187	0.071	0.192	0.063	0.183	0.058	0.183	0.074	0.208	0.067	0.198	0.065	0.189	0.088	0.225
	192	**0.089**	**0.207**	0.109	0.248	0.102	0.238	0.101	0.236	0.102	0.237	0.092	0.227	0.105	0.237	0.116	0.252	0.102	0.245	0.118	0.256	0.132	0.283
	336	**0.114**	**0.240**	0.141	0.289	0.129	0.274	0.133	0.278	0.130	0.274	0.119	0.261	0.133	0.280	0.143	0.295	0.130	0.279	0.154	0.305	0.180	0.336
	720	**0.156**	**0.310**	0.190	0.343	0.176	0.324	0.183	0.332	0.179	0.328	0.175	0.320	0.181	0.324	0.197	0.338	0.178	0.325	0.182	0.335	0.300	0.435

The results of ultra-long-range time series forecasting under multivariate settings.

Models		Ada-MSHyper (Ours)		iTransformer* (2024)		MSHyper* (2024)		TimeMixer* (2024)		PatchTST (2023)		TimesNet (2023)		DLinear (2023)		Crossformer (2023)		SCINet (2022)		FEDformer (2022)		Autoformer (2021)	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
PEMS03	12	**0.060**	**0.165**	0.071	0.174	0.106	0.207	0.161	0.323	0.099	0.216	0.085	0.192	0.122	0.243	0.090	0.203	0.066	0.172	0.126	0.251	0.272	0.385
	24	**0.075**	**0.184**	0.093	0.201	0.126	0.207	0.181	0.352	0.142	0.259	0.118	0.223	0.201	0.317	0.121	0.240	0.085	0.198	0.149	0.275	0.334	0.440
	48	**0.120**	**0.230**	0.125	0.236	0.138	0.265	0.222	0.407	0.211	0.319	0.155	0.260	0.333	0.425	0.202	0.317	0.127	0.238	0.227	0.348	1.032	0.782
PEMS04	12	**0.068**	**0.173**	0.078	0.183	0.103	0.197	0.168	0.344	0.105	0.224	0.087	0.195	0.148	0.272	0.098	0.218	0.073	0.177	0.138	0.262	0.424	0.491
	24	**0.080**	**0.189**	0.095	0.205	0.148	0.245	0.183	0.362	0.153	0.275	0.103	0.215	0.224	0.340	0.131	0.256	0.084	0.193	0.177	0.293	0.459	0.509
	48	**0.093**	**0.204**	0.120	0.233	0.191	0.308	0.199	0.383	0.229	0.339	0.136	0.250	0.355	0.437	0.205	0.326	0.099	0.211	0.270	0.368	0.646	0.610
PEMS07	12	**0.055**	**0.154**	0.067	0.165	0.137	0.256	0.151	0.322	0.095	0.207	0.082	0.181	0.115	0.242	0.094	0.200	0.068	0.171	0.109	0.225	0.199	0.336
	24	**0.065**	**0.172**	0.088	0.190	0.111	0.225	0.169	0.348	0.150	0.262	0.101	0.204	0.210	0.329	0.139	0.247	0.119	0.225	0.125	0.244	0.323	0.420
	48	**0.107**	**0.204**	0.110	0.215	0.137	0.221	0.196	0.384	0.253	0.340	0.134	0.238	0.398	0.458	0.311	0.369	0.149	0.237	0.165	0.288	0.390	0.470
PEMS08	12	**0.063**	**0.165**	0.079	0.182	0.113	0.209	0.162	0.337	0.168	0.232	0.112	0.212	0.154	0.276	0.165	0.214	0.087	0.184	0.173	0.273	0.436	0.485
	24	**0.109**	0.229	0.115	**0.219**	0.230	0.248	0.181	0.364	0.224	0.281	0.141	0.238	0.248	0.353	0.215	0.260	0.122	0.221	0.210	0.301	0.467	0.502
	48	**0.159**	0.238	0.186	**0.235**	0.317	0.324	0.224	0.422	0.321	0.354	0.198	0.283	0.440	0.470	0.315	0.355	0.189	0.270	0.320	0.394	0.966	0.733

pc | Pervasive Computing Lab
浙江大学普适计算实验室

Thank you for your lisenting!

Presenter: Zongjiang Shang